MOVING BEYOND MIRRORING

– A SOCIAL AFFORDANCE MODEL OF SENSORIMOTOR INTEGRATION DURING ACTION PERCEPTION

by

MARIA BRINCKER

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2010

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy

Prof. Nickolas Pappas

Date

Chair of Examining Committee

Prof. Iakovos Vasiliou

Date

Executive Officer

Prof. Jesse Prinz

Prof. Nickolas Pappas

Prof. Barbara Montero

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

MOVING BEYOND MIRRORING – A SOCIAL AFFORDANCE MODEL OF SENSORIMOTOR INTEGRATION DURING ACTION PERCEPTION

by

Maria Brincker

Advisor: Jesse Prinz

The discovery of so-called ‘mirror neurons’ - found to respond both to own actions and the observation of similar actions performed by others - have been enormous influential in the cognitive sciences and beyond. Given the self-other symmetry these neurons have been hypothesized as underlying a ‘mirror mechanism’ that lets us share representations and thereby ground core social cognitive functions from intention understanding to linguistic abilities and empathy. I argue that mirror neurons are important for very different reasons. Rather than a symmetric ubiquitous or context-independent mechanism, I propose that these neurons are part of broader sensorimotor circuits, which help us navigate and predict the social affordance space that we meet others in.

To develop both the critical and positive project I analyze the interpretive choices and the debate surrounding the mirror neuron research and show how the field is marred by highly questionable assumptions about respectively motor and social cognition. The discovery of mirror neurons - and the sensorimotor circuits of which these neurons are a part – actually empirically challenge many of these tacit assumptions. Findings of sensorimotor goal representations at levels of abstraction well beyond actual sensory information and kinetic movements challenge the idea of motor cognition as primarily output production. Additionally, the focus on 3rd person mindreading of hidden mental states is misguiding the field of social cognition. Much ‘mind-reading’ seems rooted in sensorimotor representations and a developmentally primary 2nd person understanding of actions and the mental lives of others, which precisely breaks the assumed dichotomy between mind and behavior.

I propose a *Social Affordance model* where parallel fronto-parietal sensorimotor circuits support representations not just of other people’s actions but of the overall social affordance space. It is a process that monitors concrete goals and teleological possibilities that the environment affords respectively oneself and other present agents. With this model I hypothesize that the complex spectrum of sensorimotor integrations are indeed essential not only to normal action choice calibration but also to social cognitive abilities, as the sensorimotor teleological representations let us *relate to* others and understand their action choices in a shared pragmatic and intentional context.

For the man and the kids

Desmond, Oona & Babette

ACKNOWLEDGEMENTS

Intellectual thought is always the product of many minds, long histories and the zeitgeist that engulfs us. This project is mainly inspired by the philosophy of Henri Bergson and has been cooking for quite some years. There are numerous close friends and family members that have heard more about mirror neurons and affordances than humanly digestible. Thank you for listening. But I particularly want to thank my advisor Jesse Prinz who both as a person and as a thinker in short time taught me invaluable lessons that I will draw on for a long time to come. Nick Pappas I thank for being my philosophical anchor and for his continuous support as he gently kicked me out of the nest. And many thanks to the rest of my committee: Barbara Montero, Massimo Pigliucci and David Rosenthal - thank you all for giving me chunks of your time and minds in this hectic summer of 2010. Thanks also to the pillar of the philosophy department Rosemarie Iannuzzi and to Rossen Ventzislavov for the urgently needed editing and cosmetic surgery in the 11th hour. Elly Vintiadis, Vic Crome, Keota Fields, James Snyder and Ellen Fridland thank you for all the support and discussions throughout the years. Lastly, I want to thank Esther O. Pedersen and my man Desmond Hall who taught me to be ambitious and to work hard and my parents and brother who had to deal with me before I learned any of these things.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Chapter 1

Introduction and Stage-setting

1.1. Introduction

This is a reinterpretation of the neuroscientific findings of so-called mirror neurons – a group of sensorimotor neurons responding to the execution and social perception of certain goal-directed actions. I will deal extensively with the concrete data and theoretical choices of empirical scientists, and the discussions are meant to help guide and theoretically shape the experimental work in this field going forward. But it is a philosophical project. It will speak to a series of questions at the heart of philosophy of mind such as the issues of other minds, of mental representation and how to understand intentional action choice and indirectly also point to new angles for understanding free will and the mind-body problem. The aim is methodologically to use the empirical findings and theoretical problems in the field of mirror neuron research to anchor a reformulation of many of these core philosophical issues. Thus, rather than attempting to argue for new answers to old questions, the aim is to present the beginnings of a new positive framework in the philosophy of mind, a framework in which the old issues are reshuffled to call for new answers beyond the traditional dualist and cognitivist dichotomies.

The project is therefore one that operates at multiple levels and works with several sub-questions simultaneously. One way of delineating these various strands of arguments and sub-projects running through the central chapters is the following:

1. **Empirical findings:** The discussion of and engagements with actual mirror neuron research, hereunder the data but also the experimental paradigms used and interpretive choices made.
2. **Empirical-theoretical tensions**: Analyses of underlying empirical-theoretical tensions in this research between the findings and the various applied and implied theoretical assumptions about the nature of the mind and cognition. Hereunder primarily the assumptions made regarding respectively motor and social cognition, but also more generally assumptions about ‘information-processing’ and mental

content as something that be reified, and transferred as inputs and outputs for various brain regions and neural populations.

3. **Motor cognition as output system:** The discussion of the traditional framework of motor cognition as basically an output system for the formation of motor commands – and by implication the traditional dichotomy between perceptual and motor systems and what lies in between. Hereunder important elements are also the discussion of the idea of motor simulation and the extent to which action intentions are chosen prior to the engagement of the motor system.
4. **Social cognition as 3rd person mindreading**: The traditional framework of social cognition as primarily an ability to 'mindread' and understand the hidden mental states of others. The discussion highlights the way this view of social cognition involves a dichotomy between observable behavior and mental representations as notoriously private thoughts sandwiched between perception and action.
5. **Affordances and the broader sensorimotor challenge:** The attempt to show how the broader sensorimotor research - of which mirror neurons are just one element - exactly challenges these above mentioned traditional frameworks of motor and social cognition. The notions of affordances and teleological goal representations are key to this project, as these cannot easily be fitted in the classical schema of perceptual input –central cognition – motor output.
6. **The social affordance model:** The presentation of my positive proposal for a 'social affordance' reinterpretation of these sensorimotor findings via all the above-mentioned elements. The proposal is that in fronto-parietal circuits we monitor and anticipate the shared social and physical space of potential action – beyond our momentary perception. It results in a relational, and inherently both informative and normative, affordance space understanding, which at the same time monitors the relative present and anticipates future directed action possibilities via multiple parallel schematic re-presentations of past sensorimotor engagements. In regard to social perception, such a model suggests that it is important to differentiate between 3rd person passive observation and 2nd person social interaction. Further, it stresses the importance of understanding not just the intentions behind actual actions of others but also their potential actions and thereby their personal process of actions choice and navigation.
7. **Philosophical implications:** Lastly there are pointers to the more general theoretical and philosophical conclusions falling out of the social affordance hypothesis and how the model reconfigures traditional categories of motor and social cognition and their accompanying philosophical issues of other minds, intentional action, mind-body distinctions, intentional action etc.

This differentiation is meant as a guide to the reader as I acknowledge the strain of keeping track of these sub-projects as they intertwine in the individual chapters. The discussion will not take the form of a traditional linear argument from a set of premises to a set of conclusions, but the force of the argument lies in showing empirical-theoretical tensions and then attempting to present and argue for re-mappings that hopefully can show the fly *beyond* the fly bottle[1].

1.2. A word on the methodology

But why attempt to do all these things simultaneously? Some philosophers might ask why one would write a philosophy dissertation about some obscure set of neurons. Why engage in a tedious discussion of details of empirical research and the debate it has spawned? From a distance the project might look like a piece of scientific journalism or history of science. However, it is important to note that my objective is not so much to uncover history but instead to empirically push theoretical change. My project is an example of what one might call philosophy *in* science. I am thus aware of the relative methodological novelty of an empirically based dissertation project like the present, but I stress that my philosophical objective is to deal with old philosophical issues. The methodology is meant to make it possible to significantly reformulate cognitive models and central issues of the philosophy of mind. One might ask why I need this methodology. Why not follow the traditional philosophical route and start by presenting existing theories and present my theory in contrast to these, and then roll in a few empirical references as guest appearances to prove my point? The answer is that this model of 'philosophy plus science' often seems to get bogged down by preconceived theories and sometimes forget both that theories are empirically based and that science cannot be conducted without theory. It seems to me that if one is to fundamentally change the received theories a good road forward is to get one's hands dirty with the science in all its minute experimental and interpretative details and start deconstructing and reconstructing theories in the midst of the empirical morass. The additional advantage is that the theoretical product is not simply given some empirical thumbs up, but is actually inherently shaped by the research – and, most importantly, thereby ready to get used to navigate and conquer new empirical seas. And that is my objective in regard to empirical psychology: to reinterpret the data, provide an overview and a broader framework for

[1] Wittgenstein writes in the *Philosopical Investigations* §309 "What is your aim in philosophy? - To shew the fly the way out of the fly-bottle." (Wittgenstein, 1953) I think that the role of philosophy as it engages with the world beyond itself can be greater.

understanding not only the role of mirror neurons in social cognition but more generally the role of sensorimotor integration in processes of intentional action, perception and higher cognition.

1.3. <u>The 'post-cognitivist' challenge</u>

What theory is it then that I want to change? In the theoretical landscape of cognitive science and philosophy of mind my thesis falls within what one might call the 'post-cognitivist' tradition, i.e. the attempt to move beyond the traditional frameworks of cognitive science. Although the roots of cognitive science are diverse and run deep in the history of philosophy – particularly rooted in Descartes' dualism - one can arguably see Chomsky (1966 & 1986) and Fodor (1975 &1983) as prototypical proponents of theoretical models that I would label 'traditional cognitive science' or 'cognitivism,' as they vehemently support the dichotomies that I sketch below. These philosophers have understood cognition primarily in terms of information processing of symbolic mental representation, central cognitive faculties and modular and encapsulated input and output systems and generally focus primarily on computational rather than biological processes. But be this history as it may, the aim of the current project is *not* to deal with any *explicit* defenders of cognitivism but rather to challenge certain more *implicit* and imported cognitivist leftovers in contemporary theories and research that in many ways aim to move beyond the traditional frameworks and their most vehement defenders.
The particular focus of this current project is to reinterpret the existing frameworks of motor and of social cognition, which dominate most debates of these and related areas in philosophy of mind and also the empirical research in the mind and brain sciences. I am interested in questioning the default assumptions underlying these frameworks, but also in showing how these assumptions are often imported implicitly and not explicitly discussed - much less empirically tested. The core assumptions that I shall critically expose and challenge in regard to motor and social cognition are:

- Serial information processing in input-output modules, and
- Dualism between an inner mind and outer observable behavior.

In regard to the assumption of input-output modules, I shall mainly focus on the motor system and the idea that it is unified and limited to producing outputs – be it covertly or overtly. I point to the influential but empirically unwarranted tacit notion of covert motor cognition as a full action simulation running off-line. I see this as an internalized version of behaviorism, which has resulted in a too limited understanding of motor processes and their functional roles. Furthermore, the notions of hidden and atomistic mental states have long ruled philosophical and scientific debates, but the present project aims to challenge the automatic mystification of the minds of others along with the internalization of representations and the externalizations of actions.

These points of discontent are, as I see it, linked to a broader dissatisfaction with some of the traditional frameworks of cognitive science and philosophy of mind. I shall for the sake of simplicity cluster various traditional assumptions of traditional cognitive models of the mind and mental processing under the heading of 'cognitivism.' At the center of this umbrella concept as I will use it, is the metaphor of mental processes as 'information processing' and the accompanying dichotomies of respectively:

- Inner mental states *versus* observable behavior
- Perceptual input *versus* motor output systems
- Vehicle/mechanisms/hardware *versus* mental 'thing-like' content
- Modular fixed functions *versus* unspecified general learning mechanisms

These dichotomies are interrelated in a picture of mind, intentional action and social interaction that Susan Hurley has perfectly labeled the 'classical sandwich.'[2] Prominent early philosophical critics of this picture include Nietzsche, Bergson, James, Merleau-Ponty, Heidegger, Dewey and Wittgenstein, and the revival of these thinkers in the context of cognitive science along with ground-breaking empirical research has over the last decades been spurring a cluster of counter movements that can be put under the heading of 'post-cognitivism.'[3] I share many of the critical sentiments in regard to the traditional picture – the question that I want to focus on, however, is how we reach a new positive and scientifically productive post-cognitivist model. The point is that the traditional cognitivist models and assumptions seem to stick with us, possibly mostly due

[2] Hurley (1998)

[3] For early philosophical critiques see Nietzsche (1882 & 1887), James (1890), Bergson (1896), Merleau-Ponty (1942 & 1945), Heidegger (1927), Dewey (1929) and Wittgenstein (1953), for early contributions and formulations of embodied approaches to cognitive science and the revival of these earlier ideas as a response to classical cognitivist frameworks see for example: Varela et al. (1991), Lakoff & Johnson (1999), Clark (1997), Hurley (1998), Dreyfus (1972/92), Brooks (1994), Thelen & Smith (1994). I am not certain of the origin of the term 'post-cognitivism' but it has been in use at least for a decade. See for example: Potter, J. (2000).

to the lack of an operative alternative. One way to move towards a new understanding of the mind and our knowledge of the minds of others is to probe into how it might be otherwise than these dichotomies suggest.

For example, given a cognitivist machinery of serially connected and encapsulated modules, in which perception and action production functions like input and output systems, mental representations that are not concrete percepts or actions are understood as distinct from and 'sandwiched' in between action and perception. Hence abstract representations are thought of as notoriously hidden and underdetermined by observable behavioral engagements[4]. One response to this view has been 'anti-representationalism,' the view that such 'sandwiched' pictures in the mind do not exist and are not necessary to explain intelligent behavior.[5] But one might try to retain some sort of notion of mental representations as analyzable well beyond stimulus and response, and as transcending the *present* observable engagements of people – without assuming that such cognitive processes or representations are in principle distinct from overt engagements in general.[6] Such a view of representation is what I shall attempt to argue for by the power of example in this thesis. But for now what is of importance is that, given the traditional picture of the mind and the above-mentioned neat dichotomies, the mere *possibility* of such an understanding of mental processes under a temporal or dynamic perspective is lost.

As mentioned many other philosophers, psychologists and biologists have ventured out on a similar 'post-cognitivist' endeavor over the last decades in particular, and have taken on classical cognitive models of mental functioning by way of emphasizing the problems of these frameworks as they ignore the way cognitive processes seem to be embodied, situated, enactive and biological.[7] I share many of the concerns of such approaches, and have learned from much of the poignant insight that these criticisms of cognitivist models

---

[4] As already mentioned Susan Hurley came up with the perfect label "the classical sandwich" for the traditional cognitivist input-cognition via mental representation-output framework (Hurley 1998). Here the suggestion is that this classical idea of cognition also is reflected in the philosophical 'problem of other minds' and the typical idea of social cognition as a process by which we attempt to understand the hidden mental states of others.

[5] See for example Freeman & Skarda (1990) and Brooks (1991).

[6] Bergson argues that the mark of the mental is the ability to act intentionally – and further that this ability has it basis in a temporality that reaches beyond the present to form hypothetical representations of the future via the past. (This is the central theme of Bergson's major work in philosophical psychology: Matter & Memory (Bergson, 1996)) I shall return to this idea of the mental in the conclusion.

[7] For a recent overview see example: Wallace et al (2007).

have produced. But as general frameworks these new approaches often take a particular failing aspect of cognitivism as their starting point. The labels of embodied, enactive and situated cognition suggest that these are contributing a particular essential element to the theory of cognition that the old framework left out. This is of course misleading as a more radical interpretation of the frameworks is strived for. But it has repeatedly proven tricky to tease out the *positive* alternative as a full theory of cognitive function. There is a temptation to formulate new positive theories in direct contrast to the old one. One might simply deny the existence of problematic theoretical posits - like the idea of symbolic representations sandwiched between inputs and outputs – and position oneself as an 'anti-representationalist.' But then one might be forgetting why this idea of representations was so cherished as a response to behaviorism. In other words the new theory should preferably keep the cognitive territory already won.[8] In response to this challenge of reinventing a radically new post-cognitivist framework, many have emphasized the historical insights of phenomenological and pragmatist thought, and tried to analyze those core aspects of our mental life anew, which seemed to have lost their life at the hands of information processing models. I agree with the bulk of the ideas that has come out of these projects and find that the work in these fields has been pivotal for the theoretical changes occurring in philosophy and cognitive science today.[9]

The problem is that much of the *actual research* and *scientific debates* out there are proceeding as if the traditional cognitivist frameworks were given as irrevocable facts. However, incoherencies and what one might call 'post-cognitivist findings' are popping up everywhere in the mind and brain sciences, and it is in the middle of this puddle of theoretical-empirical tension that the mirror neurons come into the picture. I shall focus my dissertation on these little neurological power players as I think they provide if not a window to the soul then a peephole into the need for radically new and better theories of motor and social cognition than those theoretical frameworks which paradoxically have brought these neurons to such fame in the first place. I see the mirror neuron research as presenting a unique philosophical opportunity, as the *theoretical* debate over their function is swamped with good old-fashioned cognitive science – but the *phenomena*

[8] For an interesting discussion of the representationalist- anti-representationalist debate and its inherent problems see Haselanger et al. 2003.
[9] For a good entry point into this literature see Gallagher & Zahavi 2008.

themselves carry the promise of a radically new framework for understanding the sensorimotor grounding of cognition, intentional action and social interaction.

1.4. Mirror neurons in the theoretical Petri dish

In the mid-nineties a group of scientists from the University of Parma discovered an intriguing kind of sensorimotor neurons in the pre-motor cortex of macaque monkeys that has revamped the study and theories of social cognition dramatically.[10]

These so called 'mirror neurons' seem to have a rather similar response to the observation and the execution of actions, and due to this *symmetric* overlapping response most researchers in the field have largely agreed that these neurons instantiate an observation-execution mirror matching mechanism. The standard story is that the observation of others' actions activates mirror neurons coding for the execution of those very actions, thus producing in the observer a sort of covert imitation or 'simulation' of the observed action, and that this simulation in turn yields a broader embodied or experiential understanding of various aspects of the action and its goal, consequences and motivations. In later studies mirror neurons were also found in the inferior parietal cortex and fMRI and behavioral findings have provided evidence for a homologue parietal-prefrontal 'mirror neuron system' in humans.[11] These findings, along with bold interpretations of the possible functions of such a mirror mechanism, caused an enormous stir of excitement throughout the world of cognitive neuroscience and beyond. The 'mechanism' of mirror neurons was theorized as automatically and ubiquitously producing an agent-neutral representation of the action/emotion observed, or in other words as producing a 'shared representation' or maybe even a 1st person experience in the observer of the action or emotion expressed by the other.[12]

Thus, the newly found mirroring mechanism was proposed as a neurological basis for action understanding and recognition, imitation, motor learning and possibly a foundation for other sophisticated social cognitive abilities such as understanding of others' intentions and goals. Furthermore, mirror neurons where instantly linked to human

---

[10] For early findings of mirror neurons in area F5 of monkeys see di Pelligrino et al. (1992), Gallese et al. (1996) & Rizzolatti et al. (1996).

[11] For early indirect findings see Fadiga et al. (1995) and recent human single cell findings see Mukamel et al. (2010).

[12] Gallese & Goldman (2001).

linguistic abilities and the evolution of such since the human homologue of the premotor area in which they were first found contains Broca's area, which is an area notoriously important for language in particular language production and syntactical understanding. Recently mirror neurons (broadly defined) have been found or implicated in other areas as well, such as in the primary motor cortex, the somatosensory cortex and in medial areas of the prefrontal cortex, and also in areas having to do with emotions and pain perception such as the cingulate cortex.[13] The latter findings have radically expanded the application and nature of the mirror mechanism hypothesis. In regard to emotions it has been suggested that we mirror the emotion that we perceive the other as having/expressing and that this mirroring process helps us understand and recognize others' feelings and further provides a basis for empathy. Clinically strong arguments have been advanced that many of the social cognitive problems of people with autism can be traced to a 'broken mirror' i.e. a dysfunction of the mirror neuron system. In short, the findings of mirror neurons and the hypothesis of a mirroring mechanism have had enormous ramifications throughout the field of social cognition and the cognitive sciences at large. The famous and vocal neuroscientist Ramachandran has even gone as far as to suggest that this discovery will do for psychology what DNA did for biology.[14] A basic web search will give an idea of the astonishing hype, attention and excitement that these neurons have garnered in the popular literature and media far beyond the traditional realm of science and philosophy.

1.4.1. A motor mechanism for shared representations?

But what is it about mirror neurons, which is so extraordinary that it is thought to fundamentally reorganize the way we understand cognition and psychology? Most people take the central tenet of the mirroring theory and also its great promise to be in the creation - by way of automatic covert mirroring – of a *shared representation* between people. The promise is in the idea that we, via our 1st person motor and possibly emotional resonance, somehow have a sort of pretense *access* to the minds of others, in that we simulate their experience or action (almost) 'as if' it was our own.

---

[13] See for example Gallese et al. (2004).
[14] See Ramachandran's essay from June 2000, "Mirror neurons and imitation learning as the driving force behind "the great leap forward" in human evolution" published online by Edge at http://www.edge.org/3rd_culture/ramachandran/ramachandran_p1.html

While various authors have come to doubt some of the suggested social functions and the exact process of this mirror matching mechanism during action observation, few challenge the idea of shared action representation or the observation/execution mirror mechanism as such. However, it is the theorized mirror mechanisms that are the central aim of my criticism. The main point is that the mirror terminology and argumentation assumes a somewhat *modular agent neutral* and *ubiquitous action simulation mechanism*, and that is an empirically problematic interpretation of the functional processes in question. I shall argue that the action observation/execution matching mechanism is a hypothesis that to some extent is created and sustained by conducting and interpreting research with blinders on, by ignoring the details of the experimental and neurological context and theorizing the functional role in isolation from a broader theory of sensorimotor integration and intentional action. Further, I shall argue that the theory relies on problematic assumptions about motor cognition as being solely about producing action outputs, be it overtly executed or covertly simulated, and about social cognition as mainly 3rd person mind-reading.

My critique is not aiming to deny the importance of the discovery of mirror neurons, but rather to suggest that they are important for rather different reasons than normally presumed, and that their fame ought to be shared with a much broader set of sensorimotor circuits. What I want to criticize is the suggested 'mechanism' and more particularly the implicit theoretical frameworks of motor and social cognition that it rests on. I reinterpret the functional role of mirror neurons in a broader sensorimotor context. But most importantly I use this resulting positive sensorimotor model not only to criticize the traditional information processing frameworks that repeatedly have haunted the mirror neuron interpretations, but also to make room for alternative frameworks of mental representation, intentional action and knowledge of other minds.

1.4.2. The larger affordance space of mirror neurons

I argue that the metaphor of mirror neurons as little modular mirrors that ubiquitously provide an agent-neutral action representation independently of the task and social context is empirically unfounded – but enormously influential in both the mirror neuron research and its wider interpretations. As the metaphor is exported from individual neurons to a broader ‘mirror mechanism’ it seems even more empirically implausible. Why should a heterogeneous group of action sensitive visuomotor neurons in a fronto-parietal circuit be understood as having a specific social epistemological function during action observation, which is independent from the functioning of the rest of the neurons in this circuit?

I shall argue on the contrary that the empirical findings make sense if one understands mirror neurons as playing a role within these broader sensorimotor circuits. More precisely I suggest that these circuits play an important role in monitoring and anticipating the larger affordance space we are in, i.e. in helping us understand what goal directed actions are invited by what and to whom in a concrete social and physical scenario. Thus some mirror neurons might be representing the actions and intentions of others, but most often *not* in an agent-neutral or context independent way. The central function is exactly to represent and predict the actions of others by way of their relations to the broader space and for the purpose of one’s own action choice and planning. The broad spectrum of visuomotor neurons in the premotor-parietal circuits thus plays an enormously important role in our pragmatic understanding of others as their actions pertain to our shared space and as they let us predict and prepare our actions in relation to this social space.

Hence, I might in certain respects want to tone down the mirror neuron fame and fever, but at the same time spread some of the enthusiasm bestowed on these particular neurons to other sensorimotor neurons and processes as well. Accordingly, it is not that I want to say that mirror neurons play no role in social cognition or that motor areas are not essential for perception. On the contrary, I simply say they play a different social role. Instead of focusing on agent neutral action simulations and their role in 3rd person ‘mind-reading,’ I hope to underline on the one hand the *asymmetric* and on the other the *teleological* and *relational* aspects of sensorimotor processes. But I argue that precisely such goal representations and relational processes can serve important roles in understanding not only goal directed actions but social relatedness and the intentional action choices of others. Further I suggest that these sensorimotor goal and affordance representations might provide us with an example that holds promise for a broader post-cognitivist theory of mental representation.

The key to my reinterpreting of mirror neurons is that I argue for a different notion of social and motor cognition. The central role of such sensorimotor integrations in many cognitive processes points to the inadequacy of the traditional neatly divided dualisms of action-perception, mechanism-content, inner-outer and 1^{st}-3^{rd} person.

I thus hope with this discussion of mirror neuron research to showcase the residual cognitivist view of the motor system and of social cognition that is tacitly assumed by most mirror neuron research. Hereunder, I want to question not only the dualistic distinctions mentioned above but also the part/whole modular and sequential compositionality of the information processing frameworks, which I see as in many ways stifling and misdirecting even some of the most inventive and inspired psychological and neuroscientific research.

1.5. Caricatures and empirical tensions - The devil is in the detail

Though often discussed in the abstract, the ‘mirroring’ or ‘action simulation’ theory is born out of and still tightly connected to the neurological discovery and studies of activity of mirror neurons and the areas that contain such neuron. And the cornucopia of empirical findings in the field paints a much more complex picture than the impression one gets from the standard mirroring story. Even the neuroscientists that first discovered, named and interpreted mirror neurons as instantiating a mirror matching mechanism have made many adjustments and corrections to their theories. Most mirror neurons research shows clear tensions in regard to and deviations from the idea of mirroring during action perception as ubiquitous agent neutral covert action simulation. As we shall see, Gallese and Rizzolatti and the other main spokespeople of mirror neuron research might rightfully say that the simple mirroring story is a caricature of what really goes on, and their own discussions often show that their view is both more complex and more reasonable than the mirroring metaphor suggests. Nonetheless, they never really explicitly break with the basic mirror mechanism idea, but often times fall right back in both this terminology and the many tacit implications of this framework. Thus, it is this caricature story that drives most of the debates and the research in the field and both proponents and opponents seem to repeatedly lean on the simple mirror matching metaphors. It therefore seems to be important to analyze empirical shortcomings as well as systematically misleading elements of this traditional mirroring story to pave the road for a more radical reinterpretation and reframing of the research.

In Chapters 2 & 3 I shall summarize some of the key mirror neuron findings along with the theoretical interpretations intertwined with these empirical studies. My main focus will be to analyze the interpretations of ‘action’ mirror neurons and the ‘mirror neurons system’ findings spanning some tightly connected areas of the inferior parietal lobe and the lateral pre-motor cortex. Action mirror neurons have received the most research attention and inspired the mirroring theory, but I choose to focus on action mirroring also because I want to reconceptualize motor cognition and the functional role of sensorimotor integration in both perception and action via the uniting concept of affordances. Chapter 4 traces the further theoretical development and drift of the mirror neuron interpretations as these are linked to Alvin Goldman’s simulation theory of mind-reading. The style of giving an interconnected presentation of the empirical studies, their interpretations and their implied theoretical frameworks might at times feel cumbersome and too detailed for the purposes of a philosophical discussion, but I see these details as essential to my project on multiple levels. As I shall discuss in Chapter 5 other philosophers and scientists such as Csibra, Jeannerod and Jacob have taken on the project of critiquing and reinterpreting mirroring stories suggesting inconsistencies between a caricature mirroring story and various empirical findings. In spite of sharp analyses and some focal insights, I shall argue that their critique in many ways represent a theoretical step back towards the individualistic input-output framework of cognitivism, and that they by a narrow focus on the standard theory overlook many of the more subtle and insightful intuitions of mirror neuron researchers like Rizzolatti, Gallese, Iacoboni and others. These neuroscientists are in many ways trying to break away from the input-output model of the mind and see their research on sensorimotor integration in mirror neurons as an essential move away from the traditional picture. Their research provides many important clues towards a ‘post-cognitivist’ model of the mind and brain, but at the same time I think that their mirror neuron story carries with it a number of unfortunate assumptions that fits squarely within the traditional model of a unified motor system and of cognitive processes as serially connected encapsulated mechanisms and of social cognition as 3rd person mind-reading.

In short, while the sky seems to have been the limit for the many hypothesized functions of mirror neurons narrowly, it seems that the broader neurological and theoretical implications and interpretations of the new findings have stopped in their tracks. If one wants, more fundamentally, to let the new mirror neuron data inform an improved model of brain function, I think it is essential to understand the scientific and *theoretical context* that the mirroring theory grew out of. To this effect the inherent tensions that from the beginning existed between the complexities of the phenomena researched and the standard mirroring story present extremely useful pin-pointers to the issues and concepts that need reexamination.

I thus want to show how while on the one hand mirror neuron theories take a big step away from traditional cognitivist models of brain functioning they still in many ways retain many of the assumptions of the traditional framework. To make this point I will discuss the conceptual and interpretative choices made in the course of actual research. Simultaneously the discussion of the research will also set the stage for my own theoretical reinterpretation of the mirror neuron findings, which I will continuously develop throughout all chapters.

I take my own affordance space theory and reinterpretation of motor and social cognition as representing a further - but of course still undeniably incomplete - step away from the traditional cognitivist model. To justify my alternative account I will again point to the many tensions that continuously appear in the research between the empirical phenomena and the standard mirroring story, and also how many authors of mirror neuron articles repeatedly show that their intuitions go well beyond the strict mirroring story they cling to in the conclusion section. So the many quantitative and experimental details might seem slightly superfluous to philosophers' ears and the terminological analyses pedantically nitpicky to scientists. However, the concluding Chapter 6 is meant to show how all these details carry the core argumentative ammunition of the dissertation. Accordingly, this final chapter stands as making explicit and arguing for what we can all more or less see for ourselves in the fine grains of the phenomena and further indicate the many larger implications that this story has for core issues in the philosophy of mind, and hopefully for the empirical study of fronto-parietal sensorimotor integrations, social cognition and beyond.

Chapter 2

Mirror neurons and the sensorimotor backdrop of the discovery

2.1. Introduction

As mentioned earlier, mirror neurons were first discovered - and also named and interpreted as instantiating a 'mirror mechanism' - by a group of neuroscientists at the University of Parma[15]. The group, lead by Giacomo Rizzolatti, was already involved in ground-breaking research in the area of the organization and function of the primate pre-motor cortex.[16] And it was while they were pursuing a single cell study[17] of complex premotor neurons that they accidentally discovered mirror neurons, i.e. individual neurons that responded not only to the execution of particular goal-directed actions but also to the observation of others performing similar actions.[18] As I shall be underlining in the following the group was already aware of various intriguing both visual and motor properties[19] in many pre-motor neurons before this new discovery. But that motor neurons in area F5 of the premotor cortex would respond to the *perception of actions* performed by *others* was indeed something new. And indeed something so fascinating

---

[15] Given the early publications the key figures in the Parma group of the 90's were Giacomo Rizzolatti, Vittorio Gallese, Leonardo Fogassi and Luciano Fadiga, Giuseppi Luppino and Maurizio Gentilucci. For reasons of simplicity I shall generally refer to the research and theories published by these and later Parma University neuroscientists under the label the 'Parma group' or 'Parma lab.'

[16] See Rizzolatti et al. (1987), Rizzolatti & Gentilucci (1988), Gentilucci et al. (1988), Rizzolatti et al. (1988).

[17] The properties of individual or small groups of neurons are investigated via inserting small electrodes into the cortex. This is an invasive procedure and as I shall discuss later a method with many ethical limitations, and cannot be used in humans - or chimps for that matter - without a medical rationale.

[18] The story goes that mirror neurons were first discovered by accident while the Parma group was doing single cell recordings of grasp related pre-motor neurons in a macaque monkey. I have heard various anecdotal reports at various talks of how during a recording session a passing experimenter grasped a peanut from in front of the monkey and the action potential of the grasp related cell started to fire vigorously. An additional episode involving a graduate student and an ice cream cone is also reported by Blakeslee in an NY Times article about Rizzolatti's mirror neuron findings (Blakeslee, S. (2006) "Cells that read minds", *The New York Times*, Jan 10th 2006).

[19] I shall as is typical in the neuroscientific literature refer to neurons as having visual and motor properties when their activity has been found to be modulated in correlation with respectively visual and motor episodes at the level of the organism at large. As I shall discuss later in regards to the terminology of sensorimotor neurons, the idea of such correlations as 'properties' of individual neurons might be somewhat misleading.

that the Parma group deemed it was worthy of a new unique name and a new unique independent neural mechanism with an essential social cognitive function. The name as we know became 'mirror neurons' and the hypothesized cognitive mechanism instantiated by these neurons was thought of as an 'observation-execution matching' or 'mirroring' mechanism. The main proposed function of mirror neurons was also based on the symmetrical idea of self-other mirroring, namely some kind of understanding and recognition of the action and goal of the other based on matching the observed action with one's own motor repertoire. The Parma group situated their new findings of action perception sensitive motor neurons in relation to their existing theory of F5 as containing what they had called a 'motor vocabulary.' But as we shall see in the following the new name and new hypothesized mechanism still signaled a clear break from the broader premotor research that gave birth to the discovery.

The central function of the lateral premotor areas has generally been thought to pertain to planning, hierarchically organizing and initiating perceptually guided actions. In short the function of these motor areas has been related to the agents' own action planning and coordination and execution. I shall argue that the narrow research focus on 'mirroring properties' assumes an isolated function of these neurons in *perceptual processes* and in regards to the *understanding* of perceived *others* that largely overshadows possible elements of continuity between the new celebrated mirror neurons and other visuomotor and sensorimotor neurons, which are thought to relate to one's own action performance.[20] Thus, dichotomies are made between sensorimotor functions *for action* versus *for perception/understanding* and equally between motor cognition for *own* pragmatic purposes versus for perception and understanding of *others*. It is a strong feature of the cognitivist 'input-cognition-output' model that perceptual and central processes are thought to serve epistemological/information gathering and reasoning processes whereas motor processes are thought to produce action outputs. Many contemporary more 'embodied' and 'enactive' theories emphasize the role of motor cognition in various other cognitive tasks and the mirroring theory is mostly seen as contributing to these moves away from this 'classical sandwich' model of cognition.[21] However, what I want to point

[20] A less symmetric and unique terminology and mechanism might have been conducive to more empirically plausible theories and more contextually sensitive research - but ironically such research would surely also have been less head-turning and less influential.
[21] The 'classical sandwich' is Hurley's notion introduced in Chapter 1.

out is that problematic distinctions often are maintained at the level of *function*, and that this is particularly obvious in the theory of a *motor* mirror mechanism *for* social action *perception*. As I see it, perceptual and motor processes are not truly reinterpreted in such a theory. Instead, a motor process is simply isolated and then sent to 'moonlight' for perceptual purposes. I argue that such modular and non-pragmatic *functioning* of sensorimotor processes is as empirically implausible as the sharp *anatomical* or *temporal* distinctions between perceptual and motor systems as unconnected peripheral input and output systems.[22]

To show the neurological sensorimotor context from which the mirror mechanism got extracted, I will before turning to the actual mirror neuron studies take a closer look at the background of the discovery and the Parma group's earlier research on the physiology of the premotor cortex and F5 in particular. I think these early findings and their interpretations are highly informative in regard to the problems of the later conceptualization of mirror neurons and their cognitive role as a mirror matching mechanism, but also in regard to understanding the broader theoretical framework changes and the Parma group's attempt to contribute to new post-cognitivist models of sensorimotor goal representation and brain function more generally.

2.2. Backdrop: Questioning the modular idea of a motor 'output' system

In the 80's the prominent conception of the cortical motor system was that it was a rather unified output system, and its functions were generally restricted to planning, organization and execution of movement commands for overt movement. In accordance with the famous findings from Penfield's 1930's stimulation studies, the cortical motor system was anatomically located just in front of the central sulcus and grossly thought to consist of 2 areas: primary motor cortex and pre-motor cortex each equipped with a full somatotopical homunculi representation of the moving body parts.[23] The underlying logic of the motor system theory was that via these body representations it could orchestrate the movements of the body, a story that fits neatly with the traditional cognitivist idea of

[22] The argument is thus that sensorimotor findings that I shall discuss underline not only the empirical implausibility of input and output systems serving only input and output functions, but more fundamentally challenges the modular reification and isolation of such cognitive mechanisms or processes.
[23] See Penfield & Rasmussen (1952) and for a typical representation of the Penfield inspired motor homunculi.

the motor system as a peripheral output system. One of the key theoretical projects of the Parma group in the 80's and 90's was the attempt to do away with this traditional conception of the motor system, and in particular the idea of Broadman's area 6 of the pre-motor cortex as a relatively anatomically uniform area with a rather uniform functional role in action planning and execution. As opposed to this traditional picture they had found by way of research on macaque monkeys that there were very interesting differences in anatomy, connectivity as well as functional properties between the medial areas now labeled F2 & F3 and the areas F4 and F5 in the lateral part of premotor cortex (Rizzolatti & Gentilucci, 1988).[24]

This project of reinterpreting the organization of the cortical motor system might seem to be simply an issue of anatomical labeling, but it represents a significant split with the overall model of the mind as an input-cognition-output system, because the idea of the motor system as an encapsulated and uniform movement output system has been challenged on multiple levels. Importantly they started to conceptualize the functional divisions of premotor areas not only according to effectors, i.e. around different body parts of an internally represented homunculus, like Penfield and colleagues thought, but rather as depending on more abstract categorizations of various *kinds of actions* and *goals*. Further, based on the systematic differences in sensorimotor connectivity within the large heterogeneous frontal motor area, they argued that it was *functionally* much less uniform than formerly thought and that many areas showed a level of cognitive complexity well beyond the kinetic movement commands, which traditionally had been seen as the only legitimate functional role of motor areas.[25] Thus, the new anatomical divisions of premotor areas posed a challenge to the traditional idea of body centered

[24] Most of the research done by the Parma lab is for ethical reasons done on Macaque monkeys rather than humans. The questions of how to compare and draw conclusions about human cognition from research done on macaques should not be down-played, especially when it comes to social cognition is it quite obvious that the differences between humans and new world monkeys are immense. However, on the issue the existence of general premotor and parietal circuits and sensorimotor integration it is clear that there on an anatomical level are quite strong parallels to the fronto-parietal integrations in humans. Again, I am not claiming to know to which extent the *local properties of individual neurons* or *scope of functions* of these areas are similar. Later I shall return to the issue of the plausible differences and the contemporary questions raised in regards to the lack of *direct* evidence of mirror neurons in humans. But for now I go along with the received view in the physiological and anatomical literature and assume that the over-all goal based sensori-motor and functional organization of the monkey premotor areas to a meaningful extent parallels that in humans.

[25] For obvious reasons I focus my discussion on the research of the Parma group but other labs for example that of Kiyoshi Kurata in Japan have similarly undermined the simple output system model of the motor system (see Kurata and Tanji 1986).

organization, but it also challenged the supposed uniform function of the motor system as simply putting a pre-specified action goal into motion, of simply producing coherent overt action sequences. The finding of many functionally specific connections between individual premotor and parietal areas underscored this challenge. The question was how the many loops and circuits between various premotor areas and parietal areas, which were traditionally thought of as multisensory association areas *outside* of the motor system, could be explained within the traditional cognitivist conception of the motor system as a unified output system? Further, how could one maintain the temporal idea that the motor output system only gets engaged *after* the full achievement of various relevant high-level perceptual and cognitive processes? It seems clear that the early research of the Parma group in many ways screamed for a radical reinterpretation of the motor system. In the following sections we will take a closer look at some of the empirical details that so emphasized the need for this theoretical move, and also some of the interpretations that might have caused some of the inertia.

2.2.1. The goal organization and 'motor vocabulary' of area F5

Rizzolatti and his coauthors reported in 1987 that many F5 premotor neurons are modulated primarily by the goal or target of actions rather than by the precise movements performed, the target location or the effectors used.[26] This shows that these premotor neurons operate at a level of abstraction removed from the kinetic properties of motor acts that had hitherto been thought essential not only to primary motor but also to premotor activity. They write:

> We became sure that goal was a necessary concept in order to understand this class of neurons…There is no way to interpret these observations in terms of contractions of single muscles or muscular groups…the concept of aim was indispensable for explaining also the activity of neurons exclusively related to arm movements (Rizzolatti et al. 1987).

[26] See Rizzolatti et al. (1987). Again, this research is based on single cell recordings in macaque monkeys and the human evidence is indirect. It is thus difficult to know what the exact properties are of such complex sensorimotor neurons in humans and what the differences are to macaque sensorimotor integration. But given the complex goal directed actions of humans it seem like a plausible *hypothesis* that if there is abstract goal representation in macaques' premotor areas then it is *not less abstract* in the homologue areas in humans. I will expound on this later.

They see the concept of action goal or aim as a necessary category to make sense of the neuronal activity even when this activity is linked to a specific effector like the arm or hand.

The specificity found in regard to the motor goal combined with the irrelevance of exact kinematics and location of the target lead the researchers to conceptualize the area as containing 'vocabulary' of motor acts.[27] Rizzolatti and his co-authors use this vocabulary terminology to draw an analogy between representations of goal-directed motor actions in F5 and 'words':

> This motor vocabulary is constituted of "words," each of which is represented by a population of F5 neurons. Some words code the general goal of an action (e.g. grasping, holding). Others code how, within a general goal, a specific action must be executed. These words select specific "motor prototypes" such as, for example, the configuration of fingers necessary for the precision grip. Finally, other words specify the temporal aspects of the action to be executed (e.g. opening of the hand). Thus, each action is represented by specific populations of neurons at different degrees of abstraction (Rizzolatti & Luppino 2001, p.891).[28]

This description of the different levels of generality and temporal specificity of the motor properties of F5 neurons corresponds also with the single cell findings that we shall see in regards to mirror neurons in this area. There clearly seems to be a division of labor between the various individual neurons in terms of which elements and levels of description of the action they are modulated by. To my mind, this division of labor not only in regard to action goals but also to the level of description and temporal segments raises a question about whether the analogy to words is more misleading than helpful. Given Rizzolatti and Luppino's analysis of the division of labor of the neuron populations it seems that a more complex metaphor than simply 'words' might have been more appropriate. In collaboration with Rizzolatti, Marc Jeannerod, a prominent cognitive neuroscientist in the field of motor cognition – of whom I shall talk much more later - uses the term 'motor schemas' rather than 'words.'[29] Jeannerod's terminology

[27] Furthermore, Rizzolatti et al. 1987 found that the activity of these premotor neurons could not be explained by motivational or attentional factors "since there was a strict relation between the neuronal discharge and specific motor acts."

[28] Rizzolatti & Luppino 2001, See also Fadiga et al. 2000, p.171 (section on 'Visuomotor neurons').

[29] "In F5 the schemas are represented by populations of neurons that code different motor acts. Various types of schemas can be distinguished. Some define general categories of action, for example grasp, hold and tear. Others indicate how the objects are to be grasped, for example held and torn. In this case each schema specifies the effectors that are appropriate for the action…Finally, a third group of schemas would be concerned with the temporal segmentation of the actions (the co-ordination of schemas). Thus, the motor

might do a better job at relating the specific differences in motor properties of F5 neurons to specific functions in action coordination and control processes. But either way they still end up with the idea that F5 forms a motor vocabulary in the sense of containing basic action schemas or ‘bits’ that can then be integrated and coordinated with larger contextually fitted action segments. However, it seems to me that it is very important to underline that the F5 neurons are not just *representing* actions as individual segment parts, which then can be activated by some sort of central intention. Rather their hierarchical and temporal diversity along with their sensory properties somehow *guides the process of coordinating appropriate actions for the present perceptual context*. Thus, the terminology of ‘words’ and ‘vocabularies’ is problematic because the non-hierarchical and atomistic ‘all or none’ idea of ‘words’ occludes the evidence inherent in the F5 neurons distributed response. If one wants an analogy to language, I think that a broader linguistic analogy might better capture the details of the heterogeneous evidence, i.e. suggesting that some neuron populations represent words, others phrases and some phonetic segments and that activity somehow gets syntactically and hierarchically organized in relation to the overall communicative goal. The analogy would then point to the dynamics needed to plan, choose and correctly execute hierarchically complex intentional actions in a way sensitive to both context and prior experience and skill. Simply a conceptual storage or representation of action ‘words’ does little to inform the many components of action choices and goals and neither does it tell us anything about the need for hierarchical and temporal orchestration. To sum up, my claim is that the specific distribution of F5 neuron responses shows a temporal and hierarchical spectrum and procedural glue that are downplayed by the ‘motor vocabulary’ interpretation which only underscores a link between *particular* goals at a somewhat *conceptual* level and various *kinetic* motor acts and commands.[30] Thus, the ‘vocabulary’ terminology leaves

---

schemas in F5 form a basic ‘motor vocabulary’ from which many dexterous movements can be constructed as coordinated-control programs.” (Jeannerod et al. 1995)

[30] One might want to note that the Parma group did choose to call it a ‘motor vocabulary’ and not a ‘motor dictionary’ as some later interpreters mistakenly have called it (e.g. Dinstein et al. 2008). I might not be a fan of the implications of seeing F5 neurons as action ‘words,’ but the term ‘vocabulary’ at least hints at a personal and pragmatic element. It is not a non-contextual ‘dictionary’ or a catalog of abstract action meanings or categorizations of action types, but rather it links goals to contexts and personal motor repertoires. It seems to code the performance of the actions ‘invited’ by the target at hand – given one’s spatial position and motor repertoire. Thus, in the pragmatic sense of vocabulary and expanding it to a more general personal set of hierarchical action performance tools, I might see how the F5 area could be seen as containing a hand-object action ‘vocabulary.’ But, of course, the question then is whether such a pragmatic

one with a rather misleading metaphor of individually represented action meaning and goals rather than a multilayered competence of goal-directed sensorimotor engagement.

2.2.2. Are goal organized neurons purely 'motor' neurons?

As mentioned it was the surprising finding of the abstract action goal dependency of F5 neurons combined with the independence of exact movement kinematics, which lead to the vocabulary terminology. Furthermore, the goal-level modulation findings seem to problematize the idea of the motor system as simply a system of action commands. The idea is that if individual neuron activity is correlated with the goals of action rather than specific movements and effectors then it might raise some questions as to what one takes to be a motor representation or a representation in motor 'format'. In other words, mental representations of actions in cognitivist frameworks are typically thought of in terms of specific motor commands, and the question now is what needs to be changed in this picture to accommodate premotor activity as organized around goals. Does this non-kinetic organization also question the idea of the motor system as simply an output system? I.e. the question is whether the idea of goal modulation can be combined with the traditional conception of a neat division between perception/input, higher cognition and action/output. Goal modulation was inferred from a *correlation* between the neuronal response and the goal outcome of the performed action, but of course this correlation leaves open the question of how the 'goal' modulation or representation is possible on a neuronal level. How, on a physiological level, does a motor goal *look* devoid of exact motor commands in a premotor system and how is such a 'goal' stored? Rizzolatti and his colleagues do not directly take on this question, but have repeatedly directly and indirectly mentioned the relationship between the specification of motor acts, the goal and in its concrete pragmatic context.

> Another important finding was that the majority of neurons related to grasping movements specify the type of grip necessary to take possession of the object. This finding indicates that in area 6 [inferior area: F4 & F5] the goal of a motor act is not coded in terms of abstract commands like "reach," "grasp," "hold" but in terms of

---

vocabulary makes sense in pure motor terms and whether it captures the massively parallel, hierarchical and dynamic sensorimotor activity of the areas in question. I argue that it is more misleading than helpful as a metaphor.

> the way in which this command can be implemented…(Rizzolatti et al. 1987, p. 224).

Thus, the action goals of the motor vocabulary seem to be related both to some kind of goal and to the concrete perceptual context, and somehow actions or acts at various levels of generality or specificity seem to be integrated with a perceptual or maybe imaginary representation of the target. I am not here going to solve the issue of how goal representations are instantiated in premotor areas, but it seems to me that some kind of 'extra motor' integration is necessary to make the goal-centered action organization work. William James suggested as part of his ideomotor principle that motor acts are initiated via a representation of their goals and learned sensory consequences.[31] More on that later, but here I just want to hint that *sensorimotor* integrations might be at the core of goal organized action representations.

Integrative sensorimotor goal coding is however not discussed by the Parma group as central to the way they conceptualize the organization of the F5 'motor vocabulary.' Even though they often write that the area is important for perceptually guided actions, on most conceptual descriptions the sensory properties merely seem to be a useful *accessory* for F5 rather than an essential *organizing* feature. Some important reasons for this view could be that most F5 neurons fire also when their target goal directed action is performed in the dark – i.e. without a visual perception of the target and, also, that some neurons were found not to be modulated by pure perceptual stimuli in the absence of movement. Of course the fact that some neurons fire in absence of *present* perceptual guidance does not guarantee that the organization of these neurons is *independent* of previously established sensorimotor integrations.

Whichever the reason behind the assumption, it seems to be assumed that premotor neurons code goal organized actions somewhat independently of their sensory connections. They are implicitly taken to be motor as opposed to perceptual or inherently sensorimotor representations.[32] I stress this feature of the F5 'motor vocabulary' theory,

---

[31] James (1890).

[32] I here use 'perceptual' rather than 'sensory' in contrast to purely motor properties, as it is not clear that premotor areas support what is traditionally thought of as basic sensory processes. These neurons are modulated in a way that corresponds to high-level perceptual features and meaning rather than sensory details. But as the terminology in the present literature that I write about is 'sensorimotor' rather 'perceptual-motor' neurons and accordingly sensory and motor properties of such neurons I shall at other times follow this traditional language use. However, I just want to stress that there is a rationale behind the terminological choice of perception and perceptual properties that does not have to do with conscious

as it is somewhat implicit and possibly unintentional but in my eyes highly problematic and influential in guiding further misleading theoretical conjectures. The problem is that the goal representation is assumed to be fundamentally anchored in some kind of motor action classification or representation independently of the various somatosensory or visual modulations that they also show. These sensory or perceptual properties are seen as serving important functional and pragmatic purposes, but they are not seen as *shaping* the 'vocabulary,' i.e. the action classifications or goal representations themselves. Thus motor and perceptual processes are still seen as independently organized 'systems,' and it is therefore clear that even though the 'motor vocabulary' theory reinterprets the motor organization in terms of goals, it still resists a more radical break away from the traditional model of the motor system as functioning in relative isolation. However, as I have tried to show here, this assumption might be unintentional and there is certainly an inherent tension between the findings of goal organization and the traditional motor system conception. The empirical findings of widespread fronto-parietal sensorimotor integration over the last decades by the Parma group and others add extra tension to the traditional picture.

2.2.3. Parallel sensorimotor integration in the fronto-parietal circuits

Well before they happened upon mirror neurons, the Parma lab was investigating the intriguing sensory properties of premotor neurons in various areas (Rizzolatti & Gentilucci 1988, Fogassi et al. 1992). It was to a large extent these various sensorimotor integrations and the mapping of their connections to specific parietal areas in particular that lead to their anatomical claim that the premotor cortex of the macaque, which had formerly simply been referred to as area 6, should be subdivided into a variety of physiologically and functionally distinct sub-areas. In a later review article Luppino and Rizzolatti write:

> Therefore, the parietofrontal connections form a series of largely segregated anatomical circuits. The functional correlate of this organization is that each of these circuits appears to be dedicated to a particular sensory-motor transformation, the essence of which is the transformation of a description of the stimuli in sensory terms into their description in motor terms. (Luppino & Rizzolatti 2000, p.220)

---

perceptual experience but the level of abstraction and 'top-down' influence that these sensorimotor processes seem to have.

Thus, on the one hand the connectivity is seen as essential to the functional properties, which suggests a break with the idea of narrowly self-contained cognitive functions of individual or small groups of neurons. However, on the other hand it is interesting to note the terminology of 'translations', 'transformations' and 'descriptions' in this and other quotes. This terminology suggests that sensory and motor representations are distinct and that there are meaningful distinctions to be made between sensory and motor systems as generating their respective representational formats. Here again one is invited to think of the idea of F5 as containing a 'motor vocabulary' where sensory information is thought to be *'mapped onto'* preexisting representations of goal organized actions, rather than as possibly *determining* the existence of such a goal organization in the first place. One can sense the tensions in the theoretical framework here as Luppino and Rizzolatti go on to point to the motor responses of parietal neurons. They conclude that "…if one defines a motor neuron as a neuron the activity of which correlates with action, there is no doubt that the posterior parietal cortex should be considered part of the motor system."(p. 220) Hence, they suggest that the parietal areas normally thought to be perceptual association areas on a traditional functional brain map should instead be seen as a part of the motor system. I assume they would not deny that these areas are *also* perceptual, and therefore the inference is that these areas are essentially sensorimotor and one might think that the 'representations' of such areas are best thought of as sensorimotor as well.

Again one senses that the Parma group has embarked on an ambivalent departure from the old motor command notion of the motor system. The label 'motor system' itself is linked to the old idea of functionally and anatomically self-contained brain areas involved in action execution as opposed to perceptual and higher cognitive functions. The cortical motor system has traditionally been conceived of as not only functionally but also anatomically distinct, as located within the borders of the frontal cortex just in front of the central sulcus that divides the primary motor and the primary somatosensory cortex.[33] Now, if one sees the functional role of motor areas as categorized by way of perceptual integrations and connections to various parietal areas, then that calls into question the idea of one functionally unified motor output 'system.' Plus, since the neuronal activity in these parietal areas correlates with various different actions, the very anatomical

[33] The classical cortical motor system is anatomically limited to Broadman's area 4 and 6; primary motor cortex and premotor cortex respectively.

location and definition of the motor system is challenged.[34] In this way the new motor research calls both the functional and the anatomical categorization and definition of the motor system into question. However, theoretically these questions do not really seem to be thoroughly addressed by the Parma group as they often simply move borders and divide areas while they maintain the old motor system terminology and categories. I shall suggest that we might have to rethink the traditional idea of a self-contained motor 'system' if we want to understand not only the purpose of the many parallel fronto-parietal circuits but also the organizing role of goals in various cognitive functions. The concept of a goal seems to straddle and unite the cognitive fields of perception, motivation and action, possibly in ways similar to how fronto-parietal circuits physiologically straddle and integrate these influences anatomically. One should perhaps think of these as 'aspects' in common functions rather than as cognitive modules that can be made sense of independently.[35] This also becomes increasingly pressing when looking at one of the most obvious empirical challenges to the 'classical sandwich' notion of central cognition as located in between peripheral input and output systems, namely the findings of sensorimotor integrations *within* single neurons.

2.2.4. Examples of visuomotor neurons

Before turning to the discovery of the fascinating new visuomotor 'mirror' neurons, it is important to introduce some of the other kinds of sensorimotor and particularly visuomotor neurons that had already been found in the premotor cortex of the macaque monkey and which had given rise to significant advancements in the understanding of premotor functions. Rizzolatti and his colleagues did not only look at large scale premotor-parietal connections and anatomy but also at the visuomotor integration on a single cell level. They have in various articles described different complex sensory-motor properties of individual neurons and small populations in the individual regions of the premotor cortex. Further, they also proposed fascinating conceptualizations of their

[34] As an alternative to the activity-action correlation definition one might want to suggest, for example, that motor areas are those that have cortico-spinal connections. However, given the new empirical findings this definition would have plenty of problems of its own for someone wanting to maintain a classical picture, since premotor areas F6 and F7 do not seems to have such connections while certain other parietal areas do.

[35] Henri Bergson makes this suggestion that sensory and motor contributions should be seen as aspects in common mental functions in *Matière et Memoire* (Bergson 1896).

functions given their exact sensory and motor properties and the interrelation between these. I shall here give 2 examples of such complex kinds of visuomotor neurons and their proposed pragmatic functions in regards to perception, attention and action.

A. Interlude – Single cells, macaques and humans

But before I turn to the description of the sensorimotor properties found in single cells of macaque monkeys I want to make a few remarks on the issue of using data from macaque monkeys to make inferences about the workings of the human brain. As mentioned earlier, single cell recordings are invasive and therefore not performed on humans or even 'higher' non-human primates like for example chimps. The fact that one is not so concerned about the macaque monkeys should probably raise some suspicion about very serious differences in the social cognitive skills of macaques and humans. There are many interesting things to note about these differences, but what I want to stress is that the points I want to make about basic sensorimotor integration and affordance based action choices does not in any way seem to be challenged by even potentially enormous differences in social motivations and higher order abilities of language, perspective taking, hypothetical thinking and of forming emotional connections. The reason is that what is important for my argument is that there exist multiple parallel non-kinetically specific goal/affordance organized sensorimotor integrations in fronto-parietal circuits of humans, and I am yet to see *any* of these core ideas challenged in *any* publication.

What has been empirically challenged by Dinstein and others, as I will discuss later, is that there are large amounts of strictly congruent mirror neurons that respond identically during action execution and action observation. I argue that my action monitoring hypothesis is not undermined even if it turns out that there are no such totally 'agent neutral' mirror neurons. Rather, my hypothesis rides not on the exact properties of individual mirror neurons but on the existence of a broader action affordance structure system where various kinds of sensorimotor neurons work in concert to monitor and engage the present environment and others in it. In other words my hypothesis is that we monitor goal - or teleological - relations and actions. Hence the existence of exact mirrors in the mind are neither predicted nor needed for the theory. But what is predicted is that there is *some overlap* between sensorimotor based monitoring of others' actions towards a certain goal and my own perception of and engagement with that goal – not that there is equivalence between neural activity within single neurons or small populations during observation and execution.[36]

[36] I shall later discuss some of the challenges of the existence of visuomotor mirror neurons in humans, and the suggestion that there might be two neighboring neuronal populations in the premotor areas, namely one

B. ‘In my space’ neurons, pragmatic maps and spatial affordances

In area F4 of the ventral premotor cortex Rizzolatti and his coworkers found neurons that seem to map the space around and on the surface of the body and face and often also the appropriate movements towards the location of the stimulus as it is entering a somatosensory receptive field or the ‘reachable’ – i.e. the peri-personal – visual space (Rizzolatti & Gentilucci 1988, Fogassi et al.1992). The details of the findings are fascinating in that most bimodal neurons for example seem to anchor visual properties around the somatosensory receptive fields in an anticipatory way such that the visual receptive field is larger for rapidly approaching stimuli than for slower ones. The motor properties of these neurons most often have to do with movements and head turns towards the somatosensory and sometimes visually mapped stimulus location.[37] All in all F4 and its circuits with the parietal area VIP seems to provide a pragmatic map that coordinates actions and spatial stimuli information. One could call these sorts of sensory motor neurons ‘in my space’ neurons, in that they seem to signal a tension of the incoming stimulus and the invitation or affordance to pay attention and act towards or away from it.[38] Rizzolatti and Sinigaglia compare this egocentric body frame map with the pragmatic eye-movement map based in the circuit between the frontal eye field in F7 and the parietal area LIP. This later map integrates visual stimulus information coded in retinal rather than body coordinates with saccadic eye movements to this location. One could say it provides a map of afforded eye movement directions whereas the F4-VIP circuit has to do with afforded limb/head movement directions. [39] (Rizzolatti & Sinigaglia 2006, p. 61) An important point here is that it is difficult to make sense of these circuits

---

uniquely motor and one uniquely sensory. I agree with Rizzolatti that this sort of total segregation is just physiologically unimaginable, in so far that if these ‘sensory’ and ‘motor’ neurons are thought to communicate and project to each other, then their response per definition becomes ‘sensorimotor.’ (Rizzolatti, personal communication) Or to put it differently, as long as nobody is willing to deny sensori-motor integration in these areas, the denial of individual sensorimotor neurons seems far-fetched. However, it should be noted that this is an argument about the existence of sensorimotor neurons as such and does not lend any evidence to the further claim about existence of strictly congruent mirror neurons.

[37] Note that Rizzolatti and Gentilucci found that only about half of the F4 neuron studies responded to active movements, which is in and of itself fascinating given that it is a premotor area. (See Rizzolatti & Gentilucci 1988, p. 279).

[38] Imagine her for example someone unexpectedly throwing a ball at you and your last minute reaction of dodging.

[39] See here also the fascinating work from Jacqueline Gottlieb’s Lab about this ‘salience mapping’ of potential saccade targets. (Gottleib et al. 2009).

without the concept of the pragmatic relation between perceiver and percept. Rizzolatti and Sinigaglia point to the "essentially active nature of the representation of coded space" and quoting Ernst Mach they write that "'the points of physiological space' are nothing other than the 'goals of various movements of grabbing, looking and locomotion.'" (Rizzolatti & Sinigaglia, p. 67) Thus, the perceived space is, by way of sensorimotor integrations, a pragmatic space.

I here pause for a moment to spell out some intriguing implications of this idea of a pragmatic map and the findings of 'in my space' neurons. First of all, as hinted in the quote above there is a question as to whether one should think of the pragmatic space as simply linking two preexisting sorts of representations, i.e. on the one hand a non-pragmatic 'objective' spatial representation of the stimulus and on the other hand an independently stored motor repertoire catalogue of actions in space. Or, conversely, whether the pragmatic spatial representations should be thought of as primary to more abstract representations of space and actions. Secondly, it is worth noting that such pragmatic sensorimotor integrations are difficult to make sense of in a traditional framework of mental representations as having a certain propositional 'content' and an 'aboutness' relation to some external or imaginary thing. Most often mental representations have been thought of as a 'redoubling something' say X in the external world by having the mental content X. The act or phenomenon of representing is then more or less thought of as an intentional 'aboutness relation' between 'real' X and the mental X.[40] Looking at these pragmatic sensorimotor integrations one might want to say that rather than representing in the sense of being *about* something, they *re*-present a relation between environment and agent. In other words, there is a sort of sensorimotor re-actualizing or bodily recollection of a previously learned teleological relation in a, so to speak, new present 'presentations.' The *content* of such a re-presentation might be said to be the intentional *relation* between agent and environment. But if some mental content thus can be seen as being internalized through previously learned sensorimotor relations then it seems to me that we have a wedge in the door to the not so secret life of the mind. It is not that such mental re-presentations are different in principle for overt observable engagements, but rather that there is a temporal difference of stored/accumulated learning

[40] This sort of picture of perceptual and other mental representations is vehemently criticized by Bergson (1896) and in a more contemporary context a century later by Kathleen Akins (1996).

as opposed to presently specified elements of actions and perceptions – and even thoughts and imaginations. I shall return to this issue as it is very important for my later argument that many action intentions, thoughts and perceptual beliefs to some extent seem to be given in the totality of our observable engagements – and that the understanding of such mental life of others does not rely on narrow action simulation but rather on a sensorimotor based understanding of the shared pragmatic 'affordance' space. My claim is that we to a large extent understand others by way of having similar affordances and teleological representations along with seeing how they choose their actions and react to phenomena in a shared present space. But to be able to explain this claim and its contrast to the view of mirror neurons as simulative, many more issues must be considered and developed.

As I have done in the previous paragraphs, the Parma group also often uses the terminology of affordances, and of how the fronto-parietal circuits code actions 'afforded' by the perceptual environment. Gibson famously coined the term affordance and with this term conceptualized the pragmatic relationship between an organism's motor repertoire, goals and their perceptual environment. The idea of affordances will be pivotal to my project of reinterpreting and resituating the mirror neuron findings in their neurological and social context and hopefully in moving further away from the individualistic non-dynamic input-output model of the mind. I shall say more about how my use of the term differs from Gibson's but I will try not to get too much ahead of myself. One could say that F4 sensorimotor neurons seem to be involved in a 'spatial affordance' integration, but the term is most often used in regard to another category of visuomotor neurons found primarily in area F5 that appears to code something like 'object affordances.'

C. F5 Canonical neurons and object affordances

As discussed above from a motor perspective the Parma Group conceptualized area F5 as containing a 'vocabulary' of goal directed motor actions. However, Rizzolatti and his colleagues also found that many neurons in this area, in addition to their motor properties pertaining to goal-directed actions, often had somatosensory and/or visual properties.[41] Where the somatosensory properties seem to relate to tactile stimulation of the effector of

[41] Rizzolatti and Gentilucci, 1988, P.277-279

the effective motor action, they found that visual responses were often related to the presentations of 3D objects. They later, as we shall see, also discovered mirror neurons in precisely this area, i.e. visuomotor neurons sensitive to actions rather than object perceptions. The object related visuomotor neurons have been labeled canonical neurons.[42] On a conceptual level they are thought to code something like 'object affordances' in that they create an integration between the visual perception of a present object and potential actions *afforded* by this object given one's motor repertoire. Motor responses to object affordances have been well documented by various researchers not only in monkeys but also in humans.[43] The visual properties of canonical neurons found in monkey area F5 seem to relate primarily to the physical dimensions of the object like size, shape and orientation, which makes sense given the hand-object action repertoire of the areas in question. Interestingly the visual response of canonical neurons often seem closely tied to the properties of the object, and in opposition to F4 visuomotor neurons the response is often independent of whether the object is actually within the peripersonal space, i.e. the reaching distance, of the monkey.[44] One might say that the F4 neurons seem to demand an instant response to the incoming stimuli, whether simply a controlled impact or an attentional reorientation to it, and they thus seem to provide an imperative of some kind of reaction. This does not seem to be the case for F5 canonical neurons. If this is so, these canonical neurons might be described as relating possible actions to specific objects rather than coding immediate action affordances on the object given the monkey's present position. In other words the action affordances integrated by these neurons seem at least in some cases to center on the properties of the object and be at a level of

---

[42] The label 'canonical' neurons suggests - wrongly or rightly - that this kind of sensorimotor integration and function is the canonical or typical for this premotor area. Thus the implication is that this kind of affordance coding is the rule rather than the exception. The decision to use this label is rather interesting, I think, because the fascination with mirror neurons often leaves only sparse attention to their context in F5 and more generally as part of the parietal-frontal sensory-motor circuit.

[43] See here for example: Grafton et al (1997), Grèzes et al. (2003), Grèzes, J., & Decety, J. (2002), Handy et al (2006).

[44] See di Pelligrini et al 1992, and Gallese et al1996. In my eyes, the distinction between one's peripersonal and extrapersonal space is in itself very interesting as it is exactly pragmatic affordance based rather than an objectively measurable distinction. The basic idea is that an animal's peripersonal space is the area of immediate action impact and thus flexible depending on effectors, tools and vehicles of actions in general, and that the extrapersonal space is that which is outside of such immediate action reach.

If this is so, these canonical neurons might be described as relating possible actions to specific objects rather than coding immediate action affordances on the object given the monkey's present position. In other words the action affordances integrated by these neurons seem to center on the properties of the object and be at a level of abstraction above the constraints of the actual present relationship between monkey and object.

abstraction above the constraints of the actual present relationship between monkey and object. Accordingly, one might suggest therefore that some of these neurons *monitor* the *affordances of the objects* in the present environment for longer term pragmatic purposes, than simply for immediate action choice.

With different canonical neurons possibly responding in a distributed way to respectively peri-personal and extra-personal objects we would be told not only something about the lasting affordances of objects and what actions are immediately afforded but also something about the broader *affordance structure* of the surrounding environment. Thus one could say that in the case of extra-personal canonical neurons a teleological relation is signaled without determining the interaction, in other words the object could afford a similar action both over time and to other agents in the shared affordance space. As I shall develop in detail later, this leaves an opening for not only mirror neurons but also canonical neurons as playing a role in social cognition that is normally not mentioned in the literature.

The visual response of canonical neurons occurs at the time of object presentation *prior* to any action, and this should be seen in contrast to the so-called ‘motor’ neuronal response *during* action or just prior to action initiation and the ‘somatosensory’ responses that mostly seem to fall at the time of hand-object impact. Canonical neurons appear to integrate visual object perception with *potential* actions as opposed to simply following or initiating the *actual* action. But, moreover, they might also via their visual-somatosensory-motor integration be understood as alerting to the likely outcomes of possible actions[45].

The motor properties of canonical neurons are reported to be distributed similarly to other F5 neurons such that canonical neurons respond to various types and segments of object-directed actions. Some of the canonical neurons respond to very specific motor acts, effectors or temporal sequences whereas others respond to multiple different kinds of acts towards the same goal. So, it seems that clusters of canonical neurons like other F5 populations have a distributed array of motor properties that span and integrate various goal levels and temporal segments of actions. In the case of action production such a

[45] As we shall later see, Gallese and his team often talk about how action mirroring leads to a broader action knowledge including for example sensory consequences. This integration of the potential and actual over different modalities could provide a basic example of how such knowledge can be had without the need to postulate elaborate forward and inverse internal models.

distribution makes functional sense as making dynamical hierarchical action organization, anticipation and control possible. The question is what the functional use of such a distribution is in regard to canonical neurons that are modulated already at the time of object presentation – and often even if the object is out of immediate reach. It seems that the canonical neurons alert to the affordances of objects simultaneously at a more abstract goal level and as inviting various certain concrete acts in certain temporal orders. Thus, one might say that when we perceive objects by way of affordances these objects transcend their concrete physical presence in our perception as our sensorimotor representations alert to teleological possibilities for the future. Given the distributed responses of F5 neurons one might suggest that object affordances are not simply providing one unified anticipation of future actions and their sensory consequences. It is rather the case that a spectrum of hierarchical goals and actions are invited. This could be seen as the basis for a sort of cost-benefit analysis for the purpose of hand-object action choices, and also as supporting our understanding of the evaluations and concrete action choices of others. One might say that given our perception of present affordance structures and other people in them, we might be able to understand something about other people's goals not only from the actions they actually do chose to perform but also from those actions they *do not* choose to perform.

Canonical neurons are mentioned in plenty of articles and books, but surprisingly few articles and studies have been devoted to a deeper understanding of their properties and functions. One of the interesting exceptions is the 2006 article by Vassilis Raos and his colleagues from the University of Parma reporting a study of the functional properties of F5 grasping related neurons.[46] They do not explicitly use the label of canonical neurons, but talk about 'object grasping related visuomotor neurons.' Among other interesting properties they found that the temporal features of not only the visual but also the motor response of these neurons seems to differ from pure motor or motor dominant neurons in this area. This is rather significant as both canonical and mirror neurons normally are described in the mirror neuron literature as *indistinguishable* from other F5 neurons from a motor perspective. As I shall discuss later, the Raos et al. findings problematize the categorization methodology normally used by the Parma group classifying visual and motor properties primarily in terms of a preferred stimulus/action kind. Their focus in

[46] Raos et al (2006).

regard to mirror neurons is also on object/action *kinds* and often ignores complex *interaction* of visual and motor properties and *temporal* features.

2.3. Reinterpreting the cortical motor system

In sum, the Parma group's early findings in many ways problematize the traditional conception of the motor system as a well-defined modular action-command-output system. Already prior to the discovery of 'mirror neurons' that made its members famous well beyond the borders of neuroscience, the Parma group was involved in ground-breaking research and based on their findings engaged in an attempt to re-conceptualize the organization and function of the cortical motor system. They had already found single neurons with various complex visuomotor properties in different areas of the premotor cortex. Due to the nature of the specific connectivity of different premotor and parietal areas they suggested that the different areas served different functions given the different sensorimotor transformations and, more generally yet, that made them suggest that the idea of the motor system as a unified output system should be abandoned. In area F5 they had also discovered that most neurons were modulated by various kinds of goal-directed actions rather than the exact kinematic movements necessary to reach the goal. These and other findings led them to conceptualize area F5 as containing a 'vocabulary' of goal-directed actions and again on a theoretical level further question the traditional conception of the motor system as involved purely in tasks relating to action execution.

2.3.1. Is a more radical reinterpretation needed?

I have been trying to suggest that the findings call for even more radical reinterpretations of motor functions and their organization as fundamentally depending on sensorimotor integration and thus questioning the sustainability of the traditional conception of - not only the pure output *function* of the motor system - but more generally of the *existence* of a unified modular motor 'system' that is meaningfully organized relatively independently of perceptual and central cognitive functions. The many parallel and functionally distinct fronto-parietal circuits and the spectrum of various kinds of sensorimotor neurons in these areas suggest that the motor processes are solicited by perceptual experiences at multiple hierarchical goal-levels and at various spatiotemporal distances *simultaneously*. Thus action planning and action choices seem to take place in the neuro-physiological environment that is 'always already' motorically engaged.[47] And not just engaged with one unified coherently interpreted perception, but rather it seems that sensorimotor integrations work in concert in rather complex and parallel multilevel ways. Thus, this is what I am getting at when I stress that these findings cannot simply be interpreted as the motor system subletting its otherwise well-contained motor processes out and, so to speak, moonlighting for perception. Rather, these findings suggest that motor processes are already intertwined with perceptual processes in *regular* action performance, and the perceptual functions seem to serve action choices. There seems to be no neatly defined independent motor system with extra perceptual – or social cognitive afterschool jobs, but rather motor functions that deeply depend on already being structured by the perceptual and probably social cognitive environment.[48]

2.3.2. My use of the affordance concept

As mentioned the concept of action *affordances* will be quite central to my theory. I use the concept not to discuss a Gibsonian or ecological theory of perception, but instead to discuss goal-organized sensorimotor integrations in fronto-parietal areas. Inspired by the empirical findings referred to above it seems very difficult to make sense of the organization of these areas without some concept that exactly straddles the traditional categories of perception and action. The affordance concept very powerfully resists the traditional input-cognition-output framework as affordances only emerge in the meeting between perception and action – or rather as the continuously evolving sensorimotor experience of an environmentally engaged agent. Given a cognitive system with sensorimotor experience and teleological motivations, affordances are the action invitations that are presently detected in the environment. An affordance alerts to a potential goal and a possible action that the organism could engage in the present environment.

---

[47] Heidegger has made this terminology famous but elements of this idea can also be found in Bergson and Merleau- Ponty. There are strong parallel between m affordacne space model and certain Heideggarian ideas, but I am not able to develop these historical nuances in the present project.

[48] It would be interesting to link this discussion to the issues pertaining to the functional and anatomical distinction between the dorsal and ventral visual pathways (Miner & Goodale, 1996). The question is whether my discussion only pertains to the dorsal pathway. At one level this might be sufficient for my reinterpretation of the motor system. But since STS plays such an important role anatomically and functionally in regards to the fronto-parietal circuits I would guess that both pathways would be important contributors to the affordance space understanding that I hypothesize. However these questions must be left unanswered for present purposes.

I should note that my use of the affordance term in relation to the sensorimotor integrations of the fronto-parietal circuits points to a different psychological status and extension of the concept than how it was conceived of by Gibson and by Gestalt psychologists such a Koffka, who have famously contributed to the development of the idea of affordances. Koffka did not use the notion of affordances but talked about the 'action demand character' of objects. He wrote in *The Principles of Gestalt Psychology* that "…each thing says what it is … a fruit says 'Eat me.'"[49] The gestalt psychologists were mostly interested in the perceptual experience and thought of affordances as a quality of the experience – about what the environment was perceived as demanding or inviting given the present needs and motives of a given perceiver. Gibson was very interested in these and other gestalt ideas but did not like the focus on actual conscious perception, or the suggestion that the 'demand character' was a projection made by the perceiver onto the object – much like secondary sensory qualities have been thought of as projections. Gibson coined the term 'affordance' to suggest a relation of action *possibility* based in the relation between an agent and its environment. He wrote:

> The *affordances* of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill. The verb *to afford* is found in the dictionary, but the noun *affordance* is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment. (Gibson 1979, 127; Gibson's emphasis)[50]

He thought of affordances as relational but, as Andrea Scarantino points out in her article "Affordance Explained,"[51] Gibson as opposed to Koffka thought that they were objective relational qualities of action possibility independently of whether these were actually experiences or even interesting given the present needs of the agent. I fully understand the motivation to move away from the subjective and conscious experience based notion of the Gestalt psychologists. But it does seem to me that the notion of affordances loses some if its value by being extended to *all possible* actions of an agent in a certain context. There might be a tension in Gibson's own view regarding the objectivity of affordances

[49] Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt Brace, p.7.
[50] Gibson, J.J. (1979), *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
[51] Andrea Scarantino (2003), "Affordance Explained", *Philosophy of Science*, Vol. 70, No. 5, Proceedings of the 2002 Biennial Meeting of the Philosophy of Science Association. Part I: Contributed Papers, pp. 949-961. This article gives a very good introduction to Gibson's notion of affordances and how it is different from Koffka's. Scarantino in this article goes on to present a dispositional account of affordances that I to some extent must disagree with, but that is a separate discussion.

and the notion exactly aiming to capture a pragmatic and goal-oriented perspective of an agent in relation to the environment. But this ‘perspective’ of the agent seems to be lost in a too objectively defined notion of affordances as action possibilities.

Further, Gibson often discusses rather concrete and physically measurable relations such as the size of an object as compared to the size of a handgrip and thus the quality of an object as graspable lies in the relation between the object’s size and resistance and the kinematic possibilities of the agent’s hand. He for example writes that "to be graspable, an object must have opposite surfaces separated by a distance less than the span of the hand.”(Gibson, 1979, p.133) But how do we translate such discussions to less concretely physically based and more culturally informed or skill based affordances? It seems that one needs to go under the skin to get an appropriate story of these learned affordances.

I use the notion of affordances as it relates to sensorimotor integrations and by way of this neurological foundation my understanding of affordances somehow lands at a middle position between Gibson’s objective notion of affordances as physically possible actions and Koffka’s subjective notion of consciously perceived action demands and invitations. I suggest that the perceptually triggered fronto-parietal activity for example consists of many simultaneous sensorimotor activations representing many simultaneous action affordances, many of which are clearly not consciously considered or even pertaining to our focal task at hand. However, not all objectively or logically possible actions by an agent in a given environment are translated into sensorimotor affordance relations. Hence, I think that my way of using the notion creates a new neurological conceptualization of affordances that incorporates insights of both Koffka and Gibson while avoiding some of the objections raised against their respective interpretations.

2.3.3. Action choice in a ‘field of affordances’

Now let me turn to another feature of the way I want to use the notion of affordances. Most often affordances are discussed in relation to perception and in a piecemeal fashion considered one by one. My sensorimotor account suggests that we have massively parallel sensorimotor activity and thus that we can detect and entertain many different affordances simultaneously.

Erik Rietveld has recently written a fascinating PhD thesis on the nature of skillful unreflective action, and he makes a forceful argument that such unreflective intentional actions are based on acting in what he calls a "field of affordances."[52] The central idea is that useful and skillful action choices can only be made if we are faced with a spectrum of affordances simultaneously, that engages us and our experience and serves as the 'horizon' of our ongoing action judgments and choices. Rietveld suggests - with the help of Merleau-Ponty, Wittgenstein and Dreyfus - that most skillful actions are not based on explicit preconceived intentions but are chosen via an attempt to get an 'optimal grip' on the situation at hand. The idea of 'optimal grip' is also used to show that being in a field of affordances is not only a question of simply choosing between alluring affordances but also the motivation to ease 'tensions' in our relation to the situation as a whole. In other words intentional action choices are often temporally emerging and evolving in response to the concrete 'tensions' felt in the present 'field of affordances.' I see this as a very important point for making sense of the complex hierarchical and parallel integrations of the fronto-parietal areas.[53] We are addressed by the present environmental situation at many levels and from many sides *simultaneously* and it is due to this complex parallel affordance engagement that we can choose an appropriate spatiotemporal course and sequence of actions. In other words, intentional action choices seem to be motorically grounded and specified not just by way of pure perceptual judgments but in the context of a pragmatic sensorimotor analysis, which already to some extent has engaged our motor experience.[54] This is not to say that we do not have endogenously formed intentions and distal goals that reach beyond the present situation or sensorimotor involvement; we

[52] Rietveld, Erik (2008). *Unreflective Action: A Philosophical contribution to Integrative Neuroscience*. PhD Thesis University of Amsterdam, Department of Philosophy/ILLC. For the idea of action in a field of affordances see chapter 7, 'Embodied existential normativity and the responsiveness to the field of relevant affordances,' and in particular pp. 220-247.

[53] In Rietveld's account it is essential that our behavior is 'concernful' and emotional. I agree. However, I have had to bracket the emotional aspect, which is so central to both action choice and social cognition, as it is thesis project in its own right. But the question of how emotions and reward systems relate to sensorimotor activity is obviously a project that lies in natural extension of my present one of showing how, if we let go of traditional frameworks, findings of mirror neurons and sensorimotor integration in front-parietal circuits can lead us towards a post-cognitivist theory of motor and social cognition.

[54] In vision research it has long been an issue that to make saccades and orient to new stimuli we in some way have to *notice before we really see.* Sub-personal mechanisms of target detection and saccade coordination have been proposed as a solution to this problem. I suggest that if intentional actions depend on being in a 'field of affordances,' then we similarly *act before we really act*. Like pre-attentive visual detection is under different temporal and load constraints than full attentive visual perception, I think we must similarly distinguish constraints of sketchy sensorimotor affordance activations from a full simulated or performed action. I develop this idea in detail in connection with my discussion of Jeannerod's theory of motor cognition and the resulting critique of the role of mirror neurons in social cognition.

clearly do.[55] And, these more distal goals and abstract action affordances most certainly also depend on areas well beyond the fronto-parietal sensorimotor circuits. My point is not to say that the sensorimotor integration is the whole story of intentional action, but that a theory of action and motor functions, which does not take this level of concrete affordance based engagement into account, cannot explain how the more abstract action intentions are turned into actual goal-directed physical actions and engagements. It seems to me that we must *explore* and *monitor* the affordance structures around us in order to make appropriate action choices and perform goal-directed actions in a temporally and contextually appropriate way.[56] Further, my suggestion is that such an affordance understanding cannot be purely perceptual or purely motor for that matter. Rather, it seems hard to conceive of independently of sensorimotor integration as it exactly is a pragmatic interpretation of the environment and of what actions the situation invites.
I shall speak much more about this later, as many interpreters and critics of mirror neurons seem to ignore the role of the sensory in sensorimotor integration, and often have a conception of motor cognition that not only ignores but even tacitly denies the possibility of multiple simultaneous parallel sensorimotor affordance activations. The parallel activation simply seems to be an empirical fact, and for now let us just be reminded that affordance exploration and monitoring of the concrete spatiotemporal environment at least is a possible and seemingly likely function of fronto-parietal sensorimotor circuits at large.

2.4. The discovery of mirror neurons

So with all this background setup I can now turn to the actual neurological main characters: the world-renowned but notoriously tricky ‘mirror neurons.’

2.4.1. Action sensitive visuomotor neurons before the ‘mirror’ label

The first report of the discovery of mirror neurons - i.e. visuomotor neurons sensitive to action perception - was published in 1992 by di Pellegrino, Fadiga, Fogassi, Gallese

[55] See Pacherie (2008) for an interesting analysis of the hierarchical interplay of motor and distal intentions. I would stress also that choices among distal intentions are guided by perceptual elements.

[56] As will hopefully become more clear as my argument develops, I use the notion of ‘structure’ because I expand the idea of a ‘field of affordances’ facing the individual to the idea of a social space in which we can monitor not only the development and change of object affordances over time but also changes and the relational engagements of others. Thus 'affordances structures' refers to the developing relational and teleological dynamics of the action space beyond one’s own present pragmatic perspective.

and Rizzolatti in an article titled "Understanding motor events: a neurophysiological study." This article is fascinating because it is written by the same group of Parma lab researchers, but *prior* to their invention of the term 'mirror neuron' and therefore less bound by the theoretical suggestion that mirror neurons instantiates a mirroring mechanism.

As mentioned earlier, the discovery of this new class of visuomotor neurons happened somewhat by chance. It happened in the context of a study of single premotor neurons in macaque monkeys. Originally the aim of the study was to distinguish the stimulus or the perception related and the motor execution related responses of F5 neurons by way of single cell recordings. To isolate the visual response from the motor response they used a 'delayed response' paradigm, in which the monkey had been trained to switch a light on an object in a box (make it visible), and when a glass door opened after various lengths of delay to release the switch and then grasp the object (motor response) and under it find a food reward.[57] Neurons sensitive to the execution of hand movements were first isolated and then tested for their precise motor and visual properties in the behaviorally controlled situation, where the overt movements of the monkey also were recorded. They report:

> After the initial recording experiments, we incidentally observed that some *experimenter's actions*, such as picking up the food or placing it inside the testing box, activated a relatively large proportion of F5 neurons in the absence of any overt movement of the monkey. (di Pelligrino et al, 1992, p.176, *my italics*)

After this discovery the original delayed response setup was supplemented by a test of the action observation sensitivity of the recorded neurons where the experimenter performed various motor acts in front of the monkey. In this way they could isolate and investigate the properties of this intriguing new group of visuomotor F5 neurons.

In their 1992 article di Pelligrino and his Parma lab co-authors give detailed examples of two such individual neurons that were modulated both by hand action observation and execution. One of these neurons showed a rather specific and strong response to the observation and execution of a precision grip with thumb and index finger. The other, which had a high spontaneous firing rate, was strongly inhibited by the execution or

[57] Given the later findings by researchers like Cecilia Heyes and Caroline Catmur of plasticity and malleability of 'mirror system' responses, it could be really interesting to take a closer look at this paradigm which demands extensive prior training of the monkey – and its mirror system.

observation of various hand object grasps.[58] The difference in the generality of the response of these two neurons can be seen in continuation with the earlier finding that various F5 neurons respond to different action goals levels or segments. An important finding was that these action perception sensitive neurons – in contrast to the previously described canonical neurons - were not modulated simply by the visual presentations of the *object* (even if this was food), nor as the monkey *prepared* its action to grasp the object. Rather, they only seemed to be modulated in a time-locked manner as the grasp was performed or observed. Thus, the timing of the single neuron activity seems to be closely tied to the ongoing action or a segment of it. This specific action sensitivity was the essential new property that had not hitherto been observed, and it was an important reason that these visuomotor neurons were hypothesized as constituting a category of their own, i.e. as importantly different from the canonical 'object affordance' class of visuomotor neurons, which had already been found in this premotor area.

A. Categorizing the action sensitive visuomotor neurons

Additional control experiments were made that further specified the kind of effective actions, and they found that goal-directed hand-object interaction was the key element necessary for the modulation of the F5 action perception sensitive neurons. As mentioned neither the observation of the hand action or object separately were sufficient stimuli to modulate these neurons. Interestingly, object grasping with tools like forceps[59] or threatening actions towards the monkey were also found to be ineffective as stimuli. But they found that as long as the stimuli were goal-directed hand-object interactions the neurons were modulated by action observations even well beyond the monkey's peri-personal space[60] and to a large extent independently of the side on which the stimulus

[58] The dual existence of low vs, high base rate sensorimotor neurons and the modulation of such neurons by target stimuli respectively increasing and decreasing their activity, in my eyes, raises a serious and largely ignored problem in regards to fMRI interpretations. Namely, that the *overall average* of action potentials is measured indirectly via level of oxygen use in the area, and thus different increased and decreased in firing rates might cancel each other out and not 'light up' in the fMRI data mapping.

[59] These results are now contested by some researchers as later experiments have shown that tool mediated goal-directed actions are effective stimuli for some mirror neurons in macaques. (See for example Ferrari et al. (2005) Furthermore, the human 'mirror systems' appear more flexible in terms of what actions they are modulated by, and in specific there is evidence of tool-object interactions being effective.

[60] Peri-personal space is the 'within reach' area around an animal and interestingly it seems to be judged flexibly depending not only on the object movement but also ability to reach, i.e. the availability of tools and of one's own limbs. (See also Iriki 2006.)

was presented.[61] Thus goal-directed hand-object interactions seemed to be essential to the modulation of these new neurons, but they appeared to operate at a level of abstraction above the observing monkeys' own concrete and immediate action possibilities and motivations. However, looking at the actual response it is important to note that the modulation during action execution and observation might happen in relation to *similar* kinds of goal directed actions – but the actual firing rate is exactly symmetrical or identical in any of the histograms published by the Parma group. In other words, if one takes the intensity into account then these cells do not necessarily respond in an 'agent neutral' way for action perception and execution, but rather distinguish the two scenarios.[62] I shall return to the issue of 'agent neutrality' later as this is one of the problematic implications of the mirror metaphor - that the response during observation and execution is indistinguishable.

To prepare the ground for a discussion of the interpretation of the action sensitive visuomotor neurons as 'mirror neurons' instantiating a 'mirror matching mechanism' it is also helpful not just to case-study the fascinating properties of a few such neurons but also to see the bigger picture of both these sorts of neurons and the other neurons studied in area F5 and connected areas. I have already written about various general premotor and parietal sensorimotor findings and now I shall indulge myself in a bit of quantitative data analysis of the new 'mirror neuron' findings. Di Pelligrino and coauthors report the following data. They studied a total of 184 F5 neurons; hereof 87 had visual properties, of these 48 were reported to respond to 'simple meaningful stimuli,' which I take to mean that they had non-action visual sensitivity (probably mostly to objects, but unfortunately there is no further information on these neurons and whether they would fall under the previously introduced classification of 'canonical neurons'). 39 neurons were found to have 'complex visual properties,' apparently meaning that they responded to some kind of visual action stimuli. Thus, these 39 neurons are the focus of the article and they were further sub-categorized as falling into one of the 4 following groups.

1. 12 neurons "in which the effective observed action and the effective executed

[61] This should be contrasted with the typical properties of F4 visuo-motor neurons that exactly seem to integrate spatial perceptual information with appropriate effector specific actions towards these spatial locations.

[62] From personal communications with Rizzolatti I gather that there is some dispute over whether the difference in intensity of the mirror neuron response between self and other is statistically significant – but most published histograms do suggest a systematic asymmetry.

action corresponded," (p.178) given a classification of action kinds that was sensitive to both movement type (grasping, rotating, manipulating of object) and the effector used (hand or mouth).

2. 6 neurons "in which the effective observed action was the one effective when executed by the monkey (e.g. grasping) plus other actions visually similar to the executed one (e.g. hand placing object on a table)."(p.178) I take this description to mean that a broader range of visually similar actions were effective than the exact action that was effective when the monkey itself executed the action.
3. 11 neurons "in which the effective observed actions were logically related to the effective executed actions and could be seen as preparatory to them. For example, the effective observed action was placing an object on a table, whereas the effective executed action was bringing food to the mouth." (p. 179)
4. 10 neurons "which responded to observed actions similar to those described above, but had no activity correlated with animal movement."(p.179)

I shall return to a discussion of these categories later, one should notice the overlap between the first group and what the Parma group in later articles characterize as 'strictly congruent mirror neurons,'(scMN). Group 2 and 3 might both fall under the later label 'broadly congruent mirror neurons,' in which the relation between the effective visual and motor stimuli is not one of identity but rather some kind of other 'link' or 'logic connection.' In this case the link is respectively a category overlap plus scope difference or a logical temporal continuation connection.[63] The 4th group of visually dominant premotor neurons is interesting in that the existence of such non-motor premotor neurons certainly challenges the traditional motor system conception. It should be noted, though, that given the later definition of mirror neurons as modulated by <u>both</u> action observation and execution these visually dominant neurons would <u>not</u> be included under the mirror neuron category as such, but might sometimes be referred to as 'mirror-like neurons.' Thus, to make a quantitative summation of the di Pelligrino et al. mirror neuron findings in percentages:[64] only 21% of the premotor neurons studied showed any sensitivity to observation of actions in this experiment and 16% of the studied neurons were given the later labels mirror neurons, i.e. sensitive to both some kind of action observation and

---

[63] Still,given the later focus on levels of congruency some of these temporal or logical links might be seen as simply incongruent.

[64] Percentages are calculated based on the numbers also given above: study of 187 neurons, 87 of these had visual properties, 29 were sensitive to both observation and execution of goal-directed actions and therefore would fall under the later definition of mirror neurons. Of these 29 neurons 12 had scMN properties, 6 bcMNs with category scope difference, and 11 bcMNs with logical temporal connection between effective visual and motor stimuli, a total of 17 bcMNs. Further 10 neurons were found to respond to visual action stimuli but without showing any motor properties, hence they were perceptually dominant action perception sensitive pre-motor neurons but not MNs.

execution. Of these 16% or 29 neurons, which would later be defined as 'mirror neurons,' 12 showed what has been labeled 'strictly congruent mirror neuron' properties responding to the observation and execution of similar actions[65] and 17 showed some kind of 'broadly congruent' or 'non-congruent' mirror neuron properties. I think it is important to keep these numbers and details in mind in regard to the later hypothesis that mirror neurons instantiate a specific 'mirror mechanism' or 'observation/execution matching mechanism' with a specific function of yielding *shared representations* and thereby supporting social cognition. One would think that the heterogeneity of the studied premotor neurons should caution against hypothesizing an isolated functional mechanism primarily based on the properties of 12 out of 187 studied neurons in an area.[66] These numbers and the heterogeneity of visuomotor F5 neurons should be kept in mind in regards to the theory of a neural mirror matching mechanism – which is based primarily on the findings of strictly congruent mirror neurons – as Csibra has forcefully pointed out[67] and as I shall extensively discuss in what follows. Further, these single cell findings are also critical for the interpretation of the indirect human mirror neurons studies. Increases in fMRI BOLD signals are often interpreted simply as evidence for mirror activity, but given the neuronal heterogeneity it is unclear which exact cells are responsible for oxygen use.

B. The preliminary functional hypothesis

As mentioned the term 'mirror neuron' is not used in di Pelligrino's article, and I want to suggest that something happens to the understanding of these action sensitive visuomotor neurons with the later introduction of the 'mirror neuron' terminology. Given the intriguing findings of the specificity of some action sensitive neurons on the one hand and of the independence of these neurons from one's spatial position on the other, it does

[65] The numbers reported in later articles are slightly different and the ratio of strictly congruent MNs are rarely reported to be more than 30-35% of the total number of MNs. One explanation for this difference could be that a looser notion of 'similar action' is used in di Pelligrino article than in for example Gallese et al. (1996) and Rizzolatti et al. (1996).

[66] As mentioned above another issue is whether the 'strictly congruent' mirror neurons generally show the *same intensity* of response during action execution and action observation. Since it has been argued that the difference in intensity is not statistically significant I shall not use this peculiarity of the data in my main argument, but simply flag that mirror neurons are characterized as 'strictly congruent' not due to their response being the same per se, but simply because they do show *a* response to observation and execution of same action type. For a typical histogram see Rizzolatti et al. (1996).

[67] See Csibra (2005) for example.

seem like a sound idea that some of these F5 neurons could be involved in monitoring or recognition of the others' goal-directed action. But the idea of action monitoring is a much more modest suggestion than the idea of an agent neutral observation/execution mirror matching mechanism, which to a large extent *separates* the function of action sensitive F5 neurons from other premotor neurons and their role in action coordination. In other words the question is whether the fascinating new findings necessarily call for a *separate mechanism* and a *functional distinction* between understanding and prediction of the actions of others on the one hand and the use of these in one's own action selection and planning on the other. And in this early article, despite the excitement of the new discovery, the focus is indeed still on the *continuity* between the new action sensitive neurons and previous findings and theories of the overall function of these premotor areas. The authors write:

> One of the fundamental properties of the pre-motor cortex is that of *retrieving appropriate motor acts in response to sensory stimuli*. Evidence has been provided that action retrieval can occur in response to two-dimensional patterns, color, and size and shape of three-dimensional objects. The present data indicates that in addition to these *physical factors*, retrieval can occur *in response to the meaning of the gestures* made by other individuals. If one considers the rich social interactions within a monkey group, the understanding by a monkey of the actions performed by other monkeys must be very *important in determining action selection*. Thus, the capacity of inferior premotor neurons to select actions according to gesture meanings fits well in the conceptual framework of current theory on the functions of premotor cortex and *expands it to include movement selection related to interpersonal relations.* (p. 179, my italics)

As my overuse of italics indicates, this quote is quite extraordinary in light of the later 'mirror mechanism' theory of these action sensitive premotor neurons. Note first that di Pelligrino and his coauthors understand the new findings of action perception sensitive neurons as related to the general sensorimotor function of the premotor cortex of "retrieving appropriate motor acts in response to sensory stimuli." The notion of 'action retrieval' is difficult because it leaves ambiguous whether it means that a full context appropriate action specification is produced or simulated or whether – as I shall argue - the action goal or sub-goals are simply retrieved and *re-presented* at some more abstract and schematic level.[68] But independently of the interpretations of action retrieval, it is

[68] The idea that mirror neuron activity supports schematic action goal representations should be contrasted with the implications of many mirror matching and motor simulation theories, according to which mirror neurons instantiate or initiate a covert action simulation that operates under the quantitative and temporal

important to note that they here suggest that the general function of the premotor areas is related to appropriately choosing and perceptually guiding actions and the function of the newly discovered visuomotor neurons is conceptualized in this context. This should be seen in contrast with the later mirroring theory formulations, which focuses on the role of mirror neurons in action understanding while saying next to nothing about how this understanding connects to traditional premotor functions of action planning, choosing and initiating. One could wish that di Pelligrino and his collaborators had spent more energy on explaining and justifying the idea that these neurons 'respond to the meaning of gestures,' since it is not clear how one is supposed to understand this notion in regards to the described properties of F5 neurons.[69] But it should be noted that exactly because this idea is much vaguer than the later claim that the action sensitive neurons respond in a 'mirroring' way, i.e. retrieving the very action observed – this phrasing does not exclude or oppose itself to a sensorimotor integration of different or *complimentary* actions. The next sentence seems to imply that two elements are needed to retrieve an appropriate action response to the actions of other, namely action understanding on the one hand and a response to that action on the other. As we shall see the Parma group later linked 'mirror neuron' activity narrowly to action understanding and representation, but here again they do not explicitly say anything about whether these two elements should be seen as instantiated by different processes or neurons. They rather take these action sensitive neurons to be involved in "movement selection related to interpersonal relations," and thus they continue to focus on the overall function of producing an appropriate action response to the social situation. I agree with this early assessment according to which mirror neurons could have an important role to play in relation to one's own movement selection, but this pragmatic function seems to have been effectively ignored since under the spell of the 'mirror mechanism' and its focus on the *epistemic* function of action understanding.

---

load constraints of overt actions. (See Jeannerod 2006, Jacob and Jeannerod 2005) In summary, my argument is that it would not be possible to monitor multiple affordances simultaneously if each was to be specified in spatio-temporal terms. However, we have multiple simultaneous affordance activations. Hence, I argue that it is plausible that we simultaneously can monitor multiple actions schematically as long as we are not engaging in a full spatio-temporal specification of these actions. Accordingly, I suggest that the 'action retrieval' of visuomotor neurons in the fronto-parietal circuits should be seen as schematic sensorimotor goal representations rather than full action commands.

[69] Clearly these neurons did not respond to *all* meaningful gestures as the scientists specifically noted that the neurons were *not* modulated by threatening action against the monkey.

2.4.2. Mirror neurons under the influence of the mirror metaphor

The first articles by the Parma group in which they officially refer to the newly found action sensitive motor neurons as ‘mirror neurons’ came out four years later in 1996.[70] Interestingly, Gallese, Rizzolatti and their coauthors themselves do not make any great stir or fanfare out of labeling and classifying these visuomotor neurons as ‘mirror neurons.’ In the Gallese et al. article they simply imply that the name is based on the finding that there ‘frequently’ is a ‘similarity’ between the effective observed and executed movement.[71] The broad working definition of a mirror neuron in both articles seems to be *a neuron that is somehow modulated in response to some kind of action observation and some kind of action execution*. In other words the *similarity* of the observed and executed action is not included in the official definition of mirror neurons – but is obviously very strongly implied by the name chosen. One might sniff a problematic tension already in these game-setting choices. A reason for the very general action observation and execution sensitivity definition could be that mirror neurons were conceptualized in contrast and opposition to the canonical visuomotor neurons already found in F5, which respond to object perception. Rizzolatti and his coauthors summarize the earlier object perception data in AIP and F5 as suggesting that these “form a circuit which transforms visual information on the intrinsic properties of objects into hand movements that allows the animal to interact appropriately with the objects.”(p.131) They then proceed to describe mirror neurons as a subgroup of F5 neurons that “from a motor point of view are undistinguishable from the rest of the population” but visually respond to meaningful hand movement rather than objects.[72] However, even though a

---

[70] Jeannerod et al. 1995 refers to the findings of the Parma group published by di Pelligrino et al. 1992 as ‘mirror neurons.’ Thus, the term ‘mirror neuron’ probably had been used by the group in some formal settings prior to 1996. The two 1996 articles that I shall refer to here are Gallese, Fadiga, Fogassi & Rizzolatti, ‘Action recognition in the pre-motor cortex’ and Rizzolatti, Fadiga, Gallese and Fagassi. ‘Premotor cortex and the recognition of motor actions’, Cog Bran Res 3:131-141.

[71] “Recently, we discovered a particular set of F5 neurons, which discharged both during monkey's active movements and when the monkey observed meaningful hand movements made by the experimenter. Frequently there was a clear similarity between the effective observed movement and the effective executed movement (di Pellegrino et al., 1992). The aim of the present article is to give a detailed description of the properties of these 'mirror' neurons.” (p.594)

[72] In relation to canonical neuron findings the suggestion that motor properties are indistinguishable from those of the rest of the F5 population might be a rather premature conclusion stemming from the narrow focus on the ‘kinds’ of effective actions that overlap with those of other F5 neurons. Actually, the motor properties of e.g. strictly congruent mirror neurons seem different from those of most canonical neurons in

'similarity' between the observed and executed action is not required for a neuron to be classified as a mirror neuron, the observation-execution overlap is clearly not only a terminological focal point but essential also to the overall theoretical and functional interpretation of these new visuomotor neurons. Rizzolatti and colleagues report the finding that there 'always was a link between the effective observed movement and the effective executed movement.' And from the vague notion of a 'link' they simply jump to the much more specific conclusion that "these data suggest that area F5 is endowed with an observation/ execution matching system."(Rizzolatti et al, 1996) However, given that there also is a 'link' between the effective visual stimuli and motor actions of the so-called 'canonical' F5 neurons and other *non-mirror* visuomotor neurons, it is in my eyes rather puzzling why the Parma group mostly ignores the *complexity* of this link in the case of at least 60% mirror neurons and proceeds to suggest that they instantiate a simple matching system rather than being involved in more complex sensorimotor integration processes, which would be more consistent with the general hypothesis of the function of the AIP-F5 circuit. By using the name 'mirror neurons' for all action sensitive visuomotor neurons an essential choice is made to highlight the 'frequent similarities,' and further to suggest or even assume that the essential functional properties of MNs are to be found in the similarities or overlap between perceptual and motor properties rather than in other possible 'links.'

The data that the Parma group report in the 1996 Gallese et al. and Rizzolatti et al. articles is based on studies of F5 neurons that are fairly similar but numerically larger than those of the earlier di Pelligrino et al. article, and their objective was to further characterize the properties of mirror neurons and theorize their possible functions. However, I think that their experimental and classificatory approaches show a clear bias towards finding that during action perception these mirror neurons indeed did somehow 'mirror' the observed action. Furthermore, with the choice of one common name the bias is not only towards finding mirroring properties but also towards seeing all the different action sensitive neurons as involved in a mirroring function or as being part of one *mirroring mechanism*.

---

their *temporal* features, i.e. the modulation of strictly congruent mirror neurons seems to be time-locked with the execution of an action of a particular part of it, whereas canonical neurons are modulated already prior to the actual initiation of the action. See di Pelligrino et al. 1992, and Raos et al. 2008.

That the mirror properties are emphasized and the heterogeneity of action sensitive visuomotor neurons to a large extent is ignored becomes particularly obvious when one looks at the actual interpretative choices made when conducting and classifying the empirical studies of mirror neurons properties. To set up my re-interpretation of the findings under a broader theory of affordance and action monitoring, I will go through some of the empirical details and try to make explicit some of the many implicit theoretical choices made by the Parma group.

2.4.3. Naming and sub-categorizing mirror neurons

In the Gallese et al. 1996 study, they recorded from 532 F5 neurons, of which they found that 92 (17%) were sensitive to both action execution and observation (MNs). Similarly to the previous study mirror neurons were primarily activated by the execution and perception of various object-directed (transitive) hand and mouth actions. When characterizing the properties and types of mirror neurons Gallese and his coauthors have been focusing on the effective kinds of perceptual and motor stimuli of the neurons. Consequently, like in the early purely motor studies that led to the conceptualization of F5 as a motor vocabulary, they here classify and quantify the neurons in terms of the kinds of action/s that serve as effective stimuli (i.e. as grasping neurons, holding neurons etc.), and then compare the visual and the motor response categories for each neuron. The action 'types' are in this way used as the main organizing principle and – which is very important - are maintained as such also when they turn to the analysis of the relationship between visual and motor properties in individual neurons. The data report for individual neuron properties then becomes a comparison between the effective visual and motor action kinds respectively. In my view, these choices of classifications in addition to the terminological choice of 'mirroring' clearly show that the interest is not so much in the details or complexity of the visuomotor relationship, but primarily in the degree of congruence or similarity between visual and motor properties. In other words, the findings are described primarily using the single dimension of effective action kind overlap ranging from 0-100% in congruency. Accordingly, they characterize the neurons as either 'strictly congruent,' 'broadly congruent' or 'non-congruent.' Under this definition they found that of the 92 mirror neurons studied respectively 31.5% were strictly congruent, 60.9% were broadly congruent and 7.6 were non congruent. The

category of strictly congruent mirror neurons is somewhat self-explanatory and covers neurons where there is an overlap between the effective observed and executed action both in regards to the kind of 'general action' or goal and more narrowly in regards to how it is performed. The category of 'non-congruent' mirror neurons covers the action sensitive visuomotor neurons in which "no clear-cut relationship" was found. The group of broadly congruent mirror neurons is the largest group and covers all neurons where "there was a link, but not identity" between the effective observed and executed action.[73]
These choices of the subdivisions of the mirror neurons in general and broadly congruent mirror neurons in particular are rather thought provoking. If one compares the grouping made here to that made in the di Pelligrino et al. 1992 article, a few things that spring off the page are 1) the disappearance of the subcategory of broadly congruent mirror neurons that have a 'temporal link' between the effective visual and motor action, 2) the appearance of the 'non-congruent' MN category and 3) the new break down of the scope-difference category into 3 different sub-categories.
It is possible that the Gallese at al. 1996 study produced radically different results than the earlier study, but another reason for the new findings could be that they now simply have a new agenda and therefore a different focus and motivation in categorizing the results. The main goal of Gallese and his coauthors in the 1996 article is to explore mirror neurons and their mirroring properties, which leads to a classification that focuses on the degree of overlap (from 0-1) between the various 'kinds of action' that give rise to respectively a visual and a motor modulation. Consequently there is even less attention paid to the kind of structured difference and interrelation there might be between visual and motor properties in the individual neurons. The concrete changes in terminological categorizations seem to be a clear expression of the new focus. In short, it looks as if they asked only two sorts of questions of the data; namely, on the one hand what kinds of actions and object-effector interactions are effective stimuli and on the other hand to what extent there was an overlap of *kind* between the activity during observation and execution. This would explain why the broadly congruent mirror neurons with temporal connection did not receive any attention. They either disappeared or fell under the

[73] The broadly congruent group is rather heterogeneous in terms of the relation between visual and motor properties, and Gallese et al. 1996 break it down into 3 subcategories. But in doing so they still only look at which action types and effectors are effective during respectively observation and execution.

category of 'non-congruent mirror neurons,' since these neurons might respond to the observation of 'grasping by hand' and the execution of 'mouth opening' and hence show no overlap whatsoever of either action category or relevant effector. However, as I will discuss further down, these temporally linked visuomotor neurons later become pivotal in the theory of mirroring as a process of 'action chain' activation. For now, though, the classification choices seem to be an expression of the preoccupation with 'action kind' categories, and maybe an exaggerated fascination with the properties of strictly congruent mirror neurons, in which the effective action(s) in the case of observation and execution have a close similarity.

What I want to show with this analysis is how the early version of the mirroring theory is based narrowly on an idea of simple action mirror matching inspired primarily by the idealized mirror properties of strictly congruent neurons. The bottom line is that there is no real interest in the details and functions of the heterogeneity of low or non-congruent mirror neurons. And note that this disinterest is maintained *despite* the fact that at least 2/3's of the neurons with 'mirroring properties' were found to be 'broadly congruent' or 'non-congruent,' meaning that their visual and motor properties, albeit possibly overlapping, were found to differ significantly. If one took the predominant broadly- and non-congruent properties of mirror neurons seriously, it would seem important for a theoretical interpretation of mirror neurons to get a detailed analysis of the various visuomotor property combinations and the possible logical, categorical, or temporal relation between effective perceptions and action executions. However, Gallese et al. 1996 do not dwell on such experiments or analyses but simply refer to Rizzolatti's earlier characterization of F5 as containing a 'motor vocabulary' and classify the neurons according to action types as 'grasping neurons,' 'manipulating neurons' etc.[74] Thus, they simply assume that the important functional role of these neurons has to do with their

[74] As mentioned in the early part of the chapter, it is important that the idea of a motor vocabulary can be interpreted in different ways, and that Gallese et al 1996 seemingly pull the attention away from the idea of the vocabulary as fundamentally *sensorimotor* in the sense of providing the appropriate response given a perceptual situation and more towards a more abstract *context independent action repertoire* representation. Another intriguing sign of the isolation of 'mirror neuron' from other functions and neurons of the premotor cortex and also within the F5 can be found in the introduction to the article where the premotor cortex is presented as consisting of a "mosaic of areas with distinctive differences in structure and connectivity." While this is true, it is interesting that there is no mention of possible functional commonalities between such areas such as 'action retrieval given sensory information,' which paves the way for an understanding of mirror neurons that does not have to do with an action response to others' actions but rather a more abstract action representation.

motor action type classification and with matching of action observation and execution. Their choices of classifications give the impression that the large number of broadly congruent mirror neurons can be treated as something like ‘less specific’ aberrations of the strictly congruent mirror neurons. I think that what is happening here is a typical example of letting a preconceived logically simplified theory – like mirror matching – control the interpretation of the data. The risk of this of course is to some extent analogous to throwing the baby away with the bathwater, in the sense that much of the detailed data that should have informed us and helped create, guide and improve theories is being ignored.[75]

In brief, it seems that the mirror-matching metaphor in many ways was driving the experimental and categorization choices of the Parma group in regard to all action perception sensitive neurons, and thus it seems that the *common* function of all these neurons as instantiating some kind of mirror matching mechanism to a large extent is assumed rather than empirically tested. The mirror matching mechanism in many ways seems to have been conceived of on the basis of an even more idealized version of strictly congruent neurons. Thus, most articles and books about mirror neurons – including this present thesis - start with a definition of these along the lines *of neurons that respond to both the observation and the execution of the same action*. Here, we see the exact focus and bias of the early classification: these neurons are *mirror* neurons because they are modulated by certain action *types* in an agent-neutral way. Furthermore, by simply focusing on the action type and the observation/execution overlap, it is tacitly implied that other parameters are irrelevant. It is of course the exclusion of other parameters that allows the metaphor of ‘mirroring’ to do its real magic. By way of the mirror terminology these visuomotor neurons are lifted out of their neurological context and reified as little *modular* mechanisms matching observations and executions of their favored actions. Ultimately, in abstract terms the activity of mirror neurons is conceived of *as agent neutral, context independent and ubiquitous*; the only necessary and sufficient condition for their modulation is the observation or execution of their preferred action type/s.

---

[75] This sort of simplification also often happens when two alternative theories are fighting, and the two camps use evidence merely as ammunition against one another, rather than as the fabric for development of new or simply improved theories.

2.5. Chapter summary & conclusion

I have in this second Chapter presented the empirical and theoretical backdrop of the Parma group's famous discovery of mirror neurons. I have argued that the general findings of sensorimotor integration in the premotor-parietal circuits radically challenges the traditional view of the motor system as a dedicated output system simply matching intentions with motor commands. In this context I discussed the finding of goal-organized premotor neurons and some of the concrete pragmatic properties of various sensorimotor neurons such as canonical affordance neurons. I suggested that the findings might point to the need for an even more radical reinterpretation of the motor system than those suggested by the Parma scientists themselves. Hence, the discussion is used to make an initial presentation of my view of these premotor-parietal circuits as supporting an 'affordance space' monitoring function that is used both perceptually – namely to understand and predict the pragmatic relations and circumstances that we face – but which is also used to guide and orchestrate our actual ongoing action choices and action specifications. Thus, I pointed to subtle tensions between the interpretative assumptions and frameworks employed by the Parma group and the early findings themselves. These tensions are brought to the fore by the actual discovery of mirror neurons and the subsequent terminological and classificatory choices that drive their interpretation. These choices were discussed in some detail as the general received view and research of mirror neurons continuously have been shaped by these tacit interpretive decisions. Thus, I explained how mirror neurons came to be interpreted in isolation from the rest of the fronto-parietal circuits, how the pragmatic and heterogeneous nature of sensorimotor organization was quickly forgotten as everybody got mesmerized by the seeming symmetrical observation/execution properties of mirror neurons. With these interpretative choices the idea of an agent neutral, context independent and ubiquitous mirror mechanism with the function of action understanding takes shape. Via the mirror mechanism the otherwise pragmatic motor system could be given a social cognitive epistemological function. I will now discuss how this theory has been developed, how the mirror mechanism has gradually become even more removed from its neurological and relational context and will underline the continuously mounting empirical and theoretical tensions this idealized mirror mechanism generates and faces.

Chapter 3

Empirical and theoretical tensions of the mirror mechanism

3.1. The persistence of the caricature mirror mechanism

I have already in the previous chapter suggested that the idea of a context-independent agent-neutral mirror mechanism to some extent is inherent in the early terminological and classificatory choices. It is important to note that this is not the explicit definition or terminology used by proponents of the mirror mechanism. However, my point is that this idea which I shall call the caricature model of the mirror mechanism, albeit most certainly false and possibly without explicit proponents, has been and still is immensely influential in the literature as well as in the empirical research process.

As we saw from the admittedly tedious details of the previous Chapter this caricature idea of mirroring is based on a *highly idealized* – and, given the heterogeneity of the findings, rather dreamy - version *of strictly congruent mirror neurons* as producing identical or *symmetric* responses to the observation and execution of *non-contextually* described action types. The idea is that individual mirror neurons respond ubiquitously in a symmetric mirror fashion to preferred action types. Thus, one can formulate a caricature story of the mirror *mechanism*, i.e. the neural mechanism implied by the observation-execution mirror-matching metaphor is one where certain neural responses are modulated narrowly by certain observed actions independently of who does these actions, i.e. whether they are observed or executed. Furthermore, the metaphor also seems to imply that this neural mirror mechanism is modulated consistently by the relevant action types and not by all sorts of broader practical and social contextual elements. Consequently, given this action type modulation the implication is not only agent-neutrality but also a high level of context-independence. In short, the mirror metaphor and early descriptions push the caricatured idea of a ubiquitous and rather modular, context-independent mechanism of agent-neutral action *type* representation.

I cannot reiterate often enough that not only the experimental findings but also the explicit interpretations made by various members of the Parma group mostly suggest a *much more complex* picture than the caricature model as I have spelled it out here. But the problem is that the Parma group never explicitly abandons the metaphor of particular action execution/observation mirror matching mechanism, and it seems to me that the caricature version in many respects is dominating the research from the early classifications and to the present day constrains the sprawling number of mirror neuron publications, discussions and applications. In this chapter I will back this claim up by showing how the various elements of the mirroring metaphor are doing tacit work in the further experimental investigations and interpretations of mirror neurons and their broader hypothesized mechanism. The details of the examples might seem overlabored but I want to show the tensions between the driving mirroring metaphor and discussions of the findings and thereby also the often much more sophisticated and squarely 'post-cognitivist' intuitions of the Parma group. In other words, it is not just the fact that there are tensions, but the nature of these tensions that will guide my reinterpretation of the cognitive 'mechanism' of mirror neurons and their hypothesized social functions and importance.

3.2. <u>Empirical and theoretical tensions in single cell findings</u>

With the awareness of the underlying caricature metaphor of the mirror mechanism, I now want to introduce more of the even muddier reality - both in regard to findings but also to the Parma group's own interpretations thereof.

3.2.1. From mirror neurons to an observation/execution matching mechanism

Gallese and his coauthors argue for their functional interpretation of mirror neurons as instantiating an observation/execution matching mechanism. But how more specifically do they conceptualize this mechanism, and what more precisely is thought to be 'matched'? They suggest "that their [MNs'] discharge generates an internal representation of the movement." (Gallese et al. 1996) In regard to mirror neurons they make a subtle shift in terminology from a focus on action retrieval to one on action *representation*, or rather a distinction between action retrieval *for* representation versus action retrieval *for* execution is introduced to pave the way for a less pragmatic and more epistemological role for mirror neurons than F5 in general. Rizzolatti and his coauthors thus write:

> These data suggest that area F5 is endowed with an observation/execution matching system. When the monkey observes a motor action that belongs (or resembles) its movement repertoire, this action is automatically retrieved. The retrieved action is not necessarily executed. It is only represented in the motor system. We speculated that that this observation/execution mechanism plays a role in understanding the meaning of motor events. (Rizzolatti et al.1996, p. 132)

This description might capture what one could call the basic idea of action mirroring – and note that this is a much more vague idea than the one spelled out above as the caricature model. Mirror neurons are thought to instantiate representations of goal directed actions in one's own motor-repertoire. The action observation triggers this representation and at some abstract level 'retrieves' the goal directed action. However, whether this action is actually to be executed - and maybe whether it is to be specified in all kinematic details in relation to a concrete context – depends on cognitive processes beyond the observation/execution mechanism proper. The idea that matching goal-directed actions are being 'retrieved' based on the perception of hand-object interactions is not in and of itself inconsistent with my affordance structure monitoring account of the front-parietal circuits. As it will hopefully become obvious in what follows, the extent of the disagreement very much depends on how one understands the nature of the 'action retrieval,' 'action representation' and more broadly how one conceptualizes this process in the broader context of the motor system. One question is whether the action retrieval is to be thought of as a fully specified spatiotemporal action sequence, which I would deny in the case of normal action perception and understanding. A further question is whether we are dealing with an *encapsulated* process of *symmetric* mirroring. And, consequently whether one should see the social function of this process as tied simply to the 'matching' or, as I would suggest, the matching gains its cognitive value through the broader sensorimotor 'affordance space' monitoring circuits.

In short, if perception induced action retrieval is not understood as caricature mirroring, i.e. ubiquitous modular context-independent agent-neutral mirroring, then I am not necessarily opposed to the idea. But as I will show in the following chapter this is not the direction things have taken. Rather the differences between my interpretation and the mirror-matching hypothesis are further attenuated when Gallese and Goldman in their problematic and highly influential joint article of 1998, which launches the theory of ubiquitous and agent-neutral mirroring simulation of action related mental states.

Gallese, Rizzolatti and their colleagues mostly show a much more neurologically nuanced, contextually integrated and complex view of mirror neurons, than the caricature and simulation theory of mirroring suggests. What I for brevity call their 'post-cognitivist' intuitions are particularly evident in their broader attempt to reinterpret our conception of the motor system and sensorimotor circuits, which I discuss in the previous chapter.[76] However, the problem is that with the choice of the mirroring metaphor and the focus on individual neurons and non-contextual actions they seem in many ways to have plunged right back into the traditional cognitivist frameworks of modular and serial information processing and distinctly organized perceptual and motor representations. In other words, the caricature theory of mirroring, which is implied already by the idea and concept of 'mirroring,' shapes and creates tensions prior to and independently of its more explicit formulations. And this is the problem that I will be elucidating and attempting to untangle in this chapter.

3.2.2. What is mirrored: kinetic movements, goals or what?

The initial mirror neuron findings already show that the properties of individual action perception sensitive premotor neurons are much more complex and heterogeneous than the 'mirroring' label and interpretation might let on. In Chapter 2 I discussed how the focus on 'action kinds' might overshadow more subtle temporal and contextual visuomotor properties in many neurons. In this section I will refer to some of the later and more recent studies of mirror neurons and by way of these try to shed some light on what it actually is that various mirror neurons might be said to mirror.

[76] These ideas are suggested to be 'post-cognitivist' in the sense that they challenge the traditional notion of the motor system as an output system and the idea of serial as opposed to parallel processes. Further, the relevant research raises questions as to whether one can meaningfully distinguish neurological process vehicles and connections from their representational contents. I have discussed how the sensorimotor organization might be essential to goal representation, and it is a key element in my later argument for the role of these areas in social cognition that such sensorimotor goal representations in a way break down another cognitivist and old philosophical dichotomy of the mental as inner and behavior as outer. Ultimately, these observations begin to dissolve or at least reframe the notorious paradox of how we can have knowledge of other minds.

A. Goal-directed impact - not all actions are mirrored

The concept of action mirroring implies an automatic and kinematically detailed reproduction of any observed movement. From the very discovery and even before the choice of the 'mirror neuron' label it was clear that the premotor neurons in question were not instantiating such an exact ubiquitous mirroring process. Most obviously it was apparent from the beginning that the range of effective actions was limited and that goal-directed hand-object interactions seem to be the most - if not the only - effective actions. (di Pelligrino et al 1992, Gallese et al 1996, Rizzolatti et al 1996) It has later been shown that there is some flexibility in both what can count as a hand-object interaction and even more broadly that some familiar intransitive actions might be effective as well – i.e. digestive mouth actions and grasps by robots and with tools.[77] As I shall discuss later the human mirror neuron system's response might be even more flexible, but even here it is obvious that the range of effective actions in the fronto-parietal circuits is much narrower than for example perceptual action sensitive neurons in the anterior superior temporal sulcus (STSa).[78] For most mirror neurons the important feature seems to be that the action is understood as goal-directed. This explains the finding that mirror neurons in monkeys initially do not respond to actions performed by robots and tools but after having been familiarized with such effectors in successful goal-directed actions, many mirror neurons start responding to these actions as well.[79] The finding of a more flexible mirror neuron system might be a reflection of a much more liberal notion of what can count and be organized as a goal in the fronto-parietal circuits.[80] Be that as it may, the point for now is that various goal-directed actions with a peripheral effector - most typically the hand - seem to be the most effective perceptual stimuli for mirror neurons.

B. Mirror neurons as triggering actions from motor vocabulary?

The second part of the question of 'what is mirrored' is about how to characterize the evoked mirror neuron activity. The conceptualization of the motor representation in F5 is

[77] See here for example Umilta et al. (2008).
[78] These STSa properties will be discussed in more detail in the section on the transition from F5 mirror neurons to a mirror neuron system.
[79] See the Umilta et al. (2008).
[80] Action organization in general might be characterized as goal and sub-goal organized even for simple motor commands in primary motor areas. The idea is that even bending a fingertip might be initiated by way of a sensorimotor representation of the goal – the anticipated visual and somatosensory feedback of the successful movement. (James, 1890)

rooted in Rizzolatti's earlier suggestion that this area contains a 'motor vocabulary.' I have presented this idea in the previous chapter and the core idea of this motor vocabulary is that motor actions are to be thought of as 'words' organized around meaningful goals rather than the specific kinetic movements. F5 is thus described as representing actions at a level of abstraction well above the various actual motor commands necessary to execute the action in a concrete context. What is now added is that via the activity of mirror neurons these motor 'words' can be triggered not only endogenously in the context of action execution but also simply by a perception of the very actions represented in the vocabulary – in the absence of overt action.

The motor vocabulary story seems to be that mirror neurons and other visuomotor neurons do not themselves integrate perceptual information that would be helpful as guide for action execution, but merely trigger the retrieval of action representations already organized in the vocabulary.

But what is actually the argument that individual mirror neurons are responsible for or contribute to action execution in their 'active mode' during action execution? Via the notion of a 'motor vocabulary,' mirror theorists often almost seem to take the role of mirror neurons in action execution as a given. But as I see it the evidence for their precise causal role in the production of overt actions is actually not very clear as the interpretation generally simply relies on the correlation of activity. I think it is crucial to further investigate the action production role for a better understanding of these neurons' general function. It is normally suggested that the reason why mirroring does not result in overt action is due to an active counter inhibition which I shall discuss later.[81] The idea is that all activated 'words' lead to action if not inhibited. As opposed to this idea, I hypothesize firstly that action initiation under normal circumstances of many parallel action invitations and motivations needs much more than a few active sensorimotor neurons. Secondly, I hypothesize that strictly congruent mirror neurons, which are modulated at the onset of both executed and perceived actions, might be monitoring rather than initiating actions – and that the role of these neurons during action execution is to coordinate the temporal onset of motor acts or sub-goals in goal directed action sequences. In other words, for one to plan an upcoming act one must know which act one is performing at this very moment. Thus, I would argue that that there is a problematic

[81] See section 2.3.4. C.

tacit move from sub-personal findings of individual mirror neuron activity to global action outcomes, and that it is a rather wild jump to simply assume that the activity of single neurons would automatically lead to overt action if not inhibited. More on this later, but the point here is that if the mirror theorists want to 'have their cake and eat it too' one would expect at least an explanation of how.[82]

C. Abstract sensorimotor representations

The idea of mirror neurons and F5 activity in general as linked to somewhat abstract representations of goal-directed actions has also been backed up by various later findings. For example the Parma lab found that there are also multimodal mirror neurons in F5. In a study reported by Kohler et al. in 2002 they found premotor audiovisual neurons that respond to the sound of actions similar to those that are effectively modulates these neurons when executed or observed.[83] Thus these neurons seem to be modulated by the presence of a given action irrespective of perceptual modality. Another very important finding was that the monkey actually might not even have to see or in any other way directly perceive the hand-object interaction to get a mirror neuron modulation. Umilta and her coworkers report in the 2001 article 'I know what you are doing: a neurophysiological study' that the modulation of some premotor mirror neurons simply depend on the knowledge that the action is going on.[84] This conclusion was based on an ingenious single cell recording study where the monkeys watched various fully visible and partially occluded hand actions. This study provided strong evidence that the activity of many mirror neurons is driven by the inference of the presence of a hand-object interaction rather than the actual direct observation of it. Another important finding in this study was that pantomimed hand actions without the presence of an object – whether the pantomimed action was fully visible or whether the objects absence was detected simply on the basis of prior information – did not modulate the mirror neurons activity. It had been reported from the first di Pelligrino et al. study in 1992 that mimicked or pantomimed actions without objects were largely ineffective as stimuli. Now, the finding

[82] I will in Chapter 5 discuss how Jacob and Jeannerod accuse the Parma group of wanting to eat their cake and have it too. The problem lies, on the one hand, in the caricature notion of mirroring and, on the other, in Jacob and Jeannerod's relying on implausibly restrictive views on motor and social cognition.
[83] Kohler et al. (2002). See also Keysers et al. (2003).
[84] Umilta et al. (2001).

that pantomimed or mimicked actions in the absence of a target are rather ineffective for many mirror neurons has been robustly replicated. However, it might be that this finding is particular to certain neurons on the central convexity of area F5, as Nelissen et al. 2005 reported that mimicked actions were effective for neurons in the anterior part of F5 folding into the acuate sulcus.[85] Therefore, these findings underline the problem of trying too hard to understand all action sensitive visuomotor neurons under a single heading as 'mirror' neurons or as modulated by one coherent set of stimuli such as visually perceived hand-object interaction. However, irrespective of the heterogeneity of sensorimotor integration, what in my eyes is pivotal about the new occlusion study was that for many mirror neurons the exact same observation – of a hand reaching behind an occluding screen – yielded a different response within single neurons given the prior contextual knowledge of the presence or absence of an object behind the screen. This is a truly fascinating finding that fits very poorly with the caricature model of mirroring as simply matching the motor action observed in a rather ubiquitous and modularly insulated way, depending narrowly on the motor action 'word' represented by the neuron and the visual perception of a similar action. What this finding suggests instead is that the response of individual mirror neurons is modulated by the broader temporal and contextual knowledge and perception of the present environment. And the occlusion study fits very well with my proposal that mirror neurons are part of a broader sensorimotor circuit – including also canonical neurons – that supports an ongoing monitoring of the contextual affordance structures around us. In other words, I am suggesting that the presence or absence of the object is monitored by classical canonical neurons and that the modulation of mirror neurons attuned to goal directed hand-object grasping is dynamically dependent on the overall front-parietal information about the shared affordance space that the perceiving animal is in.

Overall, it can be concluded that many mirror neurons seem to 'code' the action of the other at a level of abstraction where the actual visual perception of the hand-object interaction is not necessary and that the perception of the motor act alone is often insufficient for mirror neuron activity. In neurons such as those studied by Umilta et al.

[85] See Nelissen et al. (2005) They write: "Observation of mimicked human actions also activated both F5a and 45B. The interaction between mimicked and goal-directed actions was significant in 45B, but not in F5a. In 45B, the signal was significantly stronger, compared with static controls, during goal-directed action than during mimicked action."

what seems to be important is whether a goal-directed interaction is *deemed* to take place in the broader shared space at this very moment. It is very plausible that other visuomotor neurons would have other preferences as suggested by the fMRI findings of Nelissen et al 2005. My point is that the simple mirroring metaphor does not capture the complex pragmatic and context sensitive tracking of the various features of the affordance space that both the action and object perception sensitive sensorimotor neurons seem to contribute to.

3.2.3. Interlude: affordance space and relational action perception

The terminology of respectively 'monitoring' and locating actions and goals in a broader shared 'affordance space,' that actions are understood in relation to monitored 'object affordances' and of observed actions' ability to represent potential 'social affordances' is mine. And shall now pause to explain what I mean by these expressions. When the Parma group interprets the findings reported above they mostly talk about 'representing' and 'retrieving' goal-directed 'action kinds'. Thus, in spite of the findings reported by Umilta et al. 2001 and Kohler et al. 2002, they normally speak as if the modulation of mirror neuron activity depends simply on perception of actions, and not on judgments about the locations or the pragmatic or social relevance of these actions. I use the concept of an 'affordance space' and a 'social affordance space' rather than simply affordances to capture the expanded idea that we monitor not only the concrete immediate action invitations one by one, but the broader pragmatic and potentially social structure of the environment we are in. As reported earlier, many both canonical and mirror neurons are modulated by observations of actions and objects both inside and outside one's peripersonal space. This has been taken as evidence that mirror neurons represent actions in an allocentric/other-centered or at least agent-neutral context-independent way. However, one might also simply suggest that these neurons contribute to a monitoring process not just of actions that are afforded us immediately in our present concrete position or from our concrete perspective, but that premotor and parietal circuits help us monitor and anticipate the broader action and affordances *structures* of the space we are in, i.e. the space our ongoing hierarchical and temporal action choices are made in relation to. In other words, I have coined the notion of an affordance space and the idea of affordance structures to open up the possibility to talk about something as an affordance even if it is not currently perceived directly, or out of immediate reach, or if it is perceived as an affordance for someone else. The point is to move beyond the rather presentist and somewhat solipsistic tendencies of Gibsonian uses of the affordance term and explore is powers in regard to social cognition and our broader understanding of lasting teleological possibilities in the environment over time.

However, to make use of such a broader social affordance space it is important to be able to not only understand others' actions but also how precisely they relate to one's own action choices – i.e. actions and objects should also be represented in an egocentric way. Thus, given such a hypothesis one should expect various modulations of these areas to depend on contextual and pragmatic factors of how the observed action relates to one's own possible actions. A growing number of recent findings suggest that the premotor-parietal circuits are indeed modulated by various factors of social relevance. As many of these studies are based on human imaging and behavioral data on the overall modulation of the areas, I shall discuss these findings in relation to the move from the single mirror neuron based theory to a broader theory of a mirror neuron system in humans. For now, I want to point to a single cell study recently published by Caggiano and other colleagues from the Parma group, which seems extremely supportive of the hypothesis that mirror neurons help track the affordance space and the pragmatic and social relevance of observed actions.[86] They found that many mirror neurons differently encode the peripersonal space in which they can make immediate reaching/grasping actions and the extrapersonal space that is too far away for such immediate action. Thus, their experiment provides clear evidence that these mirror neurons do not simply and abstractly encode action representations in an agent neutral and context independent format.

Another interesting finding is that monkeys do not respond very well to televised actions. The difference between live and televised action perception was observed from the very early studies of mirror neurons in macaque monkeys, and raise a question as to whether monkeys have a problem with pictorial perception of television sets or whether the televised actions are not pragmatically relevant enough to modulate the monkeys' mirror neurons. I propose that many premotor mirror neurons might only respond to and monitor actions within a spatial or abstract zone that is deemed relevant to the concrete

[86] Caggiano et al. (2009).

spatiotemporal affordance structure. On this hypothesis one should expect that monkeys and humans in particular can learn to see televised actions as relevant. This hypothesis is backed by the general finding that the fronto-parietal circuits and motor evoked potentials can be modulated by televised action perception, and a study by Shimada & Hiraki providing evidence that human infants respond differently to live and televised actions.[87]

3.2.4. Why an equivocal understanding of what is mirrored?

Given the findings of Umilta and her colleagues, one can conclude that it is often not the visual perception of the action in and of itself but the perception or inference of the others' engagement with the target object, which modulates many mirror neurons. The findings of differential modulation of mirror neurons depending on the relative location of observed action suggest that some mirror neurons monitor the observed action not only as it relates to the target but also as it relates to the perceiver as an agent of possible future action. I hypothesize that mirror neurons contribute to an affordance space monitoring function that does not simply mirror the actions of others but helps us anticipate change and choose and prepare our own actions.

One might ask: What if anything is mirrored by mirror neurons? But maybe this is the wrong way of posing the question as various mirror neurons do indeed seem to respond to various rather specific elements of actions and levels of goal generality, some to different perceptual modalities, some to occluded interaction inferences etc. Maybe rather than coming up with a strict definition of what all mirror neurons - i.e. action sensitive visuomotor neurons - do, one should rather try to specify the spectrum of preferred modulation of various neurons. And maybe the findings that the modulation of mirror neurons are distributed over a spectrum of preferred goal-directed action elements and are modulated by pragmatic and social contextual factors are essential to understand the functional role of these sensorimotor areas. Perhaps the very variation serves an important cognitive function that is overlooked when thought of as a homogeneous group of action observation-execution matching neurons, with one homogeneous name, and a homogeneous function of mirror matching. But Gallese and Rizzolatti and their coworkers have seemed rather determined to think of mirror neurons as instantiating a specific mechanism, and in many ways determined to think of their function in isolation

[87] Shimada & Hiraki (2006).

from the general processes of the premotor and later parietal areas, in which they have been found. That specific mechanism was thought in the 'perceptual mode' to be modulated specifically by the actions of others.

A. 'Agent neutrality' and perceptual modulation by other's actions

One of the key features of the mirror neuron discovery was, as mentioned, the astonishing fact that these premotor neurons where modulated not only by the monkey's own actions but also by the simple perception of actions performed by others. The question, is how the Parma group actually knew that these neurons were truly modulated by the action perception and not something else? After all we are talking about single premotor neurons very far away from primary visual areas. As we have already seen various issues can be raised in regards to how exactly mirror neurons are modulated by external actions performed by others and which elements of the goal-directed actions modulate various action sensitive neurons.[88] Here, I would like to discuss the evidence that many of these visuomotor neurons during perception truly seem to be modulated by the judgment of the external action and not simply by initiations or preparations of the observer's own actions. The question is what the evidence is that these premotor neurons are modulated by external actions and not simply by the monkey's own internally generated plan to execute a similar action. I shall discuss this issue in some detail as it also raises the further issues of 1) why the perceptually induced motor activity does not in normal cases lead to overt action imitation in the observer and 2) how this perceptually induced motor response is understood as pertaining to the action of the other rather than the observer's own action preparation. In other words it serves as a wedge into the big theoretically important question of how motor processes can serve perceptual functions at all, and whether these processes must in some sense cease to be motor to be truly perceptual, or

[88] The discussion of how exactly mirror neurons are modulated by the external actions is also related to the issue of whether this process is best thought of as direct or mediated by various interpretative or central processes. By way of the studies by Kohler et al. 2004 and Umilta et al. 2005 I have discussed how the concrete visual perception of actions often does not seem to be necessary for mirror neuron modulation, but that this modulation often seems to depend on our broader knowledge of the context and its objects. In Chapter 5 I discuss Csibra's arguments, which more radically challenge the existence of low-level/ unmediated perception-action mirror matching. He suggests that the perception is always already interpreted by purely perceptual processes to determine the level at which to engage the motor system. The present point is that sensorimotor modulations often do not reveal the precise action-perception relation.

whether - as I shall suggest - it is possible for sensorimotor processes in some sense to serve both action and perception at once.

As we shall see in the upcoming discussion motor processes are thought to be used 'off-line' when serving perceptual functions – and thus in this 'receptive mode' motor representations are thought to cease to inform ongoing action and instead contribute to our perceptual understanding of the actions of others. I call this the idea of 'motor cognition as moonlighting for the sake of perception.'

To make sense of the possibility that sensorimotor processes in some sense serve both action and perception at once it is pivotal to let go of certain in my eyes wrongheaded tacit assumptions about motor cognition. I argue that it is essential 1) not to see all motor processes as having to do with fully specified actions, and 2) not to think of motor activity as limited to the preparation of one action at the time but to see the important function of the massively parallel sensorimotor activity. I suggest that a lot of motor cognition has to do with entertaining multiple simultaneous sensorimotor based rough sketches representing teleological relations and affordances. I have already touched on these ideas in regard to the described findings of widespread parallel sensorimotor integration in canonical neurons and in the fronto-parietal areas in general. However, in regards to interpreting mirror neuron findings these key findings seem largely to be ignored, and in my eyes this lack of attention to the broader neurological context repeatedly leads the relevant interpretations astray. With these theoretical objectives in mind, I will now look at the research and the proposed interpretations in regard to this matter of the perceptual nature of the mirror neuron responses in the 'receptive mode.'

B. The false dichotomy between perceptual and motor functions

I agree that the discovery of goal-directed action sensitive neurons in F5 really did emphasize that this premotor area has other functions than the typical action planning and initiation 'output' functions, which are normally attributed to the pre-motor cortex. The finding of canonical neurons had already pointed to potentially perceptual functions of object affordance representations not only for objects in the peripersonal space, but also beyond that space. In other words, there was evidence of monitoring of more distal affordances probably for purposes of more distal action planning, but hereunder also providing a pragmatic understanding of the perceptual environment even beyond one's own concrete or immediate action decisions. One can therefore say that already the finding of canonical neurons in some way undermined the traditional pure output conception of the motor system. Why would the motor system monitor and sketch multiple potential actions if this was a system that in general was only activated after action choices were made and effective intentions were formed? To me the parallel object affordance representations even of objects in the extra-personal space clearly suggest that these sensorimotor integrations play a role in understanding the pragmatic relevance of the perception and in shaping intentions and guiding the choice of action. In the case of canonical neurons a perceptual function is suggested – but a pragmatic perceptual function that is still thought to inform action choices and planning. With the finding of mirror neurons this pragmatic perceptual function could be expanded to include something like a relative 'other-centered' goal-directed action monitoring-anticipation-understanding function. Thus, the overall affordance space might be ego-centric, but with a level of sophistication that allows us to represent and to some extent anticipate the causal behaviors of objects and the engagements of other people. For the anticipation to be accurate and thereby useful for our own future action planning (ego-centered), we need some level of understanding of the lasting causal properties and affordances of objects and the informational and motivational perspective of others (other-centered) in this affordance structure. Consequently, the idea of ongoing and anticipatory social affordance space monitoring functions suggests the pragmatic nature of such other-centered representational processes not only for perceptual understanding but also for ego-centered purposes.[89] More on this later, but here the important point is that this is not how mirror neurons have typically been interpreted. And one reason for this lack of focus on the pragmatic role in regard to the perceivers own action is that the discovery of mirror neurons in a way reverted the terminology and the debate back to a strict functional dichotomy between perceptual and motor functions. In other words, the question that was posed was whether mirror neurons are modulated by the current action of the other or the current action plan of the perceiver. But this question might be too simplistic as it excludes the possibility that mirror neurons could be modulated by the

[89] De Vignemont (2008) presents an interesting analysis of the distinction between an egocentric 2nd person perspective and an allocentric 3rd person perspective in social cognitive processes. Without getting into any details, my view here is that the distinction is important but might in many cases be blurred.

current action of the other but as it pertains to my possible future actions. However, given the false ultimatum between my own and the other's present action it does seem understandable why the Parma group interprets mirror neuron activity during perception as other-centered action representations unrelated to the observer's own action preparations.

Like in the earlier di Pelligrino study, Gallese et al. 1996 specifically tested that mirror neuron activity cannot be explained away as simply being a product of the monkey's own motor preparation to perform a similar act. They wanted to make sure that the premotor activity was a product of the action perception and that it could not be linked simply to the monkey's concrete spatiotemporal preparation to perform the observed action itself. To prove this point and rule out the possibility that the monkey was simply preparing its own action, various control studies were performed in parallel with action observation studies. Gallese and his coauthors report that EMG recording control experiments were performed and that no heightened muscle activity was found during action observation. They take this as evidence that mirror neuron activity really is a product of perceptual stimulation rather than a preparation to act. That might be so, but one should note that this finding of the independence of mirror neuron activity from muscle tension raises some questions in regard to the nature of 1) the positive relation between mirror neuron activity and overt action and 2) the idea of mirror neurons as having a symmetric and agent-neutral response during action observation and execution.

C. The sub-threshold theory versus agent-neutrality

The idea of symmetry and agent-neutrality is at the core of the caricature model of mirror neurons, but what we might call the 'sub-threshold theory' of mirror neuron activity seems to pose a pivotal challenge to this idea. Further, as we shall see, muscle activity and motor facilitation during action observation is precisely taken as evidence for a homologue human mirror neuron system parallel to that found in monkeys.[90] Gallese and his coauthors acknowledge that there might be a 'contrast' 'at least at first glance' between motor facilitation studies (Fadiga et al. 1995) used as evidence *for* a human

[90] Gallese et al. 1996, p. 603. I return to these findings in connection to the discussion of the evidence for the human mirror neuron system, where muscle tension and facilitation are often referred to as evidence in support of a human observation/execution matching mechanism in general or mirror neurons in particular.

mirror system and the lack of primary motor cortex activity during action observation they report. However, they suggest that both findings could be due to "a sub-threshold activation of primary motor cortex,"[91] which again is a product of pre-motor activity, and that the seeming inconsistency then disappears. The sub-threshold activation story might be correct. After all, it is clear that even *within* area F5 not all the premotor neurons that are active during action execution have 'mirror' properties – the numbers that I have reported above repeatedly confirm this. Thus, though there are 'visually dominant' premotor neurons overall there are fewer neurons in this area that are modulated by the perception than the execution of a given goal-directed action. Additionally, although it is not explicitly quantified and by some suggested to be statistically insignificant, looking at the histograms of individual mirror neurons; they generally seem to show a weaker modulation in response to action perception than action execution.[92] The robustness of such issues need more focal empirical attention, but for now one can merely point out that all these findings *would* fit with a sub-threshold explanation in regard to the lack of muscle activation during action observation. That is, it would fit with the idea that action perception induced mirror neuron activity is not enough to produce an overt action response. However, these indications of a weaker modulation during observation than execution is very rarely mentioned or discussed. Why? One obvious answer could be that it is because it does not fit very well with the simple summary of mirror neuron activity as being 'the same' for action observation and execution. In other words, if one starts by assuming that mirror neurons instantiate an agent-neutral and symmetric mirror mechanism that by the core definition produces an identical response during action execution and action perception, then the sub-threshold activation is not really tenable. By looking less into the details of the findings another story has, as we shall see, made it to the center of the mirror neuron stage: what one might call the 'inhibition-who system' theory.

When mirror neuron activity is explicitly hypothesized as an automatic and agent neutral mirror 'mechanism,' the lack of overt action in the case of action perception induced premotor modulation is normally explained not as due to sub-threshold activity, but

[91] Gallese et al. 1996, p. 603.
[92] See Chapter 2 for the quantitative findings and Rizzolatti et al. (1996) for an example of one of the histograms of a typical strictly congruent mirror neuron.

rather attributed to *a separate inhibition process*. The point is that by positing an *external* inhibition process one can maintain the assumption that 'from within' the mirror mechanism there is no difference between the executed action of oneself and the perceived action of the other. The mirror mechanism is, by way of the observation-execution matching, thought of as symmetric and thereby agent-neutral, and it is exactly this agent neutrality, which is thought to be important for the social cognitive function of creating a shared action representation between self and other, or to some extent a 1st person-like glance into the mental world of the other normally perceived only from the outside as a 3rd person. I shall argue that the idea of other minds as radically 3rd personal and hidden behind the overt behavior is riding on a too strict dichotomy of inner and outer, which is implausible if sensorimotor mental representations are inherently relational and teleological.[93] As I see it, the sub-threshold theory – as opposed to the caricature mirror theory plus an 'inhibition-who system' - could fit with such a relational and asymmetric account of fronto-parietal sensorimotor integration. The sub-threshold theory could still support a weak sense of observation/execution matching in that a *rough representation* of a particular goal-directed action could be 'retrieved' either by external perception of the action in question or when the action is chosen for execution. However, the process of executing a goal directed action and specifying it in relation to a spatiotemporal sensorimotor context is different from simply *recognizing* that the action is being performed by another or in the case of canonical neurons that a goal-directed action is afforded by something in the environment. Consequently, one can on this story expect an overlap but not a broader identity or strict symmetry of the neurological activity during observation and execution

D. Inhibition of observed actions and the 'who' system

If the idea of shared representations and in particular agent neutrality is seen as the key to social understanding, and thus the social importance of mirror neurons seem to ride on the indistinguishability of the response during action observation and execution. A lot is therefore at stake theoretically. Mirror neurons have already themselves been

---

[93] I have already pointed to various evidence for such relational, asymmetric and teleological sensorimotor integration and I shall in Chapters 4 and 5 extensively discuss the notion of social cognition as 3rd person mind-reading and how a social affordance account would differ from this.

conceptualized as responding 'the same' during action execution and observation. Now this feature of identical response during observation and execution is applied to the broader mirror matching 'mechanism.' This is a very important detail as it also exports the idea of relative anatomical and functional modularity of single neurons to a larger mirror mechanism that presumably it dispersed over many thousands of neurons and multiple cortical areas. Accordingly, the mirror mechanism is thought of as relatively uniform, modularly encapsulated and as producing an agent neutral response to the perception and execution of action.

Given this notion of the mirror system it is inferred that these neurons' externally triggered activity must be *inhibited* in order for them not to cause an automatic overt imitation. Further, the agent neutrality of the mirror mechanism activation implies that our knowledge of an action's agent must come from a separate mechanism outside of the mirror neuron system. The assumed modularity and agent neutrality of the mirror neuron mechanism therefore not surprisingly often lead (especially more theoretically inclined) cognitivist researchers to theorize two supporting external mechanisms exactly for these purposes: 1) A 'who' system that determines the agent of the action, whether it is oneself or another,[94] and 2) an inhibitory system to ensure that actions deemed to be others' are not automatically overtly executed.[95] Any reader of the mirror neuron literature is familiar with the discussion of these two hypothetical mechanisms. However, I want to stress that if the sub-threshold theory is right then these extra mechanisms might be obsolete in regard to the purposes they were originally proposed to serve. In other words, in regard to answering the questions of 1) how we know who is the agent of an action and 2) why we do not openly imitate all the action we see. The alternative sub-threshold theory exactly suggests that the activity is *sub-threshold* for action and that the activation of premotor and other mirror areas and even within the single mirror neurons themselves are in fact NOT identical when an action is executed and observed. One should note that this of course also implies that mirror neurons are not totally 'mirroring' perceived

[94] See here for example: Georgieff & Jeannerod (1998) and de Vignemont & Fourneret (2004). Additionally it is very interesting that Susan Hurley who has been so critical of the 'classical sandwich' of cognitivism still buys a lot of the traditional framework for motor cognition and hereunder also the idea of a separate 'who' system. See her comprehensive and sadly posthumous BBS article: Hurley (2008).

[95] Interestingly, Iacoboni has in his 2008 book *Mirroring People: The New Science of How We Connect with Others* been gesturing at a broader mirror theory that includes 'super mirror neurons' in medial areas that might be important for this often hypothesized inhibition mechanism.

actions but are also only part of the story during action execution. The sub-threshold theory questions not only the need for external automatic inhibition and agent identification systems, but also, as already mentioned, the implicit fable of a relatively modular and isolated mirror mechanism.

Gallese and his colleagues often do take the integration of mirror neurons in a broader functional context for granted. For example, as we shall see in the next section, when they interpret mirror neurons as extending a broader extra-mirror neuron action knowledge/know-how to the case of action perception, it is clear that what gets extended goes beyond the mirror neuron activity as such. I don't actually think that Gallese or Rizzolatti would explicitly support a modular agent neutral mirror mechanism. But this is exactly why it seems so important to bring the tensions between the various assumptions of the mirroring theory out in the open, and thus push the Parma group to explicitly deny the caricature mirror theory that is perpetually reinforced at least on a tacit level by nearly everybody in the field.

These issues regarding sub-threshold activations, symmetry and agent neutrality are also important to me as I want to propose that there could be an important difference between some local premotor functions like for example mirror neuron activity and fully spatiotemporally simulated actions. The big question is whether 'sub-threshold' sensorimotor activations could still have important functional roles to play in our essential cognitive processes. My claim is yes, and I take the existence of F5 canonical neurons to provide clear support for this point. As observed, premotor canonical neurons seem to monitor object affordances even outside our immediate reach. In other words, their activity do not lead to a direct action response of the kind they are linked to during action execution. The question is, on the one hand, whether we fully simulate each of these potential actions in their exact spatiotemporal detail, and, on the other, whether we actively have to inhibit each perceived affordance not to perform the solicited action. It seems very unlikely that this is the case as we presumably must monitor many different object affordances simultaneously to make informed action choices.[96] Thus, it seems that in the case of perception of object affordances we can have motor processes that have important cognitive functions, but which do not involve full covert actions that are

[96] See for example Ellis and Tucker (2000) and Cisek (2007). I shall discuss Cisek's article and proposals in more detail later.

indistinguishable from actually self executed actions.[97] I shall discuss my theory of motor cognition beyond full covert simulations in greater detail in the subsequent chapters.

E. Action 'monitoring' and the neglected timing findings

Let me now return to Gallese and his colleagues' immediate point that the lack of muscle activation during action observation might provide some evidence that the mirror neuron activity is not just a product of the monkey's independent plan to execute the action itself. As mentioned above, they present this as evidence supporting that the observed premotor activity is actually a product of action perception and not action preparation. However, this is not very definitive and might point to as many new questions about the positive relations between mirror neuron activity and action preparation and initiation as it answers.

Proponents of a mirror mechanism would not want to argue that mirror neuron activity is unrelated to action preparation, as they exactly want an overlap between the neural circuits recruited in both action perception and execution. Hence they need to show on the one hand that there is an overlap between neural activity during action and perception and yet at the same time that the activity during action perception is truly caused by the perceived action and not a parallel endogenous action initiative. In other words, what needs to be shown is that mirror neurons really can be modulated by perception of others' actions, rather than always being mediated by an independent endogenously initiated action preparation process.[98] I think that important evidence that mirror neuron activity can be modulated in this externally induced way can be found in the timing data of mirror neuron activity.

---

[97] Both Gallese and other mirror theorists often point to the clinical syndrome called imitation behavior to support the idea of automatic covert imitation and the need for the an extra process of inhibition to avoid spontaneous imitative execution of observed actions. As I shall explain briefly at the end of this Chapter and further in Chapter 4 (Section 4.2.7), they hypothesize the syndrome of imitation behavior (and the related utilization behavior) in some frontal lesion patients simply as a product of 'released' mirror neuron and canonical neuron systems. I suggest that this is a very simplistic and implausible story of the relevant syndromes. For one thing, it completely ignores the fact that these patients in general are very apathetic and overall produce fewer actions than others. A good contrast here is presented in the often ignored differences between these patients with frontal lesions and people suffering from an 'anarchic' or 'alien hand' syndrome due to parietal lesions.

[98] The mirror neuron activity might be related to a *later* endogenous action initiation, but the question here is whether these neurons are really *initially* modulated by the post-hoc observation of the action of the other.

Gallese and his colleagues clearly report that the mirror neurons they studied only fired during the observed or executed movement – and not in between actions or during the preparation to move prior to the executed or observed action.[99] Earlier premotor research – including the work of the Parma group - distinguished between 3 classes of neurons given the timing of their response in various versions of 'delayed response' experiments: 'stimulus-related,' 'movement-related' and so-called 'set-related' neurons.[100] These are respectively premotor neurons that fire at the time of stimulus presentation, at movement onset or during the delay period after the stimulus information indicating the appropriate action but before the movement initiation. I think it is informative to see Gallese and his colleagues' findings in the context of this classification - which was of course made long before much was known about the visual properties of premotor neurons. Most mirror neurons seem to show 'movement' correlated timing, or rather because the effective stimulus for mirror neurons is an action, one could say that they are both 'stimulus' and 'movement related' neurons. The point is that they are <u>not</u> 'set-related', i.e. they are not simply holding on to and preparing a parallel action to that observed for future execution. Rather, the activity of mirror neurons - at least in the case of studied strictly concurrent mirror neurons - seem to be narrowly tied to the timing of the ongoing action whether it is performed by oneself or someone else. These temporal results do indeed indicate that the activity of mirror neurons somehow 'monitors' ongoing actions. A further strong indicator of such an action monitoring role is that they in opposition to canonical neurons do not respond to object affordances. Gallese and his coauthors report:

> …mirror neurons stop firing when the object grasped or manipulated by the experimenter is moved towards the monkey and becomes more available to it. They start to discharge again only when the monkey makes the movement. (Gallese et al. 1996, p.606)

This finding is important. Not only does it relate the activity of mirror neurons narrowly to the timing of the action, but it might also provide the basis for a different distinction

[99] Given the often ignored heterogeneity and complexity of mirror neuron responses, it might be a somewhat unwarranted simplification to say that no F5 mirror neurons, i.e. neurons modulated both by action observation and execution, show any response except during movements. But this does not diminish the importance of this finding in relation to the group of mirror neurons that was found to exhibit this time-locked activation pattern.

[100] See for example the delayed response experiment of di Pelligrini et al (1992) described above. For more elaborate information on the classification of set-related neurons as opposed to stimulus and movement related neurons, see Gentilucci & Rizzolatti 1990, "Cortical control of arm and hand movements," in *Vision and Action* edited by Melvyn Goodale.

between mirror neurons and the above mentioned canonical visuomotor neurons than the one most often explicitly used.

3.2.5. The problematic dichotomy between canonical and mirror neurons

Often Gallese and Rizzolatti and their colleagues simply stress that canonical neurons are modulated by *object* perception whereas mirror neurons are modulated by *action* perception. Thus, the distinction is thought of as 1) based on the kinds of effective stimuli (objects of actions), and 2) as all-inclusively dichotomizing the relevant fronto-parietal visuomotor neurons.

A. <u>Why only two kinds of visuomotor neurons?</u>

However, given the temporal findings just mentioned various distinctions and categorizations of sensorimotor neurons seems to be appropriate. One such distinction might be between visuo- or more broadly sensorimotor neurons in which the sensory modulation precedes and prepares the motor related response and thus might alert to *potential* interactions, and others where both sensory and motor properties seem *time-locked* with the *actual* movement or perception of such. One would in this way be able to make a distinction between on the one hand certain canonical/'affordance' neurons where perceptual stimuli solicit, monitor and prepare *potential* actions, and, on the other hand, mirror neurons in which the perceptual modulation seems to monitor *actual* ongoing actions. Such a distinction would then again set the stage for a much more complex set of functional distinctions between various groups of visuo-motor neurons beyond the simple 'stimulus kind' distinction. The simple chart below is made to illustrate the *present division* between canonical and mirror neurons:

Potentiated action/ event representation / **Effective Perceptual Event/stimulus**	Potential actions Afforded observer by perceived stimuli	Actual/present Goals/Action types
Objects	Classical Canonical Neurons	
Actions		Classical Mirror Neurons

Figure 3A: Typical action-object dichotomy categorization of visuomotor neurons

My hypothesis would be that the complexity of visuomotor neurons goes well beyond not only the object-action dichotomy of traditional mirror and canonical neurons, but also beyond the actual-potential distinction implied by the findings referred to above. For example I would expect that there are different kinds of neurons, which react also to immediate *social affordances* such as an extended hand inviting a hand shake, i.e. a social affordance not to produce the *same* or 'mirror' action but rather a *complementary* action. Such 'social affordance neurons' are a speculative suggestion on my part, but I think a rather modest addition to the existing categories given the heterogeneity already documented in visuomotor neurons' properties. As I read it, the evidence suggests that the properties of sensorimotor neurons in fronto-parietal areas seem to be *much more complex* than what can be dealt with by simply adding one potential category of social affordance neurons. Based on my hypothesis, I contend that a thorough revision of the existing categories is needed. I have already discussed the definition of 'mirror' neurons as pertaining to all action sensitive visuomotor neurons and the sub-categorization in terms of 'congruence' obscures the heterogeneity of the empirical findings. I discussed particularly how the finding of temporally or 'logically' related visuo-neurons got obscured under the congruence based mirror neuron subdivisions. Just adding that group and the hypothesized social affordance category would make the chart above a little more complex:

Potentiated action/event representation / **Effective Perceptual Event/stimulus**	Potential actions Afforded observer by perceived stimuli	Actual/present Goals/Action types	Anticipated actions and events
Objects	Classical canonical neurons		
Actions	*Social affordance "mirror" neurons?*	Classical mirror neurons	"logically connected" mirror neurons

Figure 3B: Expanded categories including action anticipation and affordances.

I argue that the existing categorizations might obscure the affordance aspect of action perception. But that is not meant to say that the time-locked modulation of many mirror neurons does not provide evidence that these support an other-centered action monitoring function. My proposal is rather that this other-centered representation might itself be modulated and related to a larger ego-centered affordance space. In other words, the 'strictly congruent' mirror neuron activity might be other-centered, but not be ubiquitous, modular and agent-neutral but dynamically penetrable and regulated within a broader framework that situates the perceived actions of others within an ego-centric action perceptive. As I see it, the evidence support such a broader dynamical framework for understanding not only the actions of others and their goal directed teleology but also how these actions are situated in the shared affordance space.

I have already mentioned how many canonical neurons are modulated by action affordances also outside of one's peri-personal space, i.e. outside of one's immediate reach. Thus there is here to some degree a monitoring of the affordances, causal and teleological properties of the *object* itself in its present relational place within the broader affordance structure. This means that also within so-called object sensitive canonical neurons one ought to make distinctions between neurons that are more *object-centered* monitoring lasting pragmatic properties and their *actual* location in the broader space as opposed to those that are more directly *ego-centered* monitoring our *immediate* pragmatic relation to the object in question. Unfortunately, I have not seen this sort of immediate affordance versus actual action/object monitoring distinction explicitly made by mirror neuron researchers. Be this as it may, the essential thing to note is that a consequence of such a temporal and 'ego vs. other/allo/object-centered' distinction would be that the existing canonical and mirror neuron classifications would get joggled and rather intertwined. Clearly certain broadly/non-congruent or logically related mirror neurons would fall within various potential affordance and prediction categories and underline also the possible *functional* heterogeneity of mirror neurons broadly defined as action sensitive visuomotor neurons. I have already speculated about the existence of various *new categories* of, for example, social affordance neurons (such as the handshake affordance) and object-centered canonical neurons, which as I hypothesized above might be found if one's theoretical and categorical framework allowed for it. Given the heterogeneity of the data so far, it seems almost inevitable that new properties would be found if the classificatory tools and experimental paradigms were used to ask questions beyond issues of congruence.

It has already been shown that there is a spectrum of ego-centered and object-centered canonical affordance neurons and other-centered action observation sensitive visuomotor neurons. I am now suggesting that given such an array of visuomotor neurons and functions it would be reasonable to hypothesize that there would also be social affordance neurons that respond to socially engaged actions directed right at the observer calling for *inter*action.[101] The main point is that the 'type of stimuli' classifications seem incapable of getting at the functional complexity of F5 sensorimotor neurons. The various combined timing and stimuli and context findings suggest that mirror neurons should not be seen as instantiating a uniform or isolated mirror mechanism that can be identically activated by action execution and observation. Rather, it seems that the strictly congruent mirror neurons work in tight cooperation with wider circuits of sensorimotor integrations that is important for both understanding and acting into a broader affordance space. Thus I suggest that one should not try to conceive of an isolated mirror 'mechanism' but rather rethink the mirroring properties in a wider neurological context and in regard to a wider environmental context. I suggest that in order to understand not only the overall function of the fronto-parietal circuits, but also the potential *social* cognitive function of the strictly congruent mirror neurons themselves it seems essential to think of the *collaborative* functions of various sensorimotor neurons. And by sensorimotor neurons I do not only mean to add typically described canonical neurons to the typically described mirror neurons, but also analyze and differentiate the more complex array of visuo-motor and sensory motor properties of the neurons in these circuits.

The question is whether mirror neurons can be said to function in isolation from the broader affordance structure monitoring and representing functions of the fronto-parietal circuits. I suggest not. It is not very plausible that we understand or predict the actions of others independently of our pragmatic understanding of the scene and objects they are interacting with and looking at. I have already mentioned that a recent finding by Caggiano and colleagues from the Parma group shows that various mirror neurons are modulated *differently* by actions inside and outside of the peripersonal space. This is a clear indication that even on a single cell level all mirror neurons should not simply be seen as mirroring or monitoring the action of the other. These new findings show that already at the single cell level there is an integration of the *other-centered* representation of the perceived action with pragmatic information about how this action is placed in the *ego-centric* affordance space of the observer. Given this sort of evidence I propose an

[101] In the case of social affordances one might say that here it is not merely an action *affordance* but almost an action *imperative*. But that is a discussion that goes well beyond my present project.

'affordance space' theory of fronto-parietal function, i.e. that these sensorimotor neuronal circuits support a social and physical affordance space understanding that helps us predict, guide and choose upcoming actions.

Given this hypothesis one can both understand the hitherto discovered sensorimotor properties but also predict possible other new findings of sensorimotor integration types. The chart below is meant to illustrate possible new categorizations of already found neurons and point to hypothetically suggested neurons that could be found in the future – or might already have been found but slipped through the existing categorization grid. The most basic point is that seeing this heterogeneity and complexity should raise concerns about the idea of a functionally distinct and agent-neutral mirror mechanism based on a very small - and idealized - subset of fronto-parietal visuo-motor neurons[102].

[102] Note that I have not included here the extra complexity of potential multimodal sensorimotor neurons beyond visuo-motor neurons. However, given the already mentioned findings of auditory-visuo-motor 'mirror' neurons I would expect plenty of yet unstudied multimodal sensorimotor integrations.

Effective Perceptual Event/stimulus \ **Potentiated action/event representation**			Potential actions afforded observer by perceived stimuli		"Mirroring" Actual/present Goals/Action types	Temporal/causal event and action anticipation
			Ego-centered	Body-centered Self-other neutral		
Objects (including tools)		Extra-personal space	*"Canonical Neurons" sub-group? ego-object affordance neurons*	*?Body-centered object affordance monitoring neurons*	*? Maybe neurons modulated not simply by present action perception but by affordances turned actual actions. Thus temporally integrating canonical object affordances with action goal monitoring.*	*? Sensorimotor neurons for prediction of changes in extra-personal space & upcoming indirect affordances*
		Peri-personal space	*"Canonical Neurons" sub-group? direct object affordances*	*?Body-centered direct object affordance monitoring neurons*		*? Sensorimotor neurons predicting changes in peri-personal space & upcoming direct affordances*
		Object- centered independent of location	**"Canonical" Neurons Action-object affordances relations**			*? object-centered neurons predicting environmental changes*
Goal-oriented Actions -possibly also abstract goal directed actions such as speech or gestures	Local/ sub-Action goals	Extra-personal Space	*? Ego-centered action/social affordance neurons*	*? Body-centered action affordance neurons*	**Context/space relational extrapersonal "mirror neurons"**	*? neurons predicting actions in extra-personal space (& upcoming indirect social affordances)*
		Peri-personal space	*? direct ego action/social affordance neurons*	*? direct afforded body-centered sub-action goals*	**Context/space relational peripersonal "mirror neurons"**	*? neurons predicting actions in peri-personal space & upcoming direct social affordances*
		Action- centered independent of location	*? Social Affordance neurons Action-action affordance relation*		**Traditional Mirror Neurons**	**"logically connected" action anticipation "mirror neurons"**
	Over-arching Action goals	Extra-personal Space	*?social/action affordance neurons*	*?Body-centered action affordance neurons*	**Context/space relational extrapersonal "mirror neurons"**	*? neurons predicting actions in extra-personal space & indirect social affordances*
		Peri-personal space	*?social/action affordance neurons?*	*?afforded body-centered overarching action goals*	**Context/space relational peripersonal "mirror neurons"**	*? neurons predicting actions in peri-personal space & direct social affordances*
		Action- centered independent of location	*?Social Affordance neurons action-action affordance relation*		**Traditional mirror neurons**	**"logically connected" action anticipation "mirror neurons"**
Other agent-attention/ eye movements		Body centered eye movements	*Sensorimotor neurons coding directional saccade affordances*		*Eye-target relational 'mirror' neurons?*	*Target direction anticipation neurons?*
		Target centered attention	*Sensorimotor neurons coding saccade affordances to target*		*Eye-target relational 'mirror' neurons?*	*Target anticipation neurons?*

Figure 3C: Broad spectrum of hypothesized sensorimotor categories

Bold = neurons already described in mirror neuron literature,
Italics = neuron types hypothesized based on existing findings.

B. The anatomical distinction between canonical and mirror neurons

This analysis of temporal features and the heterogeneity of the properties of fronto-parietal neurons has to a great extent been under-emphasized or entirely missed by the Parma group. They clearly have been interested in making a neat and functional distinction between canonical and mirror neurons. This project stands in clear opposition to my suggestion that the functions of these two groups of neurons should be seen as functionally integrated in a larger teleological ‘affordance space’ monitoring function. Rizzolatti and Fadiga have published an important article in 1998 focusing exactly on the distinction between canonical and mirror neurons. In this article, instead of basing the functional distinction on differences in properties or functional connections, they argued for the separate function of canonical and mirror neurons by way of anatomical evidence that the location of the two kinds of neurons differed and that neurons of each group were more prominent in distinct areas within F5.[103] Through single cell recordings they found that neurons modulated by 3D object perception were clustered mostly in the part of F5 that folds into the acuate sulcus, whereas action perception sensitive neurons were mostly concentrated on the central convexity of F5. (Rizzolatti & Fadiga, 1998, p.85) I think the finding of this anatomical division is important for the argument for a separate mirroring mechanism since it makes it easier to distance the function of mirror neurons from the more general theories of F5 function in visually guided actions.[104] However, later findings complicate this picture of a neat distinction between object and action sensitive visuomotor neurons. The fMRI study published by Nelissen and colleagues in 2005 suggests that there are also action sensitive neurons in the depth of the arcuate sulcus and, interestingly, they seem to respond to very specific actions in a much less contextual way than the ‘mirror’ neurons on the central convexity of F5.[105] Thus, the neat anatomical division between canonical and mirror neurons might be a bit idealized and biased by the

[103] See Rizzolatti & Fadiga (1998).

[104] This division also hints at the many unknown complexities of F5 and unfortunately the division is rarely mentioned and never really discussed or further investigated in other studies.

[105] They write, “The two premotor representations differ in their properties. F5c is active only when the observer sees an action that includes a view of its agent. The observation of a grasping hand alone is insufficient. Note that the mirror neurons were discovered and subsequently studied by testing them with the experimenter in full view (4, 5, 16). The second F5 action representation (F5a), located in the depth of the arcuate sulcus, appears to code actions in a less context-dependent way. The observation of an isolated arm action is already an effective stimulus for this representation. Similarly, the observation of a mimicked action, which is not effective in activating F5c.” (Nelissen et al. 2005)

aim of arguing that there is a clearly demarcated 'mirror' mechanism that has a specific social function of understanding the actions of others by somehow reliving these actions via retrieving the action in one's own motor system. My view is that more anatomically precise research is needed – and again that one should look not only for mirror neurons versus canonical neurons, but rather also look for more complex properties and commonalities of visuomotor neurons within each area. In this way the research might not merely support the traditional mirror theory but further inform our understanding of these areas.

3.2.6. Habituation and the assumed modularity of the mirror mechanism

Gallese and his coauthors write in the 1996 article that the response of mirror neurons were "highly consistent and did not habituate." (p.595) This finding might not seem like a big deal. But I want to dwell on it, as I see this finding of non-habituation as being implicitly used to fortify the assumption that mirror neurons display a simple context-independent, automatic and ubiquitous mirroring function. In other words, it supports what I have claimed to be a tacitly held idea of individual neurons as relatively independent little automatic and mechanistic modules simply triggered 'all or none' by the execution or observation of the appropriate action. I am suspicious of the idea of individual neurons with such an insulated context independent action-type 'mirror' behavior. We have already seen that the mirror metaphor is problematic in that it suggests that mirror neurons are modulated by the precise surface features of the observed movement, and that from even before the discovery of mirror neurons it was clear that the responses of many F5 neurons mostly seemed to be related to the goal and targets of interaction than the exact kinetics of actions. Another feature of the metaphor of a mirror is that it implies a rather direct, context independent and ubiquitous process. No one is trying to argue that the narrow issue of habituation settles or proves these larger mirror features. But it implicitly might serve as support for the very metaphor of mirroring, i.e. the imagery of an automatic modular ubiquitous mirror response. I shall argue that non-habituation would in certain ways challenge the caricature model. And additionally I think it is worth discussing the non-habituation finding simply because, if it is a robust finding, this would be a rather extraordinary quality particular to mirror neurons as sensory and sensory-motor neurons normally do habituate. As a matter of fact visual habituation is such a consistent phenomenon that it has long been used as a reliable parameter to test novelty.[106] Habituation, adaptation and repetition suppression phenomena are taken as facts by many researchers and are used as a methodological paradigm in fMRI studies to probing whether a stimulus is regarded as the ‘same’ or ‘different’ by various neural populations. This methodology has in later years also been found to apply to motor areas and sensorimotor processes. For example, Scott Grafton’s fMRI lab use the ‘repetition suppression’ phenomenon as the basis for experimental protocols probing into the visual properties of the various parts of the ‘action observation network’ – which is their term for the exact circuits of the premotor-parietal areas thought to host the human mirror system.[107] One might suggest that these studies are due mostly to habituation of sensory neurons, but Hamilton and Grafton have a new fMRI study in press showing repetition suppression “across the motor network” already at the second trial in which an action was performed.[108] Even Rizzolatti himself happily refers to these repetition-suppression studies as evidence for the goal coding of the mirror system. (Rizzolatti & Fabbri-Destro 2008, p182)[109]

Due to this tremendous inconsistency between the conceived view of perceptual adaptation and the absence of mirror neuron habituation found in the Gallese study, one could certainly wish that there were more studies replicating the finding. But rather than expressing surprise regarding the lack of habituation, Gallese and his coauthors later in this very article seem to simply dismiss some possible evidence for habituation. They write:

> The responses to meaningful objects like food were the same as those to three-dimensional solids, the only difference being the constancy of the responses. It is likely that this was due to the fact that the monkey tended not to pay attention to

[106] For example Dinstein and coauthors write in their 2008 article ‘A mirror up to Nature:’ “A common method for assessing neural selectivity using fMRI takes advantage of the fact that sensory neurons adapt/habituate when their preferred stimulus is presented repeatedly.”(Dinstein et al 2008b) See also Grill-Spector (2006).

[107] For Grafton’s labs’ use of the repetition suppression paradigm, see Hamilton & Grafton (2006). For articles showing goal rather than kinetic movement habituation in infants see: Woodward (1998). For habituation in monkeys, see Desimone (1996).

[108] A longer quote: “The results of this study clearly demonstrate that suppression of the BOLD signal occurs across the motor network when an action is repeated for a second time. This effect was observed throughout the motor system, including primary motor cortex, premotor cortex, supplementary motor area parietal cortex and cerebellum.” (Hamilton & Grafton, 2009 p. 4)

[109] A new study based on repetition suppression has also just come out, which questions the very existence of mirror neurons in humans. See Dinstein et al. (2008a). I shall discuss this later.

> uninteresting objects after a few or even the first presentation. (Gallese et al.1996, p.605)

They suggest in this way that the responses of the mirror neurons themselves are independent of the nature and task relevance of the object acted on, and further that sometimes mirror neurons stop firing due to lack of visual attention to the stimulus, i.e. the action on the object. And in this way the integrity of the tacitly operating metaphor of an automatic mirror survives, as it is claimed that the lack of response is due to the fact that the stimulus is, so to speak, not really in front of the mirror. Now it is of course very possible that they are right in this diagnosis and that there would be less attention to actions directed at geometrical shapes than to food. But, attention is not a simple neurological phenomenon and it is at least also possible that one element of what happens is that mirror neurons like most other visually sensitive neurons quickly habituate to repeated stimuli – with the exception of special cases of for example food where the response is particularly motivated and extra attentional resources might be engaged. Also it seems that attentional regulation of where to direct perceptual explorative actions is an integral part of the very fronto-parietal processes of which mirror neurons are part. In other words, it might be the continued response rather than the habituating one that needs an attentional explanation.[110]

All these suggestions are speculative and there are other possibilities than those proposed here. But this is exactly the bigger point I want to get to with this discussion: it seems to

---

[110] One could hypothesize various explanations, which of course would have to be tested. I propose that mirror neurons play various roles in monitoring and predicting ongoing actions within the affordance space around us. Mirror neurons and canonical neurons contribute to the ‘permanence’ of our experience of the environment around us, but also to our actions choices and to the specification of our actually chosen actions. It would come as no surprise if these sorts of processes are tightly linked with processes of attention and saccade eye movement modulation. Habituation might be reserved for processes of discovery and monitoring where no further attentional exploration is thought necessary. (This would seem to fit with perceptual/attentional phenomena like the well studied inhibition of return (IOR) effect for example.) In general, I would expect the habituation to depend on the role the stimulus in question plays in the further action process, i.e. whether one just needs to notice that it is there (objects)/it has happened (actions) or whether this stimulus still needs to be explored or used for online sensorimotor guidance etc. Furthermore, it has been found that parietal mirror neurons generally need foveal presentation of stimuli to respond whereas premotor mirror neurons respond to extra foveal and thus relatively unattended actions. “An interesting aspect of IPL mirror neurons is that the intensity of their response, or even the response itself, is rather frequently conditional upon the monkey’s active observation of the experimenter’s motor act, while this phenomenon is not common in F5. It appears therefore that while in the premotor cortex the peripheral vision of a motor act is sufficient to trigger mirror neurons, a considerable proportion of those in the IPL require foveation.” See Rozzi et al. (2008). Also, it could be that F5 neurons actually do not habituate, but even so, the mirror mechanism as such would probably not be free of habituation since parietal neurons seem to habituate.

me that the experimental protocols and data interpretations of the Parma group in certain ways assume a localized, modular and non-contextual mirror mechanism – rather than truly testing it. Their empirical conclusions are therefore to some extent included in the premises of the experimental questions. In this section we have seen that it seems that Gallese and his collaborators might be a little too eager to conclude that the mirror neuron response is highly consistent and does not habituate. An obvious explanation for their excitement with this finding is that it fits nicely with the concept of 'mirroring' and the abstract model of these action sensitive neurons as simply providing an automatic and ubiquitous mirroring response to the observed action completely context, task and motivation independent. I should note that I have never seen anybody explicitly using the non-habituation finding to justify the mirror mechanism, but my point is that it nonetheless plays a role implicitly in justifying the preferred experimental protocols used in the studies of mirror neurons. One can get at the importance of this issue by construing the point negatively: What would it mean for mirror neuron theories if these neurons did indeed habituate? I speculate that it would be more difficult to conceptualize their function on the basis of individual neurons in response to individual action exposures. Without the 'highly consistent' mirror neuron response one would have to always interpret the neuronal response not only in regard to the present stimuli but also in regard to prior stimuli and possibly a larger event and neurological context. This is to say that the consistency of the response along with the narrow link to the stimulus type is used tacitly to justify the typical non-contextual and often non-engaged experimental protocols used for mirror neurons research. I shall discuss these research paradigms and their theoretical assumptions to greater length in connection with the behavioral evidence (Section 3.5.2) but before then there are some more research and theory expansions to be presented.

3.3. The functional role of the action mirror mechanism

With all these tensions in the empirical findings and the theorized structure of the mirror mechanism in mind, we can now finally turn to the issue of function. I have purposefully gone over these basic issues first as they provide pivotal information for an assessment of both the standard interpretations of the social function of mirror neurons and the various later and alternative proposals.

Armed with definitions of mirror neurons and canonical neurons in terms of their effective stimuli, and the finding that they mainly occupy neighboring rather than overlapping areas of the premotor cortex, one might say that the ground was laid for an important functional distinction. Having put in a conceptual and to some extent an anatomical wedge between mirror neurons and canonical neurons, and between mirror neurons and action planning and initiation, a functional differentiation is also made between the 'expressive mode' and the perceptual 'receptive mode' function of action representation in F5.[111] The idea is that mirror neurons might be involved in action initiation and organization in the 'expressive mode' but have a rather different function in the 'receptive mode' when modulated by action perception. In other words a split between the overall functions of F5 and premotor areas and mirror neurons is introduced. Combining this with the focus on the '*congruency*' or execution-observation overlap of mirror neurons it is not a great leap to the idea of a mirroring 'mechanism', which in its perceptive/receptive mode can retrieve action representations for the *epistemological* purpose of action *understanding* rather than the *pragmatic* purpose of action *production*. Motor processes can in this way be thought to have perceptual and 'extra-motor' functions, and thus the idea that perceptual processes are always fully perceptual and that the motor system is a pure output system is clearly challenged. That being said one should note to the contrary that with the sharp expressive/receptive mode dichotomy a neat dualism of motor versus perceptual *functions* and their respective cognitive inputs and outputs is maintained. There is still an intact notion of separate perceptual and motor systems and accordingly between perceptual and motor representations. As I have already tried to argue in various ways, I think it is a mistake to think of sensorimotor processes as simply performing a *translation* between perceptual and motor representations.
Such integrative processes, rather than merely translating, seem to me to do an original constructive job of their own. In regard to the fronto-parietal circuits, I suggest that integrations *re*-present (in the sense of present again) learned goal organized sensorimotor engagements at certain levels of abstraction from actual kinetic movements and perceptual stimuli. These sensorimotor re-presentations then allows that new perceptions can be interpreted teleologically via earlier action experience and further that actions performed earlier can be initiated endogenously via sensorimotor imagination and

[111] The terminology of 'receptive mode' and 'expressive mode' is from Gallese (2001). See Section 4.3.4.

redirection of perceptual attention. I shall further develop this notion of sensorimotor re-presentation in Chapters 5 & 6 after the introduction of the notion of simulation in Chapter 4. Here the basic point is that the sensorimotor integrations seem essential to the *existence* of these teleological representations in the first place and not simply a bridge between *two formats* of mental representations (See discussion in Chapter 2). And, now I argue that if we think of the sensorimotor integrations as essential to both intentional action and goal perception, then it does not seem a far-fetched suggestion that the same sensorimotor process might be serving both functions of action and perception *simultaneously*. Thus, my view opposes the Parma group's at times explicit and at others implicit distinction between receptive and expressive mode functions of mirror neurons.

In the 1996 article Gallese and his coauthors ponder the function of mirror neurons and suggest that the visuomotor action representations of mirror neurons "may have different complementary functions, among which motor learning and the understanding of meaning of the observed action."(p.606) Thus, they here opt for the multiple functions rather than merely one key function. As we shall see in the following, their main focus is clearly on the possible role of mirror neurons in regard to action understanding.
But they also suggest that mirror neurons and the observation-execution matching mechanism could possibly play an important role also in social cognitive functions such as action imitation. I want to start by dwelling a little on this somewhat paradoxical issue of the relation between mirror neurons and imitation.

3.3.1. The imitation issue

The idea that mirror neurons during action perception automatically retrieve a representation of the action in question is at the very core of the observation-execution matching mirror mechanism theory. Accordingly, imitation seems to be an obvious functional candidate for this cognitive mechanism. Mirroring is even sometimes described as a process of 'covert imitation' or as we shall see later as 'embodied simulation.' Further, as I have mentioned, it is even suggested that we need to postulate an extra inhibition process that works in conjunction with the mirror mechanisms such that we do not simply spontaneously imitate all the action that we see. Thus, given the caricature mirror story – one might say that mirroring *is* imitation minus execution.

One should note however that Gallese and his colleagues go beyond this caricature imitation reasoning and speculate further how the imitation process could be helped by mirror neurons. They write:

> Such a mechanism can, on the one hand, extract the essential elements describing the agent of the action (hand, arm, face) and, on the other, code them directly on specific sets of neurons with motor properties like those of F5 'motor vocabulary.'(p.606)

This proposal clearly rests on the idea of a motor vocabulary that is addressed by perception. The suggestion of mirror neurons 'describing the agent of action' fits with some findings indicating that F5 has a rough somatotopical organization.[112] Mirror neurons might also be thought to 'extract the essential elements' of an action, in the sense that they seem to be modulated by action goals rather than the exact kinematics, and in the sense that their visuomotor properties in accordance with the earlier F5 findings are distributed to pertain to the various elements of the action at various levels of description. Be this as it may, the reason I quote this description of how mirror neurons could help imitation is that it points to a more complicated story than simply 'automatic covert imitation.' The broad spectrum of action sensitive neurons described would play complex roles and there are therefore *interpretive* choices/categorizations made and there is an *orchestration* to be done to succeed in the complex task of imitation even *if* one has one's mirror neurons in place. In other words, correct imitation is not guaranteed simply by a dis-inhibition of the mirror neuron system. As I have already discussed in relation to the sub-threshold theory, mirror neuron perception induced activity might fall short of the activity during overt action choice and specification. Certain mirror neurons seem to represent other-centered actions within a larger ego-centered affordance structure. Given these sorts of findings and the 'affordance space' story that I propose there would be an extra move from having *a schematic representation* of the other-centered goal-directed action to letting the other's detailed action sequence *determine one's own overall*

[112] For somatotopical organization of F5 in monkeys, see Gentilucci et al. 1988, Rizzolatti et al. 1988. Otherwise, the human fMRI study by Buccino et al. 2001 is most commonly referred to as evidence for the somatotopical organization of mirror neuron areas. (Buccino et al. (2001)) But, in a very recent article Leonardo Fernandinoa and Marco Iacoboni argue that the premotor body maps are at best very coarse and in general these areas are organized around personal goal-directed motor repertoires rather than one central body map. (Fernandino & Iacoboni, 2010). The emphasis on functional organization and multiple very rough and idiosyncratic somatotopical representations fits better with my proposal as I generally what to criticize the 'one body in the mind' conception of the motor system that is assumed by both traditional cognitivists and many contemporary researchers with an embodied approach to cognition. The latter group often assumes a rather rigid body map in the motor system and then tries to prove the importance of such bodily representations in various higher cognitive functions such as for example linguistic processes.

sequence of action choices. In short, one would often have to have *chosen* to imitate and to ignore the other action affordances that the ego-centric meeting with the environment produces.[113]

This leads us to what has been perceived as the grand paradox of the functional role of mirror neurons in imitation. On the one hand, the whole mirror matching process is often referred to as a sort of 'covert imitation' of the observed action. On the other hand, it is controversial whether macaque monkeys in which mirror neurons where initially found are actually capable of 'true' (overt) imitation at all.[114] I will later say a bit more about the fascinating issue of the differences between human social cognition and the social abilities of various non-human primates. For now we can see why mirror theorists would want to focus on a primary function of mirror neurons *other than* imitation, in particular when using the neurological evidence from mirror neurons in macaque monkeys. The point is not only that there are reasons to doubt that imitation is the main function of mirror neurons, but more fundamentally that the conception of a mirror mechanism in terms of 'covert action imitation' might be difficult to uphold given the findings concerning the cognitive sophistication and difficulty of acts of imitation. However, we also see that given my interpretation of mirror neurons as functionally (and anatomically) integrated in a broader ego-centered affordance structure monitoring process, the paradox pretty much disappears. Monkeys might have mirror neurons and be able to understand, track and predict the actions of others without having the ability – or the motivation – to choose to copy these actions in their detail. Whereas mirror neuron activity seems to depend on an automatic perceptually induced activation, imitation depends on a 'top-down' choice to attempt to reproduce the action of another.[115] It could be that monkeys are simply not able to ignore their ego-centric affordances or to take details of action sequences as action goals. But further they might more fundamentally not be able to see the ego-centered value of a temporary abandonment of ego-centered evaluation and choice of their singular movements. My point is that imitation depends on some level of

[113] I shall discuss this issue more in Chapter 5 as Csibra suggests that all imitation is emulation and that the relevant level of goal description must be understood and chosen prior to the engagement of any motor process hereunder mirror neurons. My story will be somewhat different as I operate with a more complex notion of motor cognition that distinguishes between schematic parallel sensorimotor activity as seen in mirror neurons and fully spatio-temporally specified simulated or executed actions.
[114] For a very interesting relevant discussion, see Lyons et al. (2006). And also Iriki (2006).
[115] See here also my discussion of Csibra's mirror neuron critique in Chapter 5.

trust in the value of the action choice of the other, and is often also directly motivated by the goals of establishing the mutual togetherness feeling that comes from this gift of trust in and attention to the actions of the other.[116]

3.3.2. Action and intention understanding

Although imitation is mentioned as a possible function it is clear already from the early articles that the Parma group focuses on the suggestion that mirror neurons play a role in action perception and understanding. As already mentioned, they think that mirror neurons and the mirror matching mechanism might have multiple complementary functions but mostly write about the possible role of mirror neurons in intentional action understanding. In the 1996 article Gallese and coauthors write:

> By this term [understanding] we mean the capacity to recognize that an individual is performing an action, to differentiate this action from others analogous to it, and to use this information in order to act appropriately. Self-consciousness is not necessarily involved in these functions. (Gallese et al. 1996. p.606)

The claim seems to be that we recognize action types of others by mapping them onto our own motor repertoire. Interestingly, given my concerns, they also suggest that this tacit sensorimotor knowledge can then be used to inform our own action choices. They continue their explanation as follows:

> When an individual emits an action, he 'knows' (predicts) its consequences. This knowledge is most likely the result of an association between the representation of the motor act, coded in the motor centers, and the consequences of the action. Mirror neurons could be the means by which this type of knowledge can be extended to actions performed by others. When the observation of an action performed by another individual evokes the neural activity that corresponds to that which, when generated internally, represents a certain action, the meaning of it should be recognized, because of the similarity of the two representations. (Gallese et al. 1996. p.606)

In several ways this description of the mirroring process goes beyond and is different from the caricature story of mirror matching according to which we by way of mirror neurons covertly imitate or simulate the actions of others and thereby understand them by, so to speak, going through the motions. Gallese and his coworkers here highlight something *beyond* simply mirrored motor action representations. They point to the

---

[116] For the issue of trust and social goals I think the research by Tomasello and his colleagues is very interesting (Tomasello 2008 & 2009). And I will say some more about the issue of non-human primate social cognition later in this chapter and, also, in Chapter 4 in relation to the question of cognitive continuity and the ignored importance of the ability of decoupling from one's own perspective in advanced other-centered social cognitive skills.

association between actions and their anticipated consequences, and imply that this sort of sensorimotor knowledge is integral to normal action performance. I could not agree more. The importance of perceptual anticipation and action association is at the core of William James' ideomotor principle and has been highlighted by many researchers since.[117] I think that James was onto something pivotal when he suggested that the perceptual anticipation of goals and sensory consequences is not simply associated but plays a key role in both organizing and initiating motor processes. Accordingly, I think that these action consequence associations are incredibly central not only to mirror neurons but also to the larger sensorimotor integrative role of these premotor and parietal areas. As I have suggested earlier visuomotor neurons do not simply link perceptions to motor schemas, rather the schemas instantiated by these integrations are inherently sensorimotor. This fits with James' proposal that actions are always imagined or represented by way of perceptual experiences. What I take Gallese and his coauthors to be saying in the quote above is that this ideomotor action outcome knowledge can be *extended* to others by way of mirror neurons, and maybe by way of a broader sensory-motor activity that the perception gives rise to.

However, one should note that the sensory consequence integration that Gallese and his colleagues theorize here seems to go beyond the activity of individual strictly congruent mirror neurons, since the visual properties of the observation/execution matching mirror neurons themselves seem restricted to the action observed and not temporally preceding or following perceptions. Thus, it seems that the sensory consequence integration has to rely on processes *beyond* a narrow observation/execution matching mechanism. I am obviously friendly to the idea of mirror neurons working in concert with the broader sensorimotor circuits of the fronto-parietal areas. But this expanded basis for action understood as reaching beyond the mirror neurons themselves poses some serious questions at least to the caricature mirror matching model, which is conceptually based on the idea of individual automatic mirroring modules. Now it seems that one needs to add a much more complex story about the sensory integrations in mirror neurons and mirror areas in that one needs to explain how the idea of sensory anticipations of goal outcomes gets integrated in the 'motor vocabulary.' As I see it, such a new inclusive story would likely need to change the proposal to one of a 'sensorimotor vocabulary' for

[117] James (1890).

goal-directed actions, which again would result in a more fundamental reinterpretation of the realm of motor cognition.

According to the quote above the mirror neuron activity is only thought to ‘extend’ the pragmatic understanding that we have from our own actions to the observation of others. But when thought through, this must involve more than simply an action representation referenced to the other. I am thinking that we do not simply perceive the other’s action in isolation, but use our own general understanding of the affordance space and its teleological relations and opportunities to judge and interpret the perceived actions of others. Such a *relational understanding* of the other’s action would within parietal-premotor circuits depend not only on strictly congruent mirror neurons but precisely on the dynamics between these and all the canonical and other responses to the scene at large. In sum, the action understanding that Gallese and his colleagues propose as the function of mirroring requires a lot more than one can get from strictly congruent mirror neurons.[118]

Unfortunately, I have never seen an explanation or model for such a more dynamic process where there is an integration of classical mirror neurons and ego-centered affordances and the perception of the other’s action in relation to these. Rather, the focus is normally on how we can use our own ego-centered knowledge of the currently perceived action to better understand the other. Thus, the typical explanation narrowly focuses on the *action* rather than *its place in a shared context*. Gallese and colleagues write that the actions of others generate ‘neural activity that corresponds’ to an action initiated endogenously. One is probably to take the idea of correspondence rather loosely here but, as we shall see in the next chapter, the idea of endogenous ‘as if’ action is at the core of the *simulation* idea of mirroring. And given the idea of extending one’s endogenous action knowledge to the other it is likely that they here have a sort of action simulation process in mind that at least goes beyond the retrieval of a schematic action representation. Thus, their idea seems to be that we understand the actions of others because mirror neurons start a cascade of action simulation, of covert processes ‘as if’ the action was actually endogenously chosen for overt execution. This broader simulation

[118] This is not meant to suggest that general action understanding only depends on the present affordance monitoring, which I hypothesize as instantiated by parietal-premotor circuits, but the other way around that mirror neurons are not sufficient even for the action understanding hypothesized by Gallese and coauthors.

process is thought to yield a first personal or 'embodied' pragmatic understanding of intentional actions of the other. One question is here to what degree the activity occurring during normal action execution is thought to be simulated or retrieved. Another question is about the extent of the supposed me-other and real action versus simulated action overlap. I think that a spatio-temporally defined simulation process of 'as if' action is empirically implausible as part of normal perception and that one should rather think of multiple parallel schematic sensorimotor activations as the basis for the action understanding. I will say much more on this later, but here just note that I have already in the previous section pointed to various empirical data suggesting that even *within* strictly congruent mirror neurons the activity during respectively action observation and execution is not identical. In the broader parietal-premotor populations one can guess that there could be other significant differences between the activity during observation and execution. I shall return to these questions of the level of detail that observed actions are thought to be mirrored for the purpose of various kinds of action understanding, the extent of kinetic detail of the action and the proposed agent-neutrality of the mirror mechanism. For now, let me just reiterate that Gallese and his coauthors on the one hand seem to suggest that the simulation process goes beyond the strict mirroring of the other's action and that it is exactly due to this broader propagation of the simulation to neurons and processes beyond strict mirroring, that we gain important social insight about the others action. On the other hand this simulation process is still seen as a *simulation of the other* and as resulting in a broader *shared* or *agent-neutral* representation. In short, the individual mirror neurons seem to be *entry points* into such observation based action simulation but not the sole neurological foundation. One should notice that given this explanation it seems conceivable that one could simulate or covertly imitate the actions of others *without* mirror neurons, but what the discovery of mirror neurons suggests is that the action retrieval process is *automatic* and maybe even *ubiquitous* for action perception. In other words, under this explanation, mirror neurons do not seem to instantiate the broader action simulation or action knowledge but might simply serve as an entry mechanism by which the perceived action of the other can automatically ignite our action knowledge. If this interpretation is correct it is *via* mirror neurons that our action knowledge and experience is *extended* to others without the need for extra mediating processes. One must note that this claim is significantly different from the general caricature story of agent neutral action mirroring/simulation, where it seems that the action information is assumed to be *integral* to the mirroring process. I.e. the idea that individual mirror neuron 'contains' the relevant feature of the action representation and then this feature is directly triggered by the perception with which it is integrated in this neuron.

If the self-containedness of the mirror process is challenged it also raises a question in regard to the *automaticity* and *ubiquity* of the *broader* mirroring response and the broader action simulation. The element of automaticity or ubiquity of the mirroring response is implied by Gallese and colleagues, and it generally seems to be a central element of the mirroring theory that the motor response is unmediated, ubiquitous and relatively context independent. But these claims are rarely explicitly discussed or tested. I discussed earlier the claim that these neurons do not habituate, which would indeed support the ubiquity of the response of individual neurons. Now a new problematic question is how the broader action simulation is ubiquitously guaranteed even if the individual neurons consistently fire. I shall argue that it is only plausible that actions are simulated in a spatiotemporally specified way if we are not engaged in a competing activity. In other words, I do not think that the full simulation element of the mirroring story can be ubiquitous, and further that it seems to be the exception rather than the rule that we engage in such full covert simulations of others' actions.

3.4. From F5 mirror neurons to a mirror neuron system

As we have seen, the hypothesis of a mirror neuron based observation-execution matching mechanism was already proposed after the early discovery of premotor mirror neurons in F5. However, it is rather atypical to construe a theory of neurological function based on individual neuron properties. The connectivity between areas is naturally seen as highly significant to functions, and theorized cognitive 'units,' 'modules,' 'mechanisms' and 'systems' are normally thought to span vast regions or multiple areas of the brain. Similarly, the hypothesis of a 'mirror mechanism' or 'mirror neuron system' was proposed to be instantiated not just by neurons in area F5 of the premotor cortex but instead by multiple fronto-parietal circuits, and possibly also supported by key 'extra-mirror neuron' regions like the STS, which does not contain mirror neurons as such. In monkeys, single cell recordings have shown the existence of mirror neurons in various inferior parietal areas, and begun to reveal also the differences between the properties of the neurons in these areas. However, due at least in part to the definition and classification of mirror neurons based on ‘action kinds’ and the broader implications of the mirror metaphor these differences are rarely discussed in any systematic way. Before turning to the fronto-parietal mirror neuron circuit, I will first take a quick look at the ‘mirror-like’ neurons found in the STS. The generalization of the similarity between the visual properties of STS neurons and F5 mirror neurons is informative as to the problematic theoretical framework construed around mirror neurons.

3.4.1. ‘Mirror–like’ neurons and action perception in the STS

Prior to the discovery of mirror neurons it had been shown that individual neurons in the anterior part of the superior temporal sulcus (STS) were modulated by observation of different complex biological motion stimuli.[119] These STS action perception neurons do not even fall under the broad definition of ‘mirror neurons,’ as none of them were found to have any direct motor properties. However, it has been hypothesized that they might contribute visual information to premotor mirror neurons via connections to areas in the parietal lope that again are highly connected with premotor areas. I do not have any objections to that proposal – the problem is the implicit suggestion that STS neurons prepares the very visual representation of the action that is then ‘translated’ into a ‘motor format’ by mirror neurons. As I shall discuss in regard to Jeannerod later, I generally question the empirical support for the existence of ‘similar’ or matching representations in respectively visual or motor format, and here I just point out that such representational dualism is assumed and localized in respectively the STS and F5. We saw earlier how the F5 is theorized as containing a ‘motor vocabulary’ of goal directed actions, and that the organization of this vocabulary is implicitly assumed to be independent of sensorimotor integrations as such. Differently put, it is a vocabulary of <u>motor</u> representations, which just happen to have perceptual connections and thus lend themselves to being used for perceptual guiding of actions and even for specific perceptual purposes. The STS is hypothesized as containing the visual perceptual representations of these very goal-

[119] Perrett et al.(1989) and Perrett et al. (1990). For evidence regarding faces and gaze directions see also Perrett et al. (1985).

directed actions. Thus action sensitive STS neurons are seen as the perceptual counterpart providing the pictorial analysis of the actions 'matched' in F5.

The basis for this story is the finding of motion sensitive neurons in the STS that are modulated by perception of goal directed actions. Most studies report these neurons to have visual properties that are very similar to those of F5 mirror neurons:

> The similarities between the properties of F5 and superior temporal sulcus neurons are striking. Both populations of neurons appear to code the same actions, generalize their responses to different instances of the same action, do not respond to hand movements miming the preferred action in the absence of the object, and do not respond to object-object interaction, even when the moved object is similar to an arm and hand…The major difference, however, between F5 and superior temporal sulcus neurons, is that only in F5 are there neurons that both respond to complex visual stimuli and have movement-related activity. (Gallese et al. 1996, p. 606)

This is a simplification to say the least and it might have been worthwhile to point to a few differences between these neurons *beyond* the lack of motor properties in the STS. First of all, there are neurons in STS that react to a much wider range of complex actions than hand-object interactions, such as for example walking. In and of itself this broader range of action sensitivity does not contradict the idea that some 'goal-directed hand-object interaction neurons' in the STS have 'mirror-like visual properties.' However, there seem to be a host of subtle differences that are ignored. For example, STS neurons appears to respond well to very simple indicators of motion, such as video of point light displays of people moving, whereas F5 mirror neurons seem to prefer the observation of live action. Another notable difference is that many STS neurons seem to integrate attention clues such as eye gaze with action information.[120] Lastly, I want to mention that it is *not* that STS neurons simply lack positive motor response, which is also the case for some *visually dominant* F5 neurons and a significant number of inferior parietal lobe (IPL) neurons. Rather, it has been found that most action sensitive STS neurons do not even respond to the *perception* of action, if this action is judged to be *one's own.*[121]. This is to say that it seems that in regards to action execution there is something more complex going on than simply 'lack of motor-related activity.' The 'mirror-like' neurons of the STS might very well play an important role in regard to both the understanding of actions and to the fronto-parietal mirror neuron circuit. But given these differences the question

[120] See Allison, Puce & McCarthy(2000), and also Keysers & Perrett (2004).
[121] See Keysers & Perrett (2004).

is whether it is fruitful to simply characterize them as 'mirror-like' and functionally as providing the visual or 'pictorial' analysis of the action. Gallese writes as if it is a fact that the function of the STS is to provide a pictorial analysis of the action. He seems to think that the question merely is whether this 'ventral stream' analysis is the end point of our semantic understanding or whether our understanding of actions depends on the pragmatic enrichment that is contributed when the pictorial analysis of the STS is integrated with our motor vocabulary in F5. He writes:

> The presence of two brain regions with neurons endowed with similar complex visual properties raises the question of their possible relationship. Two possibilities might be suggested. One is that F5 mirror neurons and STS neurons have different functional roles. STS neurons would code the semantic, the meaning of hand-object interactions, while F5 mirror neurons would be engaged in the pragmatic coding of the same actions. Being that area F5 is the recipient of visual information fed mainly by the parietal lobe (see Matelli et al. 1986, 1994), this hypothesis is in line with theories positing a sharp distinction between pragmatic and semantic coding within the two main streams of visual processing (see Milner and Goodale 1995). A second possibility, that I personally favor, is that these two "action detector" systems would represent distinct stages of the same analysis. *The STS neurons would provide an initial "pictorial" description of actions* that would be then *fed*, (likely through an intermediate step in the posterior parietal cortex), to the F5 motor vocabulary where it would acquire a meaning for the individual. This latter hypothesis stresses the importance of the motor system in providing meaning to what is "described" by the visual system, by positing a pragmatic "validation" of what is perceived. (Gallese, 1999 p.173. My italics)

I agree with Gallese's resistance to a too strict 'Milner-Goodale-style' division between pragmatic and semantic processes. But the point I want to make here is that it might be difficult to understand the nuances of the interrelations between semantic and pragmatic processes if one simply, as Gallese does here, describes F5 and STS as 'two brain regions with neurons endowed with similar complex visual properties.' There seems to be something more radical going on where the STS neurons are particularly devoted to the understanding of others in their specific *individuality* as exactly *different* from ourselves.[122] In contrast, the F5 neurons often seem to 'see' the actions of others through a prism of our own motor repertoire and thus to some extent eliminate or ignore the

[122] The STS is a very complex and heterogeneous area – with many functional sub-areas as well. I am not pretending to have a theory of STS function as such or even to be able to define its role in social perception. My point is here simply to show that the contrast between F5 and the STS neurons responding to goal directed actions goes beyond the simple issue of the existence of motor properties. A more complete story would take on the fascinating role of STS in face and gaze perception and in understanding the attentional direction of others in regards to the context.

individual differences between self and others. I shall later discuss the possible roles of the STS in social cognition further, but for now I just want to emphasize that the very idea of STS neurons as 'mirror-like' might be too simplistic.

3.4.2. The discovery of mirror neurons in parietal areas

In regard to 'real' mirror neurons, I have so far focused on the early discovery of these special visuomotor neurons in area F5 of the premotor cortex of the macaque monkey. However, mirror neurons were soon found in the inferior parietal lobe as well.

The posterior part of the parietal lobe was traditionally thought to be a purely perceptual associative area supporting in particular perceptually based spatial and body representations. But this view was challenged by various findings in the late 70's and 80's of motor properties in these areas[123] and, as mentioned earlier, the Parma lab itself played an important role in the redefinition of the posterior and in particular the inferior parietal areas as being part of the cortical motor system.[124] Their research also pointed to the functional importance of the heavy connectivity between various premotor and parietal areas. In a 1997 article called "Parietal cortex: From sight to action," Rizzolatti, Foggasi and Gallese challenge the traditional idea of there being an area for space perception independently of pragmatic considerations. They write:

> The posterior parietal lobe is now thought to consist of a mosaic of areas, each receiving specific sensory information and transforming it into information appropriate for action, with no identifiable 'space area'. Rather, space perception appears to be a secondary result of the activity of a series of sensorimotor circuits, each of which encodes the spatial location of an object according to its own motor purposes and transforms it into a potential action. (Rizzolatti et al. 1997, p.562)

This article focuses on the sensorimotor circuits that have to do with organization of hand arm and head movements and particularly in the parietal area AIP (anterior intraparietal) which is highly connected to area F5 of the premotor cortex where canonical and mirror neurons were first discovered. There is no mention of mirror neurons in this article, but it focuses on the role of these sensorimotor circuits in visually guided action execution. They write:

---

[123] See Mountcastle et al (1975) and J. Hyvärinen (1982).

[124] The parietal lobe of primates is normally seen as composed of the postcentral gurus (aka the somatosensory cortex), the superior and the inferior parietal lobe, and it is the latter two that are also together referred to as the posterior parietal lobe, and which were earlier simply seen as a large multi modal sensory association area.

> …We have developed a model whereby area AIP provides multiple descriptions of a 3D object, thus 'proposing' several grasping possibilities to area F5 – for example a cup can be grasped on the handle or on its upper border. Area F5 then selects the most appropriate type of grip on the basis of contextual information (e.g. purpose of the action, spatial relation to other objects etc.). Furthermore F5 fragments the selected grip into phases, such that aperture and closure and simultaneously keeps active the set of AIP neurons that also encode the selected grip. (Rizzolatti et al 1997, p.563)

Thus, in regard to object-directed actions, they hypothesize that parietal areas contain multiple pragmatically oriented object-centered descriptions affording or 'proposing' various grasps/actions then specified in area F5. Area F5 is then thought to represent various perceptually afforded actions simultaneously and use contextual information to help select among these.[125] Lastly, they add that F5 not only represents multiple simultaneous action possibilities, but also 'fragments the selected grip into phases' and 'simultaneously keeps active the set of AIP neurons that also encode the selected grip.' This explanation is interesting – and very much consistent with my affordance space hypothesis. There is obviously a very tight and dynamic relationship going on between these areas in regard to both action specification and choice, and perceptual attention and interpretation. The parietal AIP area is sometimes described as holding the pragmatically relevant features of the perceptual environment attentionally captured for the F5 premotor neurons to choose both *which* actions and sub-acts to perform and *when*. Of course action choices again determine which perceptual features will be relevant and thus the sensorimotor processes of F5-AIP circuits seem necessarily dynamic and mutually influential rather than serial with well-defined inputs and outputs.

[125] The division between parietal areas as specifying the objects for interactions and premotor areas as specifying the potential actions themselves is probably best thought of as a rough or relative distinction since both F5 and AIP contain individual neurons that are both sensorimotor and motor dominant and visually dominant neurons. In other words, there is likely a division of labor but it is implausible that it squarely follows a perceptual/motor divide, though the language of the Parma group at times suggests this. One might propose, to avoid the metaphor of perceptual/motor division, that it would be useful also to stress the seeming temporal and action specification divisions of labor. In other words, it seems that parietal area AIP focuses on the potential actions as they are invited by ongoing monitored environment and of further providing attentional specifications of such action affordance features given a 'top-down' chosen course of action. Premotor area F5 on the other hand seems to sequentially orchestrate and evaluate actions invited to further determine focus back in AIP and eventually 'choose' the actions to be initiated. Rizzolatti and Luppino suggest how the F5-AIP circuit depends on other areas: "The selection of one of these possible ways of grasping depends on preliminary object recognition and on agent intention, and not exclusively on the visual intrinsic properties of the object. Thus, a more complete model of how the F5-AIP circuit physiologically also requires information from circuits that code object meaning (inferotemporal lobe, IT) and circuits where decisions on what to do are taken (prefrontal lobe, cingulate areas)." Rizzolatti and Luppino (2001)

Given this model one might expect that mirror neurons like other visuomotor neurons in respectively parietal and premotor areas of the circuits have slightly different properties, i.e. respond to different elements of the action, maybe different levels of description. It certainly would be very interesting to take a deeper look at the exact coordination between the parietal and premotor responses. However, as mentioned earlier, most mirror neuron studies focus on a categorization of neurons according to which 'kinds of action' neurons are modulated by, and this tendency seems to have been continued in the study of parietal mirror neurons as well. The result of this categorization is that most published studies conclude that the properties of parietal mirror neurons are 'similar' to those found in premotor areas.[126] In 1998 Fogassi and his Parma colleagues published a report of their mirror neuron findings in the rostral part of the inferior parietal lobe.[127] Given exactly the above mentioned 'kind of effective action & effector' categorization they summarize their findings quantitatively in the following way: They recorded from 195 PF neurons of which 61 were responsive to hand/arm action observation while 43 also had motor properties. Thus they found 43 mirror neurons given the broad definition including any action perception sensitive visuomotor neuron. They then specify:

> A strict congruence between observed and executed actions was found in 8 neurons. Other 9 showed a similarity, but not identity, between observed and executed actions. In the remaining 26 neurons the effective observed and executed actions differed either in terms of effectors or action goal. The data indicate that a visual representation of biological actions is present in PF, in addition to those found in the region of the superior temporal sulcus and in the premotor area F5 and that an action observation/execution matching system similar to that found in the premotor cortex does exist also in the parietal lope. (Fogassi et al 1998, 257.5)

I quote this report of the results mostly to show how enormously *uninformative* their categorization is in regards to function. Given the low number of strictly congruent mirror neurons it seems particularly quick to say that these results indicate that the parietal areas have an observation-execution *matching* system, and further that it is *similar* to that found in F5. Additionally, this quick conclusion is especially surprising given the theory of a division of labor between these areas in visually directed action, which these very same researchers proposed a year earlier in the 1997 article quoted

[126] Fogassi & Luppino (2005) write for example: "These [IPL mirror] neurons have properties similar to those of the ventral premotor area F5, which is tightly connected with this IPL sector." (p.628) See also Rozzi et al. (2008) for similar observations.
[127] See Fogassi, Gallese, Fadiga and Rizzolatti (1998).

above. I naturally suspect that the reason for the changed focus again is to be found on the power of the mirror metaphor and the caricature model.

In a very recent article Fujii and his colleagues at the RIKEN Brain Science Institute in Japan provide evidence not only for a division of labor between premotor and parietal areas but also more specifically a differentiation between the representations of self and other within these areas.[128] These very fascinating findings paint a much more complicated picture than the simple idea of a mirror mechanism of agent-neutral shared action representations in the premotor-parietal circuits.[129] In regards to the possible difference and division of labor between premotor and parietal areas, the point is that the Parma group often simply focus on the fact that mirror neurons *exist* in parietal areas and that these areas are thus *part of* the 'mirror neuron system' in the macaque monkey. The discovery of mirror neurons in parietal areas thereby represents an expansion of the *already existing* mirror matching theory rather than a specification or correction of it – which is what for example Fujii and his coauthors propose. Of course the problem is that the specification undermines the narrow notion of a ubiquitous context independent agent neutral mirror mechanism.[130]

3.5. From monkeys to humans

The importance and earthshaking effect of the finding of mirror neurons in monkeys was naturally not only about the astonishing socially symmetric properties, but was also predicated on the prediction that similar neurological processes can be found in humans. The obvious problem is that ethical concerns make single cell recording from humans an extremely tricky and rare possibility.[131] Thus, in general the human evidence for an

[128] Fujii, Hihara & Iriki (2009).
[129] Iriki and the RIKEN group are in general doing fascinating groundbreaking research pointing to complex dynamics within the fronto-parietal circuits. They have for example found that holding tools will change the receptive field size of certain parietal sensorimotor neurons – only if the money *intends* to use the tool. I.e. one might say that here the body and environment is perceived through action planning. Could be very interesting to see how such dynamics would look in social scenarios.
[130] There is also some accumulating evidence of a lateralized division of labor, suggesting that the right hemisphere might be more important than the left for the understanding of intentions of others within the frontoparietal circuits (See for example Ortigue et al (2009)). This would fit with the right lateralization of other important social cognitive areas like the STS and the temporal parietal junction. Hopefully, a more nuanced story of all these findings of labor divisions within fronto-parietal circuits will be written in the future. The modest point is here to emphasize that all these differences get erased by the mirror focus.
[131] As mentioned single cell recording is torturous enough for the macaque monkeys and therefore in general not a possibility in humans. In regards, to humans Iacoboni has reported a few rare opportunities to

observation-execution 'mirror' matching mechanism has been obtained via various indirect measures such as fMRI imaging, action facilitation studies and clinical cases. As I will try to show, each of these indirect empirical techniques has its own limitations and theoretical baggage.

The list of unknowns in regards to a human mirror system and the possible social functions of such a 'mirror mechanism' is thus even longer than it is in regards to macaques. Beyond questions of cognitive continuity between monkey and human social cognitive skills, I want to raise some questions regarding the move from single cell studies to behavioral motor resonance and imaging studies. There is a question as to whether single cell mirror neurons actually *exist* in humans, i.e. whether there are single cells in humans that integrate visual and motor action information or whether the fMRI modulations are due to various kinds of neighboring cells. I will briefly discuss this debate a little later, but one should note that it would be rather shocking if the massive single cell sensorimotor integration, which is found in monkeys and which supports visually guided actions would not be paralleled by single cell sensorimotor integration in the homologues areas of humans.[132] TMS lesions for example have been used to provide evidence that the fronto-parietal areas in humans support similar functions of perceptually guided actions.[133] In other words, there is a high burden of proof for the argument of discontinuity between monkeys and humans at the *fundamental level* of sensorimotor functional organization, and evidence well beyond simply mismatches in fMRI activation between action and perception tasks would be needed to support the idea of sensorimotor *segregation* in humans.[134]

This is however not to say that 'mirror neurons' might not be an even more muddy affair in humans than in monkeys. But muddiness and lack of symmetry is in and of itself is

---

record from single neurons in medial prefrontal cortex of patients undergoing surgery for epileptic seizures. These recordings reportedly resulted in findings of action sensitive neurons or 'super mirror neurons' that were modulated by both action production and perception- but these two modes would differentially respectively decrease and increase the activity of these neurons. Thus they were proposed to be 'super mirror neurons' that might be involved in the automatic inhibition of the overt imitation of the actions of others. (Mukamel et al. (2010) have already published actual mirror neuron findings from such patients.)

[132] As mentioned, findings of single cell mirror neurons in patients undergoing surgery for seizures have very recently been published by Mukamel et al (2010). However, due to the medial location of the recordings there is still no data from the fronto-parietal circuits.

[133] For such a TMS study on the parietal areas of the fronto-parietal circuit see: Tunik et al. (2005).

[134] Rizzolatti argues that the so-called 'two population' interpretation of the fMRI data of fronto-parietal areas in humans is highly implausible as neighboring sensory and motor information must be integrated in some cells and that these cells per definition would be sensorimotor. (Personal communication)

hardly a problem for the argument that I am advancing – as I am exactly attempting to question the idea of a strictly congruent agent neutral action observation-execution mechanism. Accordingly, the question that I am interested in is what motor resonance and fMRI observation-execution overlaps and differences tell us not simply some sort of relatively encapsulated, hard-wired, ubiquitous and agent neutral mirroring mechanism, but also about the of the more complex sensorimotor mappings. It is clear that rough overlaps of fMRI activation do not provide evidence of individual strictly congruent mirror neurons, but a further question is to what extent these provide evidence for the hypothesis of a 'mirror mechanism;' i.e. a neurological instantiation of a ubiquitous cognitive mechanism yielding automatic agent neutral shared action representations. Via a discussion of James' ideomotor principle and the role of experience based sensorimotor integrations in actions production, I hope to raise a few flags, which are informative for my general project. It is important to understand that neither motor resonance nor fMRI analyses of blood oxygen levels are simply products of mirror neuron activity, but of much broader neurological dynamics. In other words, I emphasize that even if there were to be ubiquitous mirror neuron responses that would not in and of itself guarantee ubiquitous overall motor resonance. I argue that it is problematic that much of the human research is narrowly focusing on proving or disproving the action-observation symmetry in the fronto-parietal circuits, as such projects ignore the fascinating details of the found asymmetries and contextual dependencies of sensorimotor activations. Below I will discuss a series of recent studies that actually do point to a much more complex dynamics of motor resonance during action perception than what is emphasized by the traditional mirroring perspective. Beyond introducing human evidence, this section will therefore also inform my further arguments for understanding 'mirror' phenomena and their function within the broader context of fronto-parietal processes.[135]

[135] I aim to show that mirroring processes must be understood within the larger functional, social and neurological context, rather than as a neurologically modular and ubiquitous mechanism with an encapsulated cognitive function of agent neutral action representation, but the discussion is also meant to support more specific structural claims. One such claim is that mirror neuron activity should not be seen as instantiating a full action simulation but rather a more abstract rough sketch. Another such claim is that the social function of fronto-parietal circuits should not be seen as exclusively based on mirror neuron activity.

3.5.1. Social abilities in macaques and humans

Beyond the indirectness of the human evidence, a big issue in moving from the mirror neuron evidence in monkeys to that of motor resonance in humans concerns the obvious differences in social abilities between monkeys and humans. Ironically it is exactly the difference in the social cognitive status of humans and macaques, which underlies the ethical judgment that one can do these sorts of single cell recordings on the latter and not the former. Notice that this methodology is not even allowed on chimps and other great apes – thus that it actually *is* allowed on macaques could actually be due to discontinuities between macaque and human social cognition, and by implication possibly between macaque mirror neurons and the functions of human homologue areas. If similar premotor-parietal circuits play an essential role in human social cognition there is an important story to tell as to why monkeys do not seem to posses these very abilities. Many of the hypothesized social cognitive functions of mirror neurons such as for example language, imitation and empathy simply do not seem to be present in the macaque monkeys, and other core abilities thought to be supported by mirror neurons functions such as action and intention understanding certainly show great divergences as well. These differences raise basic questions about the function of mirror neurons on the one hand and about the degree of similarity between the mirror systems in monkeys and humans on the other. The poverty of some social abilities in monkeys might suggest a highly deflationary account of mirror neurons where their main function does not pertain to social understanding at all but rather simply supports for example hand-eye coordination.[136] Conversely, proponents of the idea that mirror neurons are the foundation of our advanced social cognitive skills might suggest that the difference in social competence is due to a broader difference in cognitive sophistication and self-knowledge. Thus their question is not so much *whether* there is a shared representation but *what* sort of shared representations mirror neurons yield and what sort of broader *use* the observer can make of them. I shall discuss the issues of cognitive continuity and my own proposal in regard to mirror neuron functions in monkeys and small children more in the next chapter.

[136] Cecila Heyes has provocatively suggested this to be a main function in a conference presentation, but I do not know if she has done so in print.

We should notice that the differences in social abilities in many articles and books about mirror neuron findings are neatly swept under the rug and the authors swiftly move from the single cell findings in macaques to the findings of a 'homologue' mirror neuron system in humans. Similarly to the argument I made in the previous section in regards to the issue of a division of labor between various parts of the 'mirror neuron system,' I think that the details of the social cognitive skills of macaques as opposed to higher primates are very informative for our understanding of the premotor-parietal circuits and their role in social cognition. Still, the neat idea of shared representation via execution-observation matching does not provide a very useful framework for probing these differences, as the mirror metaphor appears to suggest an all-or-nothing phenomenon.[137] With all these caveats in mind let me turn to the human mirror neuron findings, starting with a look at the behavioral findings of motor resonance during action perception.

3.5.2. Motor resonance and ideomotor action

It has long been known that the observation of others' actions can elicit various forms of 'motor resonance' and action facilitation in the observer, i.e. that in various instances movements similar or compatible to the ones observed are primed or initiated in the observer's motor system. Many mirror neuron theorists point back to William James' observations and his interpretation of these phenomena in terms of ideomotor action. Buccino and his coauthors write in a 2004 article:

> According to this [ideomotor] theory, building up an image of the sensory feedback related to a certain action is a fundamental step for the proper execution of that action. The prediction of this theory is that, when that execution of an action is guided by a stimulus, the more the stimulus is similar to the action, the more the execution of that action is facilitated…The theory of ideomotor action finds indeed a strong neurological support in the existence of a mirror neuron system, where, by definition the visual representation of an action and its motor counterpart are anatomically and functionally embedded. (Buccino et al 2004, p. 372)

They refer to the ideomotor theory and the hypothesis that actions are initiated by way of anticipatory representations of the sensory feedback. As their focus is on mirror neurons,

[137] I have already earlier referred to the articles by Lyons et al. (2006) and Iriki (2006) for interesting ideas regarding social cognition and mirror neurons in monkeys. Notably, the idea of 'intentional compression' is introduced to begin to understand the difference between human and monkey mirror neuron systems and abilities. The idea is that mirror neurons are used to represent and understand goal directed actions at various levels of abstraction or 'compression.' Thus, a compressed intentional understanding does not guarantee an ability to actually take others' perspective or imitate their actions in any detail.

they underline the elements of *similarity* between the action and the stimulus that is the key to the sensorimotor integration. However, it is important to note that the ideomotor theory presents *a much broader hypothesis* of action initiation in general as based on goal/outcome anticipation - and that a perception always, in so far that it evokes the 'idea' or 'image' of an action goal, yields some level of motor resonance. In other words, ideomotor action does not depend on a literal similarity between the stimulus and the action. That being said, clearly James' theory is very interesting to look at in relation to the contemporary mirror neuron and motor resonance findings. In my eyes, it is particularly interesting to look closer at the ideomotor theory given that the broader ideas of sensorimotor integration and action initiation might not fit so well with a narrow hypothesis of an observation-execution mirror matching mechanism. Consequently, I think that the details of James' thought might *inspire* a more nuanced understanding of mirror neurons and their functional role in regard to more complex cognitive processes, rather than simply being pointed to as historical precursor.

From the perspective of mirror theorists the important feature of the ideomotor theory is that it can explain the many reports of perceptually induced action facilitation, and, further, that the hypothesis of processes of sensorimotor integration can explain how the observation of actions directly can elicit comparable actions in the observer, i.e. without the need of an intermediary process of reasoning and action decisions. James defines ideomotor action precisely by way of the directness of the response activity. He writes: "Wherever a movement follows unhesitatingly and immediately the notion of it in the mind, we have ideomotor action." (James, 1890, p.1130) But we also see how tremendously broad his definition is. As he himself notes, ideomotor action is not some kind of curiosity of the mind – as it often was interpreted as due to para-psychological associations - but rather a *core principle at the base of voluntary action*. James thinks of voluntary action as made possible by sensorimotor associations formed through experience.

> A supply of ideas of the various movements that are possible, left in the memory by experiences of their involuntary performance, is thus the first prerequisite of the voluntary life. (James, 1890, p.1099-1100)

Thus, actions are represented and organized via their experienced consequences – or the goals that they have been known to produce. For such action 'ideas' to become action, according to James, one simply needs to think of them or be perceptually reminded of them. He writes:

> Whilst talking I become conscious of a pin on the floor, or of some dust on my sleeve. Without interrupting the conversation I brush away the dust or pick up the pin. I make no express resolve, but the mere perception of the object and the fleeting notion of the act seem themselves to bring the latter about. (p.1131)

James here claims that we do not need to form a separate intention or an 'express resolve' beyond what I would call the perceived action affordance. The simple perception for example of some dust on my jacket leads directly to the action of removing it due to the sensorimotor integration underlying the pragmatic interpretation of this very perception.[138] He continues a little further down:

> The reason why that doctrine is not a self-evident truth is that we have so many ideas that do not result in action. But it will be seen that in every such case, without exception, that is because other ideas simultaneously present rob them of their impulsive power. But even here, and when a movement is inhibited from completely taking place by contrary ideas, it will incipiently take place. (p. 1133)

This is an extremely important passage. James here stresses two elements of his theory, which are incredibly insightful given contemporary research. Namely, that A) all action representations lead to some incipient action, and that B) the degree to which it will be overtly carried out depends on the competing action plans. Or as he himself famously writes:

> …Every representation of a movement awakens in some degree the actual movement that is its object; and awakens it in a maximum degree whenever it is not kept from so doing by an antagonistic representation present simultaneously to the mind. (p. 1134)

Thus, we can sum up the key elements of James' ideomotor theory. We saw above how the theory of ideomotor action is often understood by mirror neuron theorists in terms of 1) the *directness* of the response in relation to the idea and 2) the organization of action around *anticipations* of goals/perceptual outcomes. To this we can add the further elements that 3) we can *simultaneously* entertain and anticipate multiple action outcomes and that 4) they all lead to some *degree of* motor activity by way of the representation simply having been evoked. Lastly, 5) the reason why only some action representations or 'ideas' lead to *overt movement* is that *parallel* action representations *compete* for the spatio-temporally limited action resources of the body in the present environments.

[138] It is very interesting in James' example here that more actions are going on simultaneously; it is because my hands are rather idle in the possibly boring conversation that the afforded ideomotor actions are carried out and not outcompeted by the hand gestures of emotionally involved and intense conversation.

I shall argue that each of these elements could productively inform a broader interpretation of fronto-parietal integration during action perception than the one suggested by a 'mirror mechanism' interpretation.[139] However, one major difference between James' discussion and the way I will understand the ideomotor ideas pertains to the issue of consciousness and perception. My focus in this project is on sensorimotor integration in fronto-parietal circuits at a sub-personal level, whereas James' discussion is a mix between suggestions as to the role of consciousness in action initiation and claims about sensorimotor learning probably also at a more personal level.[140] I will now turn to more recent findings and speculations.

A. <u>Behavioral findings – Motor resonance & facilitation</u>

Numerous contemporary studies also explore motor resonance phenomena and their possible relations to a mirror neuron system. Fadiga and coworkers from the Parma lab published a study in 1995 that is often referred to as early evidence for a human homologue to the mirror neuron system in macaque monkeys.[141] They showed with the use of Transcranial Magnetic Stimulation (TMS) of the motor cortex that motor evoked potentials (MEP) were significantly increased during action observation. They further found that these motor evoked potentials reflected the muscle activity of the execution of the observed action. Hence, this study is generally taken as evidence that the observation of action activates the motor system in a way that facilitates the execution of the same action. More narrowly the authors argue that the findings also point to an automatic observation-execution matching system like the mirror neuron system found in macaque monkeys.

[139] Contemporary researchers such as Wolfgang Prinz and his colleagues at the Max Planck Institute are explicitly building on James' theory of ideomotor action control and anticipatory sensorimotor organization of actions in theories such as the Common Coding theory. See Hommel et al. (2001) and Prinz, W. (1997).
[140] One might rightfully say that the mirror neuron research and my project here do indeed include both the personal and sub-personal levels as the shared focus is the role of social action *perception* in inducing *sensorimotor* activity in specific cortical areas. But the point is that the role of consciousness is not the focus. My view is that James – probably in part due to lack of experimental alternative and in part due to his view of free will – overemphasizes the role of consciousness in action initiation. However, I think that the Parma group and other neuroscientists might underemphasize the potential import of consciousness and awareness for the modulation of fronto-parietal circuits and it would be very interesting to do more studies on the role of perceptual attention. I have already referred to findings that the modulation of parietal mirror neurons depends on foveation whereas that of premotor mirror neurons does not.
[141] Fadiga et al. (1995).

Other behavioral studies use reaction time as a measurement of action facilitation and interference respectively. The basic concept is here that the reaction time reflects potential conflicts between the overt action task and covert motor priming by way of observation. Thus, various cases of motor priming and task are classified as either ‘congruent’ (similar action observation and execution) as opposed to ‘incongruent’ (different action observation and execution), and the shorter reaction time for congruent actions is taken as evidence that during action perception action plans similar to those observed are automatically activated in their own motor system.[142] In these sorts of paradigms, congruency and incongruence is generally defined in terms of the actions as being the *same* or *different*, and not in terms of the *pragmatic relation* between the action observed and those to be executed. Given my discussion of ideomotor action and the heterogeneity of sensorimotor integration this is rather problematic, since some ‘incongruent’ but teleologically meaningful and previously learned actions might under some circumstances be facilitated. Further, there might be multiple simultaneous action facilitations, and interference effects might therefore depend on the extent to which one is trying to specify more of these ‘incipient’ representations to a degree that is under more narrow load constraints. If this is the case then the motor resonance should be interpreted in light of the task at hand and the broader pragmatic context it occurs in. But, it seems that the idea of a ubiquitous modular mirror mechanism often is more assumed than it is explored by some of these behavioral studies.[143] More specifically, I think that many of the experimental protocols have a clear *imitation bias*, i.e. their pragmatic setup is such that the subject is prone to imitate (produce a congruent action) rather than engage in a complementary or incongruent action. Thus, the typical protocol of action facilitation studies is that the observer watches someone perform a simple transitive or intransitive action like a finger movement with or without an external target - or often simply a video clip of the moving body part. The subject is then tested during simple observation and while performing imitative or incongruent actions. The key is that the incongruent action is normally not simply different but arguably incoherent with the observed action and overall context. The subject is supposed to carry out their own action plan in parallel with

[142] See for example Brass, Bekkering & Prinz (2001), and also Craighero et al. (2002).
[143] For a rare study incorporating the relevance of different cues and interpretations of actions in their pragmatic context, see Brass et al. (2000).

watching the other's action, which is irrelevant to their task. In the imitation task on the other hand the other's action is a useful part of the 'affordance structure' and can be used to perceptually guide one's own action. On my affordance structure monitoring story of parietal-premotor sensorimotor integration, one could expect that if one could set up a scenario where the other's action *affords a complementary action* on the part of the observer, then the pattern of facilitation would change. This prediction is exactly what recent empirical studies seem to support. I will now look closer at some of this recent research.

Cecilia Heyes and her colleagues at Oxford University have shown that experience modulates the typical automatic imitation response.[144] They trained subjects in responding to an observed action of hand opening with the opposite action of hand closing. What they found is that the observed actions no longer facilitated the congruent or 'mirror' response. In this study the training did not lead to direct facilitation of the incongruent action, but later studies by Catmur, Walsh and Heyes showed that such a 'counter-mirror effect' could be produced.[145] In this latter study they used TMS (Transcranial Magnetic Stimulation) to produce muscle-specific responses in the primary motor cortex. In these latter studies, with the use of a somatotopical map, various effectors such as fingers were represented in anatomically different locations. By way of this paradigm they could show that the facilitation induced by observed finger movements could be moved from the representation of the same finger as the one observed to a trained 'counter-mirror' response in a different finger. These studies are extremely interesting in regards to the mirror neuron debate on several fronts. Catmur, Heyes and their coworkers mostly focus on how their findings constitute evidence that mirror responses are not wholly innate, and that they can be modulated by sensorimotor experience and learning. I think that learning and sensorimotor reconfigurations are very real and very important, but I think that Heyes, Catmur and colleagues might be a little quick in conclusion that these findings of finger response facilitations show that *individual* neurons or connections have been reconfigured. One should underline that these studies do not in fact show that single *mirror neurons* have had their response properties inverted by training, but rather simply that learning happened at the level of

[144] Heyes et al. (2005).
[145] Catmur, Walsh & Heyes, (2007). See also the more recent Catmur et al. (2008).

the overall primary motor response. In other words, just like motor resonance findings do not produce single cell evidence that mirror neurons do 'mirror,' this sort of evidence can also not show that they 'counter mirror.' What it does show is that the perceptually induced global motor response depends on the sensorimotor training of the subject and the task performed. And of course if the sensorimotor training reconfigures the perceptually induced response one might assume that this is a reflection of new or changed sensorimotor integration also at the single neuron level. But my point is that this sensorimotor learning could be instantiated by various sensorimotor neurons – not necessarily by strictly congruent mirror neurons turned born again 'counter mirror neurons.' To sum this up, there might be some opening in regards to whether these studies provide evidence against the 'hard-wiring' of individual mirror neurons, as I shall briefly discuss again below in regards to the neonate imitation findings of Meltzoff and Moore.[146] The ignored alternative interpretation of these studies is that the training leads to new action affordance integrations, and that the 'counter mirror response' relies on sensorimotor reconfigurations that go beyond traditional strictly congruent mirror neurons. Be this as it may, the studies are still of great importance. I have tried to stress how the sensorimotor training studies indicate that the global motor resonance is highly experience based and task-dependent. Heyes and Catmur and colleagues do not discuss the importance of the task and context much as their main interest is in associative learning and thus in the experience based learning rather than the dynamics of 'top-down' control or the possible more complex hierarchical flexibility of sensorimotor connections. Here, it is interesting to look at studies by Hein van Schie, Harold Bekkering and colleagues, which further investigated the importance of not only training but also the pragmatic and social context of the perceived and executed action. They did a study, which shows that reaction times can be reversed from a faster imitative response effect to a faster complementary action if the context and object invites – or affords – the complementary action.[147] In this study they found the traditional imitation effect if they prepared same or different action were cued by a color on the observed hand doing respectively a whole hand grasp or a precision grip of an object. But if the color cue was moved from acting hand to the object to be apprehended – and thereby proving a context

[146] Meltzoff and Moore (1977).
[147] Van Schie, van Waterschoot & Bekkering(2008).

affording attention to the object and complementary action – the imitation effect was reversed. Poljac, van Schie and Bekkering have since done an interesting follow up study to test whether the discovered reversed imitation effect was due simply to associative learning reconfiguration of sensorimotor connections reliant on 'counter mirror' stimulus-response practice like Heyes and colleagues suggest, or whether the effect is influenced by something beyond simple training - like the context and goal of the task.[148] They conclude:

> Thus, the present results provide further support for the theoretical notion that task representations may temporarily overrule existing long-term S–R associations in accordance with task requirements and that this flexibility in action–perception coupling cannot adequately be explained in terms of associative learning effects. (Poljac et al. 2009, p. 583)

These findings fit my broader hypothesis, in that they suggests that the overall action facilitation is a product not simply of automatic mirroring, or even other singular automatic stimulus-response codings, but rather that the perception facilitates a more flexible and contextual response. The response that is facilitated seems to depend on the broader affordance structure perceived and how this structure fits with the task one is preparing to do.

One should notice that in these studies the perceptual attention might be very different in the complementary trial where attention is on the object and the action afforded as opposed to in the imitation trails where the cue is one the moving hand. This should first remind us that action perception and perception in general are not simply all-or-none phenomena that can be treated as objective snapshots of the environment happening independently of the pragmatic task at hand.[149] Furthermore, mirror neuron research rarely controls for attentional features or eye movements, which leaves too much interpretative guesswork in regards to what is actually perceived and what more precisely is driving the sensorimotor response in various cases.[150] One might expect researchers behind an 'automatic mirroring' account of motor resonance to argue that the reversal of

[148] Poljac, Schie & Bekkering (2009).

[149] A series of recent studies have also begun to show how our action primings and intentional preparations radically influence our perceptual processes, which again points to a bi-directional interplay between perceptual and motor processes and a problem of simply seeing motor resonance as a 'later stage' of embodied interpretation that builds on an already independently completed 'pictorial' perceptual representation as the language of the Parma Group often suggests. See here: Schutz-Bosbach & Prinz, (2007) and Fagiolo et al. (2007)

[150] As we have already seen, Tthis disregard for attentional variation raisesd, as we already saw, some issues in regards to the question of adaption and habituation of mirror neurons.

the mirror effect depends on lack of attention to the action of the other, that it is not the main target of perception so to speak. In other words, one might suggest that the van Schie et al. findings are due to a focus on object affordances and that the 'action observation system' or 'mirror mechanism' is still an automatic ubiquitous system that if triggered by action perception covertly activates a similar action in the observer. Here, it could be interesting to do experiments that test whether a complementary social action can be facilitated – like for example a fist hit to an open hand. If such complementary social interaction could be facilitated, then it would show that the imitation effect is relative to task and action goal, and that action perception does not ubiquitously yield a mirror response independent of the context.

In any event the findings of complementary action facilitation and the many non-researched possibilities alert to how problematic the contextual imitation bias is in the experimental protocols normally used to show automatic mirroring effects, as the conclusion to some degree seems to have been assumed in the experimental premises. Now evidence suggests that the affordance context of the observed action and the observer's own sensorimotor experience and goals in regards to perceptual situations all influence which actions are facilitated by particular action observations. This means that at the level of motor facilitation, there is not a ubiquitous imitation response. Rather, it seems that given the task context and experience perception of the other's action can facilitate a complementary response. However, there is a need for more studies showing that there could also be complementary externally induced – and more socially interactive - responses to action perception.[151]

B. Behavioral findings – ideomotor action

The sensorimotor training and complementary action affordances studied by Heyes et al. 2005 and van Schie et al. 2008 focus on how perceptually induced sensorimotor triggering interferes with responses. However, James' view of ideomotor action is not supposed to simply pertain to perceptually triggered actions but also to voluntary and endogenously motivated actions and initiated representations of action goal/outcome

---

[151] I also earlier mentioned the possibility of individual social affordance neurons where the other's action invites a compleimentary not necessarily an imitation response. Here, the suggestion is that it would also be interesting to see studies testing for sensorimotor compleimentary action facilitation ats a global motor level.

anticipations. Anne Häberle and her colleagues at the Max Planck Institute in Leipzig have recently published an extremely fascinating article investigating the extent to which ideomotor resonance induced during action perception is determined by the goal of the externally perceived actions and/or the perceiver's own goals.[152] I shall pause and discuss this experiment in some detail as I think it raises some important questions and provides strong evidence for both a dynamically integrated top-down and bottom-up view of the motor system and also more specifically the idea that other-centered actions are represented within a broader ego-centered affordance space.

Häberle and her coworkers configured a study where they could contrast the perception of others' actions under respectively cooperative and competitive circumstances, such that they could tease apart the influences by creating opposing outcome desires in the observed and observing agent. In other words, the experiment allowed them to ask the question of whether the ideomotor action is determined *top-down* by the *ego-centered* representations of the *desired action outcome* or *the actual other-centered action* goal of the perceived agent. Their study is based on their earlier findings that both the wishful ego-centered and other-centered predictive action representations can yield motor resonance or ideomotor action. For all these experiments they used an experimental paradigm of a computer bowling game where a ball or a target could be manipulated by the player at first and then the resulting movement was watched to see if the target would be hit or not. They write:

> The results of these experiments showed that both principles were indeed effective. On the one hand perceptually induced movements "copied" the movements of the ball on the screen. When participants saw the ball moving to the right they tended to move to the right…On the other hand, intentionally induced movements were related to the intention to hit the target… (Häberle et al. 2008)

After having shown that both the observed movement of the ball and the wishful intention to move the ball to hit the target resulted in ideomotor action, they measured and found almost the same ideomotor evoked muscle tensions while watching someone else play the same game. They then turned to the question of whether the ideomotor action resulting from the perception of others' actions would be different given respectively a cooperative and a competitive social context. Häberle and her coauthors

[152] Häberle et al. (2008).

refer to how mirror theories predict that in both scenarios the intention or goal of the observed agent that would determine the induced ideomotor action. They write:

> On these [mirroring] views motor representations are mainly used to reflect other-centered actions and intentions. However although the understanding of other persons is of obvious importance, it is only one part of social communication. The other part concerns the observer him/herself. Namely she has to select an appropriate response to the seen action of the other and not only put him/herself in a position to do the same as what is currently perceived. Depending on the context the observer has to activate a certain motor representation for responding in either a complementary (e.g. when dancing together) or competitive manner (e.g., when playing tennis with another person)…From this one can raise the question of whether ideomotor movements, as reflecting the activations of these representations, are primarily guided by other-centered or self-centered intentions. (Häberle et al. 2008, p.4)

Häberle and her colleagues here show the same concern in regards to the mirror theory assumptions that I have been expressing. There is a question as to why one should assume that the global motor resonance is only a product of mirroring. I have repeatedly argued that mirror theories make an illegitimate move from individual mirror neuron activations to ubiquitous imitative motor facilitation on a global level. The tacit – and in my eyes unwarranted - assumption is that the mirror response is the only automatic action tendency that action perception induces. Here Häberle and her coauthors rightly point out that such an other-centered monopoly of the motor response would be rather impractical for the perceiving agent, who needs a pragmatic representation of the action of the other such that it can be used to inform the observer's own ego-centered action choice. This finally leads to their empirical question of whether a social competitive context would lead to a primarily other-centered or ego-centered ideomotor response. They found that in a *competitive* context the intentionally driven ideomotor movement, which wishfully 'pushes' the bowling ball after the action seized to be effective, was indeed driven by the ego-centered intention of the observer (for the ball to miss the target) and not by the other-centered intention of the agent (for the ball to hit). They conclude:

> The observation that intentional induction is, in the case of conflict, self-centered and not other-centered seems to suggest that motor representations may not only be engaged in perceiving and predicting what the other is doing but also in planning and preparing one's own complimentary action. Functionally speaking, the point of understanding others' actions may be to prepare one's own complementary doings. (Häberle et al. 2008, p.10)

This study thus provides evidence for my hypothesis that we represent the actions of others within our ego-centered representation of the broader affordance space. Häberle and her coauthors here explicitly make the interesting suggestion that other-centered representations should be understood functionally as supporting our self-centered action preparations. I would want to add also the relation between action observation and self-centered action choice. An interesting element of this particular study is that they focus on intentionally induced movements, thus they take on the issue of the observer's own action drive and intentions and how that relates to the perceived action and the (sensori-) motor activity that this perception causes. They conclude that the perceived actions of others – and movements of external objects – affect our motor activity, but that in the case of conflict our own intentions and preferences drive the overall motor response. Hence, there might also be a bottom-up versus top-down message to take away from this study. It seems to me that this study supports the idea that our motor response is influenced by low-level externally/perceptually triggered bottom-up sensorimotor modulations and, at the same time, that our endogenous action goals impose a top-down modulation of the motor response. I shall return to this issue in Chapter 5 where I explore Csibra's argument that top-down motor control excludes the possibility of low-level bottom up mirroring as important for action understanding. This study by Häberle and her colleagues points to the possibility that low-level perceptually induced sensorimotor activity can support and scaffold our endogenously driven higher-level action choice. Therefore, rather than seeing the two as mutually exclusive, it seems that this study supports the idea of a dynamically integral function of bottom-up and top-down sensorimotor modulation. I shall argue that if one – as opposed to Csibra - allows for multiple simultaneous sensorimotor activations then a theoretical framework based on such dynamic perceptual and endogenous integrations and modulation becomes possible.

C. Behavioral findings – self versus other

The last recent motor resonance study that I want to discuss here is an ingenious experiment reported by Simone Schütz-Bosbach[153] and her colleagues. They used the so-called 'rubber hand illusion' to manipulate the sense of body ownership and made it possible via this illusion to vary the belief that an observed action was one's own versus another's while keeping constant the actual perceptual stimuli and produced motor

[153] Note that Schütz-Bosbach was also a coauthor on the Häberle et al. (2008) study reported above.

actions.[154] With this method they found that "observing another's actions facilitated the motor system whereas observing identical actions, which where illusionarily attributed to one's own body, showed to opposite pattern."(Schütz-Bosbach et al. 2006) This is an extremely interesting finding. Not only does it question the idea of an encapsulated agent-neutral mirror response to action perception. It also more specifically suggests that the perception of others facilitates similar actions whereas perception of one's own actions inhibits perceptually induced motor activations. Schütz-Bosbach and her coauthors propose that the "cortical suppression is a functional response to viewing one's own actions." They mention, as I have also already discussed, the idea that we have automatic inhibitory mechanisms in regards to the action facilitation produced by the perception of others' actions. However, these findings suggest that in the case of our own actions the inhibitory cortical suppression happens *before* we even see our own actions and therefore can preempt the facilitation altogether. Schütz-Bosbach and her coauthors write:

> This would be even more important for preventing inappropriate perseveration or entrainment when viewing one's own actions. In normal voluntary action, this observation-evoked suppression may be masked by the additional cortical facilitation associated with motor output. (p.1833)

In this passage they propose that the finding of cortical suppression of perceptually induced sensorimotor activations by one's own actions might be a general process in normal voluntary action. It certainly makes sense to pre-empt being automatically prone to repeat the same actions over and over again. However, one should note that in this particular experiment the subject is performing the same action - out of view - in all the trials, and is simply induced to think that the perceived rubber hand is his/her hand performing the action in some of the trials. Therefore, to me this study further suggest that we do not simply suppress *any* sensorimotor activation of some given 'action kind,' but rather that the suppression is narrowly linked to an individually perceived body and action in our affordance space that we judge to be our own. In other words, when we think that the rubber hand belongs to someone else, it facilitates a similar motor response even if we are aware that we ourselves are already performing the very same movement.[155] This study therefore shows not only that we distinguish the actions of self

[154] Schütz-Bosbach et al. (2006).
[155] These findings also align nicely with the repetition suppression studies suggesting that habituation from perception of others to one's own action, but not from one's own action to perception of a similar action done by another. See discussion of Lingnau et al. (2009) in the next section.

and others, but also that there is a complex dynamics, which allows us to monitor the actions of others even while we ourselves are performing similar actions.[156] It seems to support my idea that we do not simply represent 'action types' in an abstract 'motor vocabulary,' but rather as these are thought to relate to the overall affordance space. These findings regarding the differentiation between self and other might also be understood in the context of the earlier reported research done by Umilta et al. 2001 on individual visuomotor neurons. They found that some mirror neurons are modulated not just by the actual action perceived but rather dynamically integrate earlier perceptions and/or judgments of object permanence. Here such ideas are expanded to the social realm: the question is whether previous perceptions and inferences lead us to identify the perceived hand as belonging to ourselves or the other.

I am also intrigued by the idea that our own actions yield a rather precise cortical suppression of perceptual facilitation. As Schütz-Bosbach and her coauthors point out such a mechanism makes a lot of functional sense also in normal voluntary action. If our motor cortex was simply automatically activated by action perceptions, whether or not they were self or other produced, it seems that we would have a tough time distinguishing self from other on an epistemological level. More basically it is difficult to see how we would be able to orchestrate our own actions in relation to the affordance space we are in.[157]

Overall, there are important differences between *single cell* findings of action perception induced activity and behavioral findings of action facilitation at a global level.

---

[156] We shall see that Jacob and Jeannerod specifically try to argue against a motor role in action perception based on the reasoning that we would not be able to represent both our own actions and those of others simultaneously. Clearly they would have some trouble fitting these new findings into that narrow picture of the motor system.

[157] Thus beyond the distinction between self and other, it seems that we would have to distinguish internally prepared action plans from the external affordance structure. I think it would be very interesting to look into the possible connection between the intricate functions of such cortical suppression mechanisms as those involved in perception of one's own actions and the notorious problems of self-other control and perception-hallucination problems of people with schizophrenia. Traditionally there has been a focus on trying to tie these difficulties to a problem with a modular 'who' system, but the problems might be due to *dynamic regulations* rather than *specific mechanisms*. Some suggest that some hallucinations and feelings of alien control might be related to failures to send 'efference copies' of one's own actions to the STS and parietal areas. (See for example Frith (1996)) The findings of Schütz-Bosbach and her colleagues suggest that the idea of an efference copy as simply a motor command is problematic. The point is that if I can perform an action and simultaneously let the observation of another performing the same action facilitates such a motor response, then it is not simply a communication of a given kinetic motor response (efference copy/internal forward model). Rather, my claim is that this study shows that the cortical suppression must be based on a modulation of *sensori*-motor representations of an action as it relates to the overall perceptual affordance space.

It seems that human mirror neuron research based on motor resonance often make the assumption that the global action facilitation is a direct product of automatic and ubiquitous action mirroring at a single cell level. Often the finding of mirroring or imitative action facilitation is taken to support not only the existence of mirror neurons but also the idea that there is a ubiquitous covert imitation process going on. This latter conclusion totally ignores the possibility of a broader Jamesian idea of competing action parallel representations. Accordingly, it might be that some mirror neurons automatically respond to action perception and lead to some sketchy representation of the observed action. However, given the pragmatic, motivational and social context this representation might never lead to further motor processes but early on be out-competed by other action representations and ego-centric affordances. Now we have seen that multiple recent findings seem to support this idea.

3.5.3. Human neurological evidence - findings and questions

Over the last decade there has also been an accumulation of neurological evidence in regards to the function and sensorimotor activity of fronto-parietal circuits in humans. The question that one can ask of this research like the behavioral findings is how exactly fMRI and other imaging techniques speaks to the theory of a mirror mechanism and the existence and nature of mirror neurons in humans. I shall in this section discuss some of the methods and findings and the issues and questions that arise from these.

A. <u>The focus on mirroring and ‘shared representation’</u>

A key feature of the fascination with the mirror neuron findings in macaque monkeys is that the visual and motor integration happen at a single cell level. The single cell findings of mirror neurons in monkeys showed that the action observations somehow lead to a motor resonance or simulation process in the observer. But the game-changing effect of the discovery owes to the idea that *individual* mirror neurons do not differentiate between observed and executed actions, and thus that these neurons each provide a mini ‘mirror mechanism’ by way of which the observer automatically and directly channels the observed action in their own motor system. As suggested earlier, it is the proposed ubiquity and agent-neutrality of the response of individual mirror neurons that guarantees that it is the action of the other that is directly matched. And it is that the mirroring activity is interpreted as yielding a shared representation of the action that is hypothesized as essential to so many social cognitive functions. I argue that the obsession with looking for a shared representation of a symmetry between action observation and execution actually obscures and derails much of the research, and also the understanding of the role of fronto-parietal circuits in social cognition. However, that is because I see the social function of mirror neurons as linked to the broader sensorimotor circuits. Given the traditional functional interpretation of mirror neurons as yielding a social mechanism based on symmetry and shared representation between self and other, things are different. Here, it becomes pivotal to show empirically that there is an overlap between the neurological activity during action and perception of actions. And this is exactly the project that guides most of the human 'mirroring' imaging studies.

I have in the previous sections already discussed various behavioral evidence for motor resonance. I tried to show that whereas motor resonance findings tell us something about sensorimotor integration and about how perception induces action, they do not guarantee that self and other are represented in the same way in the cortical motor system, and thus motor resonance does not as such support the idea of a 'shared' representation.[158] In regard to proving a shared representation, one might say that the problem with the behavioral evidence is that it is a global response that could be caused by various different low-level processes and does not need to be based on the same neural activity in the case of perception and action. To put it differently, it is an open question whether such motor resonance in humans is a product of activity homologous to monkey mirror neurons.

I have argued already on the basis of the single cell findings that the caricatured theory of mirroring as *agent-neutral, ubiquitous* and *context independent* is rather implausible. And my theory would not assume a neat agent-neutral overlap between activity during action and perception in the fronto-parietal circuits.

[158] Within the ideomotor and motor resonance literature, people often use the idea of shared sensorimotor representations and of a common 'coding' basis for perception and actions. (e.g. Hommel et al. 2001) However, as I see it these notions are here used more loosely to indicate that the same pathways and circuits are serving multiple purposes. And the 'mirror' claim that the neural activation is identical or 'agent neutral' during execution and observation or is not integral to these more basic ideas of sensorimotor integration. The main point of common coding proposals is to deny the typical dichotomy between perceptual and motor representations and not between self and other.

Lately, a series of articles have been published which stress the lack of direct evidence for mirror neurons in humans.[159] I think that these articles are important as they force the debate in the area to actually question some of the many assumptions and speculations that human mirror neuron research seems replete with. More specifically I think it is highly problematic that rather grand conclusions are being drawn about exact mirror functions from overlaps of BOLD signals in the fronto-parietal circuits in general, averaged over multiple trials and people, and contrasted with control studies with their own series of assumptions.[160] Thus, even though I find it highly implausible that the monkey and human fronto-parietal circuits would be so discontinuous that there would not be mirror neurons in humans – in the broad sense of action sensitive visuomotor neurons, I do want to look at some of the recent fMRI evidence for and against the existence of mirror neurons and mirror activations in humans. The point is that this discussion might inform a more precise picture of both the phenomena and the many assumptions that drive the research.

Marco Iacoboni's lab at UCLA has by way of fMRI produced various interesting findings in regard to action, and goal/intention representations in humans. I shall discuss some of their groundbreaking research in more detail in the Chapter 4 (Section 4.3.1). For my present purposes the point is that many researchers, based on some rather generalized findings of activation overlaps, have pretty much assumed that mirror neurons exist in humans, and thus many important labs in the field have focused on mirror neuron properties and functional correlations rather than the narrow nature of the execution-observation overlap. However, before an overlap is documented it does seem quite problematic – particularly given the heterogeneous sensorimotor findings in the fronto-parietal circuits - to assume that increased or decreased activity in the areas of the human 'mirror neuron system' is an expression of mirror neuron activity. To put it differently, it is rather rash to denote the fronto-parietal areas as the mirror neuron system.

[159] See Dinstein et al (2008) and Lingnau et al (2009). I will discuss these findings in detail in the following.

[160] Here we lean up against the big topic of what exactly fMRI scans can and cannot tell us, but I shall pass on that discussion and assume for the sake of argument that the BOLD signal can be used rather accurately to trace average neuronal activity in local areas.

B. The mirror challenges & repetition suppression studies

But how does one isolate mirror neurons or a mirror neuron system via imaging techniques? Due to the movement restrictions of an fMRI scanner, very simple experimental protocols are normally used. Most often there is a comparison between a rest condition, a passive picture or video-clip action observation (maybe a simple finger or grasping move) and an action execution condition and sometimes additionally an imitation condition. One should note that these are extremely isolated and asocial scenarios, and the results might therefore say very little about how these circuits work during actual social interaction, but let us leave that issue aside for now and focus on these methods as a means to detect and locate mirror neurons in the human brain. Rizzolatti and Craighero for example report a series of experiments suggesting that action observation-execution overlaps can be found in premotor and parietal areas homologous to the areas in which mirror neurons were found in monkeys.[161] Valeria Gazzola and Christian Keysers have also published evidence that there is an activation overlap between relatively small anatomical motor and somatosensory voxels during action and perception in individual subjects.[162] However, this sort of imaging of anatomical activation overlaps is clearly tricky in regard to single cell conclusions.

To accommodate this limitation of fMRI a new method called 'repetition suppression' is gaining popularity within the field. This sort of experimental paradigm builds on the premise that neurons adapt to repeated stimuli and that one therefore can use the imaging of adaptation contrasts to test whether a specific stimulus is judged to be new or repeated by various neurons. Interestingly, as I have mentioned earlier, Gallese and his colleagues suggest that the mirror neurons do not adapt. But now it seems to be assumed that all sensorimotor neurons do adapt and the repetition suppression method has become an influential protocol in fMRI studies of perceptually induced motor resonance. This change is in and of itself rather interesting as I argued the lack of habituation or adaptation reported by Gallese and his colleagues fit nicely with the idea of little individual ubiquitous mirror modules. The question is whether the scientific acceptance

[161] See Rizzolatti & Craighero (2004).

[162] Gazzola & Keysers (2008). The observation and execution of actions share motor and somatosensory voxels in all the tested subjects - single-subject analysis of unsmoothed fMRI data. *Cerebral cortex*. Doi:10.1093/cercor/bhn181. Notice that these areas are outside the traditional mirror neuron areas, but fit with the discovery of mirror neurons in primary motor areas in monkeys. See Koski et al. (2002).

of the adaption paradigm already is an indicator that the activity of single neurons do not easily translate to theories of modular cognitive mechanisms. At a *global level* we have seen how one's own actions often are facilitated by perceptually induced motor resonance.[163] At a local neural level adaptation findings suggest that the responses to repeated stimuli are attenuated rather than facilitated or increased. Though the behavioral studies do not directly quantify motor facilitation over multiple similar exposures, they do hint that one should be careful in making inference between findings at various levels – for example firing rates of individual neurons and motor resonance.

Several studies have lately been conducted with the methodology of repetition suppression to attempt to prove and disprove the existence of cross-modal action-execution adaptation and thus provide evidence of a human mirror neuron system. As I see it, however, a problem with these studies is the assumption that the cross-modal sensorimotor integration of mirror neurons is at the same level of description as adaptation to repeated external stimuli or overtly executed actions. In other words, on my theory there might be single cell visuomotor integration that pertain to some teleological or goal-directed aspect of the action within single neurons that is playing a role in both action perception and action execution, but which would not necessarily habituate cross-modally depending on the broader neuronal context of the activation. The empirical question is whether visuomotor integration within single neurons has to make these neurons 'blind' to whether the current modulation is caused by perception or action. Differently put, there seems to be a question as to the assumption that mirror neurons would necessarily show cross-modal adaptation. Rizzolatti has in personal conversation pointed out that repetition suppression is likely to happen at the synaptic level, which leads me to think that the process is highly contextually sensitive. But be that as it may the assumption that repetition suppression yields an accurate picture of observation-execution overlaps is again rooted in the assumptions of the narrow mirror metaphor, i.e. of agent neutrality and similarity from within the mirror mechanism on the one hand and the idea of the response depending narrowly on the action type independently of the context on the other. I have in this chapter tried precisely to problematize the empirical basis for these assumptions. In regard to cross-modal repetition suppression I contend

[163] Schütz-Bosbach & Prinz also have an extremely informative article on action-induced facilitation of perception. See Schütz-Bosbach & Prinz (2007).

that the lack of such might challenge the existence and location of potential modular and strictly congruent mirror populations. But I shall in the following argue that the existence of contextually and agent modulated action sensitive visuomotor neurons might actually be supported by the lack of cross modal action to perception adaptation. But this point is easier understood through the details of findings themselves.

Recent findings suggest that the neural activity overlap might not be exactly the same between action and perception conditions, and thus challenge the existence of an agent neutral observation-execution matching mechanism in humans. Ilan Dinstein and coauthors published an interesting essay in 2008 called "A mirror up to nature," in which they argue that the evidence for mirror neurons both in monkeys and in particular in humans is rather weak.[164] But their main focus is the weakness of fMRI evidence for a human 'mirror system.' I share many of their concerns, for example the critique of studies that interpret any activity in an area thought to contain mirror neurons as evidence for such neurons. This is of course unwarranted when it was clear even from the single cell experiments in monkeys that mirror neurons only represent a fraction of the total neuronal population in the implicated areas. In general Dinstein and his collaborators argue that fMRI imaging is too imprecise to provide evidence for individual mirror neurons. They argue along the same lines that I have that, given the typical experimental protocols used, activity change in an area could have multiple causes beyond mirror neuron activity and also that the seeming lack of activity modulation might cover an averaging out of local increases and decreases in activity.[165] The new empirical contribution of Dinstein and his coworkers is that they repeatedly have used the repetition suppression paradigm to try to investigate potential mirroring areas. In their 2007 journal of neurophysiology article 'Brain areas selective for both observed and executed action' they report 5 areas, which show adaption for both action observation and execution.[166] Still, they report that they found themselves "unable to demonstrate the existence of

[164] They argue the monkey findings "suggest that mirror neurons form a distributed representation of observed and executed movements." (Dinstein et al. 2008) I think this might be a good way of talking about the findings and it takes the focus away from single neurons and the idea that individual neurons represent actions and have individual functional roles independently of their immediate neurological context.

[165] Remember here again my earlier discussion of the findings that some of the individual mirror neurons originally found in monkeys had high spontaneous firing rates and thus when modulated decreased rather than increased their activity.

[166] Dinstein et al. (2007).

mirror neurons in the human brain as we did not find any cortical areas exhibiting cross-modal adaption.[167]

I shall here discuss a similar recent study by Angelika Lingnau and her colleagues at the center for Mind/Brian studies at the University of Trento in Italy. Their 2009 article is entitled "Asymmetric fMRI adaption reveal no evidence for mirror neurons in humans" and is already raising some eyebrows and causing a great deal of blogging activity as it claims to have produced evidence against the existence of mirror neurons in humans.[168] Briefly put, Lingnau and her coworkers used the repetition suppression paradigm to see whether action observation and execution would yield the cross-modal adaptation that one would expect if the individual neurons where truly agent neutral and treated observed and executed actions as the same stimuli. They purposefully used meaningless intransitive actions in this study as they suspect that meaningful actions might yield an indirect semantic action priming - which is a whole new can of worms by itself.[169] On the basis of this intransitive action paradigm they summarize the findings and their importance as follows:

> We found adaptation for motor acts that were repeatedly observed or repeatedly executed. We also found adaptation for motor acts that were first observed and then executed, as would be expected if a previously seen act primed the subsequent execution of that act. Crucially, we found no signs of adaptation for motor acts that were first executed and then observed. Failure to find cross-modal adaptation for executed and observed motor acts is not compatible with the core assumption of mirror neuron theory, which holds that action recognition and understanding are based on motor simulation. (Lingnau et al. 2009, p.1)

Thus, they see their findings as evidence not only against mirror neurons but also more broadly against the idea that we use motor simulations to understand the actions of others. Given this radical conclusion, I will look into the findings that are suggested to be

---

[167] Dinstein et al. 2007, p. R16. To see other of their articles with similar results: Dinstein et al. (2008). See also Agnew and Wise (2008). For general motor adaption evidence, see Kroliczak et al. (2008). Another interesting Kroliczak study includes pantomimed actions.

[168] See Lingnau et al. (2009).

[169] This is of course problematic in that the mirror neurons documented in monkeys exactly did not respond to such actions. But Lingnau et al. argue vehemently for this choice. They criticize a study by Chong et al. 2008, which via a similar repetition suppression methodology claims to have found cross-modal adaption and thus *evidence for* mirror neurons, as problematic exactly because the stimuli used where we are aware that this choice might have some consequence for activity in the premotor-parietal areas as mirror neurons in monkeys preferably are modulated by goal-directed actions. Here it could be interesting if one could perform a similar repetition suppression study in monkeys where one knows that mirror neurons exist. And in any event these difficulties in regard to meaningful actions in my eyes seem to raise some problems for the repetition suppression paradigm in general as a method for evaluating individual goal representing neurons.

so detrimental for the mirroring hypothesis. First, they found that repeated observed or executed actions each produced adaptations. Further, they found that if one first observes an action and then produces it then there is an adapting response as well. Lastly, their crucial finding was that there was no area in which action execution leads to adaptation for action execution. I think that these are very intriguing findings, but I am suspicious that the conclusions might be somewhat overblown. More specifically, they seem to rest on the *caricature model of mirroring* (and possibly on the rather specific notions of motor simulation, which I shall further criticize in great detail in chapter 5). The cross-modal adaptation from observation to execution suggests that the neurons in the areas in question represent the executed action as in some significant way similar or maybe not new in regards to the previously observed action. Lingnau and her coauthors refer to this adaptation as reflecting a 'motor priming.' They write:

> If the same motor act is first observed and then executed, the same motor command that has been covertly generated during action observation can be used for overt motor execution. This might lead to benefits for the selection, preparation, and execution of the required motor command…(Lingnau et al. 2009)

I agree, and under my interpretation the perceived action would lead to an activation of a rough sketch of the goal and a schema for motor and perceptual connections that would be represented as the action of the other in my affordance space. To produce an action this sketch would need to be specified in the spatiotemporal terms in regard to the current context as it presents itself from my perspective. This is however a more involved story than the one given by Lingnau and her colleagues, and they suggest further that the lack of cross-modal adaptation in the other direction – from action to perception - is detrimental for the mirror hypothesis.[170] Well, if we are talking about the caricature model it certainly looks like a serious blow, but if mirror neurons should be understood within a broader functional network of sensorimotor integration I think one can explain - if not even expect - this asymmetry of adaptation. We saw previously in the discussion of Schütz-Bosbach et al. (2006) similar findings of a motor resonance asymmetry between the actions of self and other. They reported that action perception only facilitated

[170] They write: "These findings challenge the direct matching account of action understanding, according to which neurons should exist that adapt if the same motor act is repeated, both when the motor act is observed and when it is executed. Our findings are compatible with the view that activation in mirror neuron areas reflects the facilitation of motor programs as a consequence of action understanding."(Lingnau et al. 2009)

action if it was judged to belong to someone other than oneself. The present finding seems to support this asymmetry. My explanation of these findings is that fronto-parietal visuomotor neurons form representation of the social and physical affordance space that we are in, and that that the action asymmetry is essential for this relational representation. More specifically it would be a problem if my own actions would prevent me from perceiving similar actions by others, or confuse me to think that my own visually perceived actions belong to others. Thus, both these findings of asymmetry make sense: when I perform an action the anticipated visual perception is disregarded (Schütz-Bosbach et al. 2006), but does not prevent me from regarding a perceived token of the same action type as new information and as belonging to others – as taking a different relational place in my social affordance space.

To sum up, in regard to mirror neurons my point is that these findings challenge the caricature model of modular agent-neutral and context independent action type mirroring – not the existence of action sensitive visuomotor neurons in humans like those actually found in monkeys.

Thus, a different question would be whether mirror neurons in a more broad sense of 'action sensitive visuomotor neurons' can be found in homologous fronto-parietal areas in humans. This is of course an empirical question – and a notoriously difficult one given the pragmatic and ethical experimental limitations. But I would guess that given the extent and central function of sensorimotor integrations in primates, it will be enormously shocking if mirror neurons in this broad sense cannot also be found in humans. Given my proposal of a broad spectrum of visuomotor integration serving the monitoring and predicting the goal-directed actions of others in relation to the broader affordance space, I would suspect that humans like monkeys will have some amount of relatively congruent mirror neurons in addition to the cacophony of other visuomotor neurons whose existence I have hypothesized in the earlier parts of this chapter (see my flowchart in section 3.2.5 A). But, again, it is important to emphasize that my interpretation of mirror neurons exactly does not predict findings of overall agent-neutrality, or precise overlaps of activation between observation and execution or context independence of the modulation. Thus, while my proposal certainly hypothesizes the existence of action sensitive visuomotor neurons in humans, it does not ride on the existence of strictly congruent mirror neurons in the narrow sense of individual modules of exact, context independent and agent neutral observation and execution matching – and cross-modal adaption.

C. A modular mirror mechanism?

Related but slightly beyond the issue of the existence of single mirror neurons is the question of the notion of a 'mechanism' for mirroring, for retrieving ubiquitous context independent agent neutral action representations. In order to prove that we have such a modular mirror mechanism one should expect a rather tight overlap between action execution and observation. As we have seen, the problem in regard to fMRI results is to tease apart the activations that are a product of the overt execution and potential perceptual differences in perspective, which might be responsible for noise and could obscure a possible central overlapping mirror response during action observation and execution.

However, a few studies seem to have questioned at least the claim of context and task independence of the mirror response. Thus, Newman-Norlund and his coauthors have reported that the proposed mirror neuron areas are more active during complimentary than during imitative action.[171] Georgiou and colleagues have similarly found diverging activations for cooperative and competitive behavior.[172] And lastly but not least interestingly, Cheng and colleagues have found an effect of motivation on mirror area activity.[173] I shall not discuss all these very interesting findings but I simply alert to the complexities they bring out. In the present context of brain-imaging research an interesting point is that the inconsistent findings regarding the overlap between action perception and action execution modulations might be reconciled under a less caricatured model of sensorimotor integration. My proposal is that it might be that perceptually induced motor resonance, as these and other studies suggest, is a much more complex and contextual process than the narrow mirroring mechanism metaphor suggests. Thus unanalyzed contextual, pragmatic and motivational factors might account for many of the seeming inconsistencies in the fMRI findings of sensorimotor resonance during action execution and perception.

[171] Newman-Norlund et al. (2007).
[172] Georgiou I., Becchio C., Glover S., Castiello U. (2007).
[173] Cheng et al. (2007).

Summing up, the neurological evidence for homologous mirror neurons and circuits in humans is indirect and mainly stems from fMRI neuroimaging studies. The main foci of early fMRI mirror neuron system research was first to show that in humans there *exist* circuits with activation overlap between instances of execution and observation of the same actions. Secondarily, the goal was to use the potential action observation-execution overlap to pinpoint the anatomical locations of the mirror neuron system in humans. Some key studies have also explored various questions about the overlap between the observed action and the observer's motor repertoire, about the motor and visual features necessary for motor resonance etc. But the problem is that these studies of activation overlap do not really prove the existence of single mirror neurons or a ubiquitous modular agent-neutral context independent mirror mechanism. Rather, recent fMRI studies seem to have seriously questioned each of these 'mirror metaphor' inspired features.

3.5.4. Clinical cases; imitation behavior, autism & neonate imitation

The last area of empirical support for a human mirror mechanism that I shall mention is the clinical evidence that is typically seen as providing evidence for a ubiquitous agent neutral mirror mechanism. I will here only very briefly introduce the evidence from imitation behavior, autism and neonate imitation, as I will revisit these issues in my discussion of the simulationist interpretations of mirror neurons and clinical findings relevant to such processes in the next chapter.

A. Imitation behavior

As shown in this chapter, the reasoning of the Parma group often tacitly assumes that the observation-execution mirror matching process is ubiquitous and agent neutral, and thus that action observation always at some level results in an activation of the same action in the observer. It is obvious that we do not always or even very often overtly imitate all the actions that we observe, and therefore an important additional element of the mirror mechanism hypothesis has emerged, as mentioned earlier, namely that we have separate processes that differentiate observed actions of others from self-initiated actions (the 'who' system) and processes that inhibit the former. Mirror neuron theorists have pointed to various clinical observations as evidence that these processes can break down and thus by implication that both a mirror mechanism and these cognitive support mechanisms must exist in normal agents and observers.

As I shall discuss in greater detail in the next chapter, the syndromes of 'utilization behavior' and 'imitation behavior' found in some patients with frontal lesions are often characterized as the 'release' of the ubiquitous object affordance and mirror matching processes instantiated by canonical and mirror neurons respectively.[174] Imitation behavior is described as a syndrome in which patients lack the ability to inhibit the impulse to imitate all the observed actions or action goals. This might be a problematically simplified story since another important characteristic of these patients, according to Erik Rietveld (2008), is that they seem to lack any kind of independent action motivation or inspiration and also seem utterly unembarrassed by their inappropriate imitative actions.[175] Hence, thinking again back to Williams James' ideomotor ideas, it could be that they find the observed action suggestive not so much because it is automatically mirrored but because there are no other competing action plans or concerns. This latter interpretation is also supported by the fact that these patients are generally very apathetic and that this syndrome might better resemble the lack of action initiative in a kinetic mutism than other inhibitory or involuntary 'release' problems like for example 'anarchic hand' syndrome. Given the idea of multiple competing sensorimotor integrations, one might suggest that imitation behavior and utilization behavior do not result from a release of actions, which in people with normally functioning prefrontal cortices are ubiquitously fully simulated but inhibited. Rather, the problem might be that motor responses which normally only produce weak affordance activations are left without any alternative competing actions or motivated and engaged action choices. These random action invitations are then carried out due to lack of alternatives in a motor system that is left relatively idle.

Thus, what I want to stress is that whereas imitation behavior does suggest that we automatically have some level of evoked imitation response, it certainly does not provide evidence that unimpaired people covertly carry out full imitations of other people's

[174] See Rizzolatti et al. (1999). In this article Rizzolatti and his coauthors refer not only to imitation behavior but also to 'echopraxia' which is another syndrome in which some severely demented patients show reflex-like imitation of other peoples gestures.
[175] This interpretation is suggested by Rietveld (2008). See my discussion of his interpretation of imitation behavior in section 4.2.7. A in the next chapter.

actions. Thus, imitative tendencies might provide some evidence of sensorimotor 'mirror' integration, but fall short of being evidence for any kind of ubiquitous agent neutral mirror mechanism.

B. The autism connection and the 'broken mirror' theory

Another enormously important use that mirror neuron theorists are making of clinical cases is the discussion of the possible relationship between the mirror neuron system and the range of social impairments in people with autism. I shall also discuss this idea in greater detail in the upcoming chapter. Briefly stated, the central idea in the literature is that since the mirror neuron system is thought to support basic action understanding processes via providing an embodied shared action representation between agent and observer, it could be that many of the social and linguistic problems characteristic of the autistic spectrum could be due to a 'broken mirror' i.e. a dysfunction in the mirror neuron system. The debate over a possible autism-mirror neuron link is obviously very complex. For now I would simply note that autism is thought to inform the possible social cognitive function of a mirror neuron system and vice versa. As I will be advancing a different story of basic social cognition and 'theory of mind' abilities, I will also give a rather different explanation of symptoms in the autistic spectrum. (See section 4.2.8 B)

C. Neonate imitation and the innateness of mirror neurons

Lastly, I want to mention the often-made reference to the findings of Neonate imitation. Meltzoff and Moore (1977) famously demonstrated that neonate infants can imitate some facial expressions in particular tongue protrusion, which they have never had any visual feedback on. This suggests that in regards to these simple actions there is an innate 'hard-wiring' of visuo-motor mirror integration and that the perception of the experimenter's tongue protrusion therefore automatically can trigger a similar action in the newborn.[176] I have earlier discussed Cecilia Heyes and Caroline Catmur's findings regarding the experiential base of sensorimotor integration and more precisely their suggestion that mirror neurons are not hard-wired but can be reconfigured.[177] Heyes has argued that

---

[176] Similar experiments have also been conducted on neonate macaque monkeys and there is a rather exciting You-Tube video of these experiments: http://www.youtube.com/watch?v=k72WFYv6WMw.
[177] See section 3.5.2.

Meltzoff's findings might have been somewhat exaggerated and that people have too liberally generalized from very limited and narrow findings of action imitation in the case of tongue protrusion to a much broader conclusion of neonate imitative abilities.[178] As mentioned, Heyes and Catmur's findings on sensorimotor reconfiguration are mostly dealing with action facilitation on a global level, and they therefore have not proven that any or all strictly congruent mirror neurons are as malleable as the overall action facilitation. On the other hand, I do think their questions to the conclusion of hardwiring based on neonate imitation are sobering. It could be that there is only little and very limited neonate visuomotor integration.[179]

I think that it is important to discuss neonate sensorimotor integration beyond the question of mirror neurons. In the cases of facial imitation and tongue protrusion one should note that there clearly also is some non-mirroring visuomotor integration at work in newborns. Looking for example at the video of imitation in new born macaques, it is clear that the imitation rides on another and maybe even more primitive visuomotor function, namely, the ability of the newborn monkey to obtain eye contact and 'locking in' their visual attention on the face of the experimenter. My point is that there is a larger story to tell in regard to the fascinating abilities of neonates that allows them to imitate tongue protrusion rather than simply talk of observation-execution matching or strictly congruent mirror neurons firing – namely a response to the social affordance of attention based engagements with the other. These intriguing developmental issues of the establishment of sensorimotor connections are entirely beyond the scope of my present project.[180] I simply what to suggest that whereas neonate imitation might point to some innate sensorimotor mapping between execution and observation, it might be very limited and not give us much in terms of a more general theory of a mirror mechanism for observed and executed actions.

[178] These views of Heyes are based on a conference presentation and not published articles.

[179] Given the myriad of connectivity in the neonate brain, neonate imitation could be a by-product of this pre-pruning state along with an imitation syndrome and an underdeveloped frontal lobe.

[180] It is known that infants are born with enormous amounts of synaptic connections that are then 'pruned' back by experience and sensorimotor learning. An interesting possibility could be that multimodal and mirror like connections are actually more rampant at birth and then more sparse and specific later. This would fit with Meltzoff and Moore's findings that neonate imitation is a phenomenon that appears in a short window of time after birth and then imitative abilities reappear in a more sophisticated version months later.

All in all, I have in these latter sections discussed the move from the view of idealized strictly congruent mirror neurons in monkeys to the idea of a general mirror mechanism in humans. I have tried to show that the metaphor of mirroring as encapsulated, agent-neutral and context independent, which was already problematic in regards to the individual mirror neurons in F5, is even more so in regards to the complex functioning of parietal-premotor circuits at large. Furthermore, this idealized mirror mechanism seems to obscure the relation and difference between monkey and human social cognitive skills. Without looking at the complexity of sensorimotor integrative processes, the human mirror neurons system research becomes a matter of simply finding activation overlaps of perceptual and motor functions, which then opens the door for the suspicion that such neurons or such a mechanism might not even exist in humans.

3.6. Summary and conclusion

I have in this sizable chapter hopefully succeeded in showing the empirical untenability of what I have called the caricature model of mirroring, i.e. of mirror neurons as instantiating a ubiquitous agent-neutral and context independent observation/execution matching mechanism. The detailed analyses of the tensions between the preconceived notions of the cognitive mechanisms in question and the actual findings were presented to support my criticism of the implicit baggage of the mirror metaphor and what I have called the 'caricature model' of mirroring. But more importantly, the traced tensions are used to carry my argument of a reinterpretation of mirror neurons as not supporting an isolated mechanism but as integral to much broader sensorimotor processes, which underpin our pragmatic understanding of the physical and social affordance space we act in. My pinpointing of the tensions is therefore meant to guide to an understanding of my alternative theory, but also to give systematic reasons for why it has so far been ignored as a possibility. I have therefore also in this chapter begun to show how my reinterpretation has wide consequences for the way one interprets the role of mirror neurons in various social cognitive abilities. The point is that even to see 'affordance monitoring' as integral to social cognitive function one needs to revise many entrenched assumptions about the nature of social and motor cognition.

This chapter is also meant to setup the next two chapters in which the cognitive ramifications of the diverging views on sensorimotor integration will be explored. Firstly, in Chapter 4, this will be done via the proposed and enormously influential link between the simulation theory of mindreading and the mirror neuron discovery. And then, in Chapter 5, it will be continued in my discussion of the structure and assumptions of the prevalent criticisms of the role of mirror neurons in social cognition.

Chapter 4

Mirroring, simulation and '3rd via 1st person' mindreading

4.1. Introduction & outline

I have in Chapter 2 discussed the interpretations of the early mirror neuron findings and the general studies of sensorimotor integration, which served as a context for their discovery. I showed how the theory of these neurons as instantiating an observation/execution 'mirror matching' mechanism, was intertwined with a series of problematic decisions of classification and terminology. In Chapter 3 I outlined the resulting 'caricature model' of an agent neutral and context independent action representation mirror matching mechanism, which I argued implicitly has strongly influenced the mirror neuron research. I traced the specific tensions between this model and the empirical findings of sensorimotor integration. I will to devote this present chapter to the primarily theoretical development and change that happened with the mirroring hypothesis as it was linked to the idea of action simulation.

This chapter has two main sections. The first section focuses on the 1998 TICS article by Vittorio Gallese and Alvin Goldman, which brought the mirror neuron findings and their proposed mechanism together with the philosophical theory of mental simulation also known as the simulation theory of mindreading. The meeting between these two strands of thinking and research has had enormous consequences for the subsequent debate about mirror neurons and their possible role in the understanding of other minds and social cognition. Pivotally, I am critical of many of the assumptions that shape the frameworks of 'mindreading' and simulation, and which are brought into the mirror neuron debate by this meeting. Most importantly, I think that it is misleading to think of social cognition mainly in terms of a disengaged 3rd person process attributing and predicting hidden/1st person mental states.

But, in order to untangle mirror neurons from the assumptions of the 'mindreading' framework, I need to show how they got tangled up to begin with, and I therefore see the discussion of Gallese and Goldman's joint-venture not only as essential to the presentation of my alternative affordance view but also to my analysis in the next chapter of how many other mirror neuron critics seem to stay within this problematic mindreading framework.

The second and shorter section of this chapter on simulation focuses on later attempts by mainly Gallese[181] to reinvent the idea of mirroring as simulation in less explicit and less Goldman-style mental state attribution centered terms. Gallese presents a broader theory of mirroring as 'embodied simulation' attempting to distinguish this idea from the higher level, explicit and intentionally controlled pretense simulation theory of Goldman. In regard to motor mirroring the mechanism is now not simply conceptualized via the properties of strictly congruent mirror neurons but based on the idea that the various fronto-parietal neurons form 'prewired action chains,' which can be triggered by action observation. On his notion of embodied simulation Gallese also wants to conceive of what he calls a unified theory of social cognition and suggests that mirroring processes ground not only action understanding but also emotional processes, empathy and conceptual linguistic understanding. I point to the many constructive thoughts and improvements to earlier versions of the mirror mechanism and the idea of mental simulation, but I also argue that Gallese's new theory still in many ways fails to break up with the key problematic assumptions. Mainly, my problem is that Gallese still focuses on what I have labeled the '3rd via 1st person' understanding of social phenomena. The 'unifying mechanism' of social cognition is conceived of as a mirroring process of embodied simulation of the other's action, emotional experience, linguistic meaning etc. In other words, the simulation yields a shared representation and experience and it is this mental symmetry that grounds our social understanding. I argue that this view ignores 1) the importance of social interactions and how we understand others in relation to ourselves, and 2) that it continues the trend of trapping mental content within the 1st person experience. Hence, I argue that despite plenty of gestures in that direction Gallese does not theoretically explain how sensorimotor integration processes like those of the fronto-parietal circuits are fundamentally relational in a way that transcends 1st person

[181] Even though somewhat similar ideas have been proposed by Iacoboni, Rizzolatti and others, I have chosen to focus on Gallese not only because he was Goldman's original collaborator, but also because he is the most outspoken and involved in the theoretical simulation issues.

introspection and to some extent externalizes the 'content' of intentions, perceptions and even 'beliefs' and 'desires.'

Overall this chapter provides an analysis and critique of mirroring as a simulative and temporal process and of social cognition as based on mindreading as '3rd via 1st person' mirroring simulation. We shall see in the next chapter that the main critics of the importance of mirroring in social cognition all maintain the idea of mirroring as 'as if' action simulation and social cognition as primarily a matter of 3rd person mindreading. Therefore, I use the discussion of these ideas of simulation and mindreading in both this chapter and the next one to try to present my alternative view not just of the social function of mirror neurons, but more broadly of what sensorimotor integrations and simulations and/or basic social cognitive abilities consists in.

4.2. <u>The affaire fatale of mirror neurons and simulative mindreading</u>

In the now classic article 'Mirror neurons and the simulation theory of mindreading,'[182] Gallese and Goldman make a push to bring together Goldman's theory of simulative mindreading and Gallese's 'observation-execution matching system' interpretation of his and his Parma lab colleagues' groundbreaking mirror neuron discoveries. This marriage – or affaire fatale if you like - of simulation theory and mirror neurons has had enormous influence on later interpretations of and attitudes to mirroring theories and research.[183] I will discuss not only the actual proposal of Gallese and Goldman, but also see the subtleties of how these two theoretical branches came to influence each other and arguably still do. On the surface Gallese and Goldman's joint-venture was short-lived and they have since each pursued their separate ideas of both simulation and mirroring. However, Gallese inherits parts of the procedural framework of simulation and of the conceptual framework underlying the idea of mindreading as mental state attribution, and this I argue complicates rather than helps our understanding of the role of mirror neurons and fronto-parietal circuits in social cognition. To make these points I will first say a bit more about Goldman's simulation theory and Gallese's possible motivations for teaming up with him.

[182] Gallese & Goldman (1998).
[183] The article is one of the most cited TICS article ever (1150+ reported citations).

4.2.1. Social Cognition via 3rd person mindreading & 2nd person interaction

Goldman has over the last two decades developed and argued for his theory of 'mindreading' as based on simulation. The central idea of simulation is that we use our 1st person cognitive resources to, so to speak, pretend to be in the other's shoes. He opposes this theory to the so-called 'theory theory' which takes our abilities to mind-read, i.e. to attribute and predict the mental states of others, to depend on a 'theory of mind' and so-called folk psychological heuristics and inferences.[184] Goldman's simulation theory is thus conceived in opposition to the old 'theory theory,' i.e. the view that we understand other minds primarily by way of some sort of theoretical inferences and heuristics. But, it is worth noticing the common starting point of the two theories; namely, that they both take the issue of understanding others as mainly being a question of 'mindreading,' which normally is described as the ability to attribute and predict the mental states of self and others. Accordingly, the dispute between the two theories is about how and by means of which cognitive processes this mental state attribution takes place. Their common key question is how we on the basis of observation of *overt behavior* come to understand others' *hidden mental states*, and both the theory theory and the simulation theory are in turn attempting to answer the traditional epistemological question about how we come to have knowledge about other minds.[185] Here one should note that the riddle of other minds is based on an assumed dualist separation of inner and outer and, also, hidden mental states and overt behavior. In other words, if minds fall squarely on the internal side of the divide and behavior squarely on the external side, then other minds are moved out of sight - and out of reach. And, it then becomes puzzling how we can understand others as even having minds given that we only have access to their observable behavior. And what becomes even more befuddling is how we can know what goes on inside those inaccessible minds of others.[186]

---

[184] For more on theory theory and the response to simulationist critics see Carruthers & Smith (1996), in particular Carruthers (1996).

[185] In regard to the focus on *knowledge* of other minds and the 'other minds problem,' it is noteworthy that Goldman came to the issue exactly from the perspective of epistemology, which was and to a large extent still seems to be his primary philosophical concern.

[186] In regard to the 'problem of other minds,' one might see Goldman's proposal as rooted in Mill's argument from analogy, i.e. the idea that we recognize others' behaviors as being like ours and hence infer mental states similar to ours. This is not to say that the idea and theory of simulation does not contribute something radically new, but that it in some ways shares the structure of the argument for analogy.

But why assume this picture of the mind? And why blindly assume that this sort of 'mindreading' or 'mental state attribution' is the essential means by which we understand others? Why assume that our basic social cognitive abilities are based on attribution (via introspection) of inner mental states to oneself on the one hand and disengaged – one-way-mirror if you like - action observation of 3rd persons on the other? The complex social skills displayed by little children and present in real human interactions are difficult to explain via such abstract and inherently individualistic processes of attribution are the primary social cognitive ability.[187] Rather, disengaged 'mindreading' seems to be a much later and secondary social ability, and accordingly appears unlikely as the foundation for seeing others as minded agents and as having rather specific things on their minds. This makes particular sense if one lets go of the picture of the mind as necessarily hidden and opposed to 'mere behavior,' and explores the possibility that we often can 'see' not only that others are minded but also to a large extent what is on their minds. The idea is that many mental goings-on are to be seen *in* overt engagements rather than always *behind* these – and thus that much social cognition can be done without simulated access or theoretical inferences.

Hence, I share the motivation that has led Merleau-Ponty and many contemporary philosophers to vehemently argue that most actual processes of social cognition and engagement simply do not seem to be based on these sorts of 3rd person attributions – be they implicit or explicit.[188] Such a distinct 'attribution process' seem obsolete in many successful interactions and dealings with the intentions, goals and concerns of others. Wittgenstein and more recently Dan Zahavi and Shaun Gallagher have suggested that mental life often is directly perceivable in behavior and thus that in most normal cases of social cognition neither inference nor simulation is needed.[189] The idea of 'direct perception' might be controversial in its own right, in that it might make problematic suggestions as to the simplicity of the cognitive processes involved perceptual and social

[187] Later in this chapter, in section 4.2.6.A, I discuss children's understanding of other minds and the experimental protocols of false belief tests that are used to assess their social cognitive abilities. A discussion of social behavior in infants and non-human primates, which seem incapable of more abstract mental state attributions, is beyond my present scope. My point is merely that it cannot be assumed *a priori* that mental state attribution is the primary or pivotal social cognitive skill.

[188] Merleau-Ponty (1962). For contemporary echoes of the view see e.g. Dreyfus & Dreyfus (1999).

[189] Gallagher (2007) and Zahavi (2008).

understanding. But that requires a separate discussion.[190] What is important for me here is not so much claims about the directness of the perception – but rather the idea that mental goings-on often can be perceived. In other words there is a question as to whether the 'objects' of social cognition, for example what other people are seeing, intending or thinking, often is available for perception rather than hidden. Gallagher also stresses the importance of the 2nd person relation and of what he with Trevarthen calls 'primary inter-subjectivity.'[191] Paradoxically, social interactions are often ignored in the discussion of how we 'mindread.' The attention is normally narrowly focused on 1st person and 3rd person understanding and prediction of mental states, very much as if these were little wooden blocks or pages of print lying around in the head. Dan Zahavi is, by virtue of his phenomenologically inspired theories of the self, inter-subjectivity and empathy, one of the most vocal and insightful critics of this idea of mental content and of the neat distinction between inner mental states and external overt behavior.[192] I will later say more about the neglect of the 2nd person and the problem of seeing mental states as some kind of internal 'content.' Here, I just want to underline that even though Goldman opposes the idea that we understand other minds by way of *theorizing,* he still focuses on 3rd person 'mindreading' and sees this as a stepwise process of, firstly, mental state recognition and then an attribution of this state to someone other than oneself.[193] As recently as in a 2008 book chapter Goldman writes:

> By 'mindreading' I mean the attribution of a mental state to self or other. In other words, to mind-read is to form a judgment, belief, or representation that a designated person occupies or undergoes (in the past, present, or future) a specified mental state or experience. (Goldman, 2008)

Thus, there is both a 'judgment' about the nature of a mental state and then an 'attribution' or a 'designation' of it to a person. His proposal differs from that of the theory theorists in that he does not think we need a theory or rules of inference to come to

---

[190] I shall return to the issue of direct perception in the final chapter. One difficulty in regards to this issue is that the discussion of the possibility of social or 'mental state' perception is interrelated with the debate over whether perception as such is based on an inferential or a direct process like the one suggested by Gibson and others. My view is that the inferential-direct dichotomy might have some serious problems. But I see al this as a somewhat separate to the issue of whether elements of our mental life is given in our overt interactions. In other words my interest here is in whether there is something mental 'directly' there to perceive rather than whether perception as such is a direct process.

[191] Gallagher (2001) and Gallagher (2004).

[192] Zahavi (2001) and Zahavi (2008).

[193] In the case of one's own intentional and emotional states the suggestion seems to be that our knowledge is based on some process of introspection. (Goldman 2006)

a conclusion about the mental states of others. Rather, we use our own mental and cognitive apparatus plus our knowledge of the uniqueness of others' situation to make low- and high-level simulations of 'putting oneself in their shoes.'[194]

Goldman thinks that we know our own minds and mental states primarily by introspection and then given this idea his proposal is basically that we by simulating others to some extent can get that sort of pretend privileged inside access to their minds as well. More concretely, he explains that we with the knowledge of others' situation make 'pretend beliefs' and 'pretend desires' that we use to make 'offline' simulations of their thoughts, feelings and decisions.[195] I propose that one might characterize his theory as a instance of what I more broadly label a '3rd via 1st person' social cognition. The idea being that one can understand the 3rd person by pretending to be them and 'reading' of their mental state from my own 1st personal simulation and then attributing it to the other.[196] Given this sort of theory of mindreading as simulation, positing an 'as if I was you' function, we can see how mirror neurons become interesting for Goldman in that they could possibly render empirical evidence for low-level simulation of others' actions, i.e. a neurological process of 'as if' I performed the action. However, the step-like process of first mental state *recognition* and then some process of attribution of this state to the other also warn of possible tensions between this conceptual and procedurally staged account of mindreading and the findings of automatic sensorimotor integrations at a low and sub-personal level. Goldman writes:

> To attribute a mental state, the judgment must deploy a mental concept or category. Thus, if 'empathize' simply means 'echo the emotional state of another,' empathizing isn't sufficient for mindreading. A person who merely echoes another's emotional state may not represent the second person at all, and may not represent her as undergoing that emotional state (a species of mental state). (Goldman 2008)

Goldman thus explicitly emphasizes that emotional or action mirroring does not amount

[194] See Goldman (2008).

[195] Alvin Goldman recently expressed that he personally has moved away from utilizing the term "offline" but, since I rely on the early version of his simulation theory, I retain the term because of its early influence on the mirror neuron debate.

[196] As we shall see, I at some level sympathize with the idea that we use our own cognitive apparatus to understand the others' perspective in relation to their specific situation – and 'how it is to be in their shoes'. But I think that the social knowledge gained from this process exactly pertains to an understanding of the other's relation to and perspective on the world - not by some miraculous jump into their heads or through 1st person experience. Additionally I think a different notion of what counts as a 'cognitive apparatus' is needed as well as a more restricted notion on what can be 'simulated'.

to conceptual mental state attribution,[197] and that additional processes are needed for, on the one hand, conceptually recognizing or labeling the mirror state/action and then for attributing the state or action intention to the specific other person in question. This means that if mirroring is truly an agent-neutral process then an extra process of attribution and representation of the other as other is needed. Already in the early 'strictly congruent' based conception of the role a mirror mechanism in social cognition, we saw that an extra-mirroring 'who-system' and an action inhibition process were suggested to deal with the question of how a symmetric process of mirroring can yield respectively action and understanding of others. I have obviously challenged this conception of mirror neurons as constituting a symmetric, encapsulated agent-neutral mechanism in the first place. But the point in regard to Goldman is that he in his recent work is clearly very aware that there are serious differences between his notion of mindreading and the mirror neuron findings. He makes a distinction between low-level 'enacted' simulation and higher-level 'pretense' simulation, and suggests that only the former might be instantiated by mirror mechanisms. Be this as it may, he originally embarked on the affair with mirror neurons in the hope of finding neurological and empirical backing for simulative mindreading, and it is indeed the higher-level pretense version of simulation that is presented in the article with Gallese, which went on to become ever so influential. Interestingly, it also seems to have been the link to higher-level mindreading that attracted Gallese to Goldman's theory.

4.2.2. Simulation as gateway from mirror neurons to higher social cognition

Gallese was interested in building a theory of human personal level social cognition from his low-level empirical evidence of action sensitive visuomotor neurons in premotor areas of macaque monkeys. He had in prior articles (di Pellegrino et al 1992, Gallese et al. 1996, Rizzolatti et al. 1996) with various colleagues already hypothesized that mirror neurons provide an observation-execution matching system and further argued that this system could play a central role in various social cognitive competencies. Goldman's simulation theory of 'mindreading' could help bridge the palpable conceptual gab

[197] It is interesting that Goldman in this quote equates empathy simply with emotional mirroring or contagion. Thus he does not see the representations of the other as other as important for the ability to empathize. Gallese at times seems to show a similar view (Gallese 2008). Zahavi has an effective criticism of this view of empathy. See Zahavi (2001).

between local single neuron recordings in the infamously 'beasty' macaques and personal-level social abilities in humans. The key uniting feature was the idea of understanding of others via a simulation process; a kind of 'as if' pretending to be or act like the other resulting in understanding due to one's own broader implicit knowledge of being in such a situation. Based on the finding that (some) mirror neurons responded similarly to one's own and others actions, it had already been proposed that they instantiated an agent-neutral or maybe even a 1st person representation of the 3rd person observed action. Mirror neurons were seen as little modules mirroring, retrieving – simulating - elements of the actions of others, and further as somehow accessing broader 1st person experiential and implicit knowledge related to such actions when executed by oneself. The 'mirror mechanism' could now borrow the framework of Goldman's hypothesized *process* of simulative mindreading, and make the move from *action retrieval* to *social understanding* of other minds. The more specific proposal, as we shall see, was that the mirror neuron system during perception is used 'off-line' in the sense that the neurons fire without an overt motor consequence. The functional role of mirror neurons in perception is thereby related to understanding rather than action per se, which supports the idea that mirror neurons in their 'receptive mode' might also have a radically different and more 'epistemological' cognitive function than other more 'pragmatic' neighboring motor neurons. And, further that they might instantiate a mirroring mechanism that can be conceptualized in relative isolation from the rest of the F5 neurons and more generally the adjoining premotor-parietal circuits.

4.2.3. The joint Gallese-Goldman theory – more virtual than real

With this background in mind – of strong motivations but also theoretical tensions - it is hopefully easier to understand why the explicit proposal of a simulation-mirror theory is extremely cautious to the point of non-existence, and why the marriage between the two theories is happening between the lines of the 1998 article rather than in the actual text. To bring together the Simulation Theory of mindreading and the finding of Mirror neurons Gallese and Goldman start by summarizing their hypothesis as follows:

> …we propose that humans' mind reading abilities rely on the capacity to adopt a simulation routine. This capacity might have evolved from an execution/observation matching system whose neural correlate is represented by a class of neurons recently

> discovered in the macaque monkey premotor cortex: mirror neurons (MNs)(Gallese & Goldman, 1998, p. 493).

Given this formulation we can distill several distinct sub-components or hypotheses: 1) the idea of mind reading as simulation, 2) that mirror neurons are the neural correlate of an execution-observation matching mechanism, and 3) that this mirror matching mechanism *might be the evolutionary precursor* of some more advanced mindreading simulation process. Thus, there is the simulation theory and the mirror mechanism, and a possible evolutionary link between the two. No offence to findings of evolutionary connections, but this is a rather weak link as mirror neurons in their present evolutionary state – in monkeys as well as humans – already were proposed to support a mechanism important for action understanding and social cognition. Later in the article they use a slightly stronger but still vague formulation; i.e. that mirror neurons "underlie the process of mind-reading or serve as precursors to such a process". With this level of cautiousness from the founders of a mirror neuron based mind-reading theory one might ask oneself if there was ever a fully fleshed-out theory. As noted Goldman in his later writings seems to suggest that only some rudimentary forms or partial elements of simulative mindreading are based on mirror neuron activity, and he uses emotion mirroring rather than action mirroring as the most clear example of such low-level 'enacted simulation' (Goldman, 2009). As we shall see in section 4.3 of this chapter Gallese moves away from the terminology of mindreading altogether and reinvents his mirror theory of action understanding as 'embodied simulation' based on 'pre-wired chains' of mirror neurons. Thus, Gallese and Goldman in spite of the impression one gets from the title and many frequently used slogans never really seem to make a forceful argument for full-blown mirror neuron based mind-reading theory. However, they do indisputably claim that 1) mirror neurons instantiate a observation-execution matching mechanism and 2) that this mechanism creates some level of simulation of the others action and the experiences that are connected to that action, and 3) that this action simulation provides an special understanding of this action and possibly the intention behind it. But most importantly, simply by attempting to link the two theories they did arguably in some way link them and change each of their interpretations in not so subtle ways. I argue that the caricature model of the mirroring mechanism, as ubiquitously producing context-independent, agent-neutral covert simulations of observed actions, was transformed from its tacit life in the mirror metaphor to being the received view by way of the simulation connection. As we shall see in the next chapter, it is this version of the mirror mechanism that typically is the target of discussions, and thus the actual findings and subtleties are generally ignored for the sake of a caricature theory – which paradoxically neuroscientists did not really explicitly defend in the first place. However, I think that Gallese by co-writing this enormously influential theory with Goldman might be guilty of leading everybody – maybe even himself and his colleagues – further a stray in regard to the nature of mirror neurons and motor resonance during action perception.

4.2.4. A simulation-friendly caricature of the mirror findings

Gallese and Goldman emphasize how the activity of neurons of area F5, in which mirror neurons where first found, is related to goal-directed motor acts rather than individual movements. They find this important as it "allows one to interpret the role of the motor system as not just in terms of the control of movement (like joint torques, etc), but rather as a possible candidate for the instantiation of mental states such as purpose or intention". They follow Rizzolatti's idea of conceptualizing F5 neurons "as a 'motor vocabulary' of actions related to prehension"[198]. In regard to the relation between the visual and the motor action properties within individual mirror neurons, they summarize the findings as follows:

> In over 90 percent of MNs a clear correlation between the most effective observed action and their motor response was observed. In many neurons this correlation was strict both in terms of the general goal of the action (e.g. grasping) and in terms of the way in which it was executed (e.g. precision grip). On the basis of their functional properties, here summarized, MNs appear to form a cortical system that matches observation and execution of motor actions. What could be the possible functional role of this matching system? (Gallese & Goldman, 1998, p.495).

Gallese and Goldman here dramatically ignore the complexity of the findings and progress from a description of an observation-execution overlap in some mirror neurons to the conclusion that these neurons instantiate a cortical observation-execution matching system. At first glance this move might seem rather innocent. However, as we saw when looking closely at the early findings the important questions about so-called broadly and non-congruent neurons are here being swept under the rug. The coauthors simply write

[198] Gallese & Goldman, 1998, p. 495.

that many of the mirror neurons are strictly congruent, i.e. responsive to the observation and execution of the same action.[199] Thus, again, if the totality of mirror neurons are supposed to form an observation-action mirror matching system, then it seems problematic that Gallese and Goldman fail to even ask why only around 1/3 of them are strictly congruent.[200]

In regard to the evidence of a homologue human mirror system, as mentioned in the previous chapter, things take on a whole new dimension of problems. Substantial evidence has been found of various kinds of 'motor resonance,' i.e. motor activity induced by action observation in human subjects over the years. Several such studies have been taken to provide evidence for a human 'mirror neuron system' similar to that found in monkeys. However, it is pivotal when interpreting the results to remember the constraints of the indirect methods of functional imaging and Transcortical Magnetic Stimulation (TMS), and make explicit the many assumptions one needs to make in order to interpret the found pre-motor and parietal activity as evidence for a direct observation-execution matching system. Needless to say, such caution is not too common in the mirror matching literature. Often it is rather the other way around; the conclusions get more generalized rather than more cautious when it comes to the theoretically scaffolded human data. Thus, Gallese and Goldman, after a reference to Fadiga and his coworkers' findings of motor resonance in humans,[201] feel justified in making the following statements:

> This study provided for the first time evidence that humans have a mirror system similar to that in monkeys. Every time we are looking at someone performing an action, the same motor circuits that are recruited when we ourselves perform that action are concurrently activated. (Gallese & Goldman 1998, p.495)

The first claim about the evidence of a similar mirror neuron system in humans might be somewhat fast on the basis of the simple evidence of strong motor resonance to action observation. However, it is almost innocent compared to the second more specific claim

---

[199] Given the earlier reports of di Pelligrino et al (1996) this number seems to be at most 1/4 of the subgroup of F5 neurons that have both visual and motor properties and approximately 1/3 of the mirror neurons, i.e. all the neurons responsive to some kind of action observation.

[200] We shall see in the last section of this chapter that Gallese has changed his view somewhat and is now with his focus on 'chains of action' to some degree addressing the possible 'not direct matching' functions of broadly congruent mirror neurons. However, the 'finding' of a direct matching system is still based on a series of assumptions and simplifications about strictly congruent mirror neurons.

[201] Fadiga, L. et al. (1995).

and its gigantic generalizations. This second sentence is immensely important because it pretty much distils the caricature mirror matching model – except they present it as a proven fact rather than the poorly supported hypothetical proposal it is. Let me pause and dissect the three key implications made in this sentence:

A) That ALL action observation in EVERY context leads to mirroring in the observer.
B) That the mirroring process uses the SAME motor circuits as action performance.
C) That the mirroring happens CONCURRENTLY with the observed action.

First, the generalization from the simple finding of motor resonance in response to observation of specific hand-object interactions to a general claim of action perception is of course empirically unwarranted and also contradicted by basic findings. Both the direct evidence from monkeys and the human study of Fadiga and colleagues referred to here present clear evidence to the opposite: Namely that not all actions are effective, and that presumably intentional performed “aimless movements in the air” and mimicked actions in the case of monkeys are not effective in creating significant motor resonance. I shall argue later that even if the claim only regards concrete goal-directed actions, it is problematic to suggest that the motor resonance only depends on the perceptual presence of an action independently of the wider perceptual, experiential, motivational and practical context of the observer and the interactional relation between the two subjects.[202] Turning to the issue of whether the ‘same motor circuits’ are active during observation and execution, it is certainly clear that it is not all the same circuits, since there is not any overt action in the observation case. The question is therefore rather which parts of the motor circuits for the execution of a given action are activated during action observation? What is the nature of the overlap and how can the functional role of these processes be understood in regard to normal action and action perception respectively? Also, is the motor activity during action perception the same independently of the task and the context? I have tried to show in the previous chapter that the idea of mirroring as a ubiquitous observation-execution matching process was tacitly imported

[202] There is also the issue of what constitutes action observation and what an effective form of action presentation is in monkeys and humans respectively. Comparing various studies it seems that many monkey MNs respond only to action produced by live full person views of action performance, whereas human motor resonance has been found also in response to video clips and observations of isolated body-part movements. See here for example Gallese et al (1996) and Brass et al. (2000). These differences might simply be due to different groups of neurons studies as indicated by the findings of Nelissen et al. (2005). The point is that it is an open empirical question rather than a settled fact.

already with the mirror terminology and categorization choices made in regards to single mirror neurons. Here it looks as if Gallese and Goldman simply *replace* the empirical findings with the mirroring metaphor. They seem to suggest that (some) motor activity during action observation and execution is nearly identical – except the neural actual directly responsible for overt execution.[203] This is exactly the idea of motor simulation, i.e. 'covert action as action minus execution' that Marc Jeannerod and Pierre Jacob seem to be defending when arguing against the importance of mirroring in social cognition, and which I will discuss in detail in the next chapter. I argue that it is wrong to think of all mirroring, sensorimotor activations, or perceptually induced motor responses as such fully specified covert actions. My proposal is that given the social task and the context the action perception could result in different degrees of action specification.[204]

In regard to the claim of the concurrent activation, Gallese and Goldman do not provide any empirical evidence regarding temporal features of externally triggered mirror neuron activity, and neither do they discuss the possible implication of concurrent or delayed activity in the observer. As discussed in Chapter 3, many mirror neurons were found to be modulated in a way time-locked with the ongoing action; thus there is evidence for some kind of concurrency of mirror neuron responses and the *performance* of the overt action but not a concurrency of neurological activity in the agent and the observer. It might be the latter neurological mirroring between agent and observer that is taken to support Gallese and Goldman's suggestion that the mirroring process can give access to the hidden mental states of the other – not just the overtly observed action. They might therefore want to establish that mirroring is simulative in the sense that the activation matches the neurological – and presumably thereby the mental - state of the other. In total, the process is suggested to tell the observer something about how it would be to be the one performing the action observed. That is, it would give a simulated 1st personal access to the 3rd person. Pierre Jacob has devoted a lot of energy to refuting exactly this idea of neurological resonance or what he calls the 'tuning fork model' of mirroring. For

[203] The neural activity actually directly responsible for overt execution presumably would be limited to primary motor cortex and cortico-spinal connections to motor neurons, thus the implication is the strong claim that there should be a nearly perfect overlap between the activities of premotor and supplementary motor areas during observation and execution, which of course is not the case.

[204] This might also explain the at times contradictory findings regarding the degree of muscle facilitation and the level of accuracy and details of motor resonance in humans and of mirror neuron activity in monkeys. See Gallese et al. 1996 and the discussion of the sub-threshold theory in Chapter 3.

Jacob this element of the mirroring theory is pivotal, because he supports the mindreading idea that the understanding of other minds must rely on some kind of inference to hidden mental states. Accordingly he uses his argument against the ‘tuning fork model’ of mirroring to conclude that mirror neurons cannot be essential to social cognition or the understanding of prior intentions.[205] I agree that neither neurological concurrency nor matching is very plausible. But, since I do not understand social cognition as primarily grounded in processes of ‘3 via 1st person’ simulative attribution of otherwise hidden mental states, I do not think that either the lack of correspondence or of temporal concurrency of covert activations preclude mirror neurons from playing an important role in social cognition. Further, I am not sure that Gallese or Rizzolatti or others from the Parma group ever meant to claim a neuronal synchronicity. However, as we see here in the Gallese and Goldman article they do unfortunately sluggishly or by some sort of mirror-mesmerization suggest both the idea of concurrency and identity.
To summarize, the overall impression one gets from the quoted mirroring claim is that Gallese and Goldman simply do not find the existence of a ubiquitous, agent-neutral mirror matching process in either monkeys or humans seriously questionable. They almost take the existence of a mirror neuron system as presented by the caricature mirroring model as a given, just as they seem to take it as a given that understanding of other minds can be equated with access to hidden mental states.

4.2.5. Simulative mindreading as a model for mirror neurons

At first glance a direct execution-observation matching mechanism seems to provide an ideal neurological basis for imitative abilities, but as mentioned the problem is that macaque monkeys in which mirror neurons were discovered show very weak imitative abilities.[206] On these grounds Gallese and Goldman suggest that the primary function of the mirror system might be one of ‘mindreading,’ of “representing specific mental states of others.” They hypothesize that “MNs are part of – albeit perhaps a rudimentary part of

[205] Jacob, P. (2009), ‘The Tuning-Fork Model of Human Social Cognition: A Critique’, Consciousness and Cognition, 18, pp. 229–43. As we shall also see in the next chapter Csibra suggests that simulation is anticipatory and that we actually do not simulate the actions we see but emulate the actions we infer. Thus, he challenges the idea that ‘mirroring’ is actually mirror matching, and thereby what it can tell us about the others’ intentions that we did not know prior to the simulation process.
[206] See discussion of imitation in Chapter 3, section 3.3.1 and of social abilities in humans and monkeys section 3.5.1.

– the folk psychologizing mechanism." (p.495) Thus, Gallese and Goldman make the jump from mirror neurons and the idea of an observation-execution matching mechanism to a discussion of mind-reading and how we come to attribute mental states to others. More precisely, they jump right into the old theoretical dispute of whether mindreading is accomplished by way of simulation or a more rational/theoretical inferential process as the so-called 'theory theory' account suggests. I want to discuss Gallese and Goldman's presentation of the simulation theory of mindreading in detail because some of the ambiguities and problematic ideas of this theory are imported into their theory of simulation at the level of mirror neurons – and from there they get adopted by other researchers in the field. In other words, even though their explicit claim is only that mirror neurons might be some evolutionary precursor to full-blown mindreading, it looks like both procedural and theoretical elements of the simulation theory of mindreading get implicitly imported into the understanding of the mirroring mechanism and how it could underlie action understanding. Thus, in what follows below I spell out those elements.

A. The pretence model of simulation

Gallese and Goldman present the simulation theory in contradistinction to the theory theory and they give two examples of cases of 'folk psychologizing' where a simulation account seems more plausible than some kind of lawful reasoning process. The interesting thing is that what is taken to be 'simulated' in each of the two examples is radically different, and hence rather than illuminate the idea of simulation the examples raise further questions of how actually to understand this process.[207] In the first example they use a 'mind-reading' task where one is supposed to judge two travelers' emotional reactions to two variants of a similar situation. The vignette from a study by Kahneman and Tversky (1972) goes like this:

> …two travellers who shared the same limousine en route to the airport and were caught in a traffic jam. Their planes were scheduled to depart at the same time, but they arrived 30 minutes late. Mr. A was told that his flight left on time; Mr. B was told that his flight was delayed and just left five minutes ago. Who was more upset?

[207] Stephen Stich and Shaun Nichols have repeatedly argued that the notion of simulation is used so 'heterogeously' that it has become 'quite useless' and thus needs to be 'retired.' See Stich and Nichols (1997) and Nichols and Stich (2003). The examples used by Gallese and Goldman certainly suggest that they use the idea of simulation ambiguously. However, I do not strictly insist on a streamlined or well-defined vocabulary; I simply want to question the plausibility of simulation under certain interpretations.

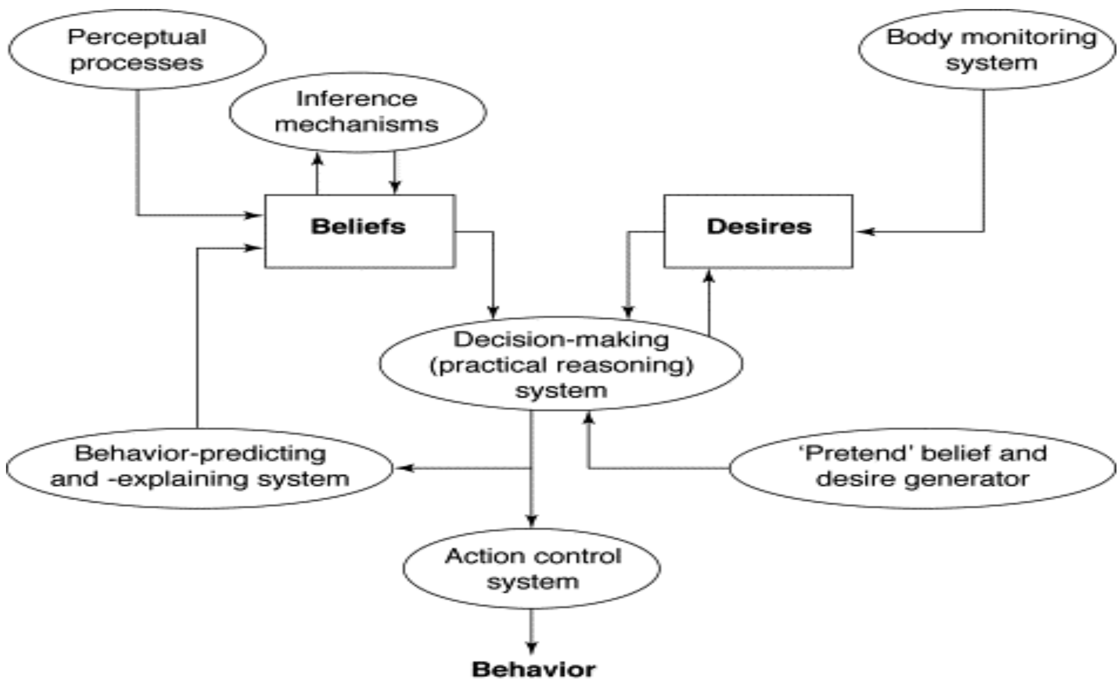
> Ninety-six percent of the experimental subjects said that Mr. B was more upset. How did they arrive at this answer? (Gallese & Goldman, 1998, p. 497)

Gallese and Goldman answer that "according to TT [theory theory] there must be some psychological law they exploited to infer the travelers' relative upsetness," and find this method improbable as the underlying cognitive process. They instead claim that on the simulation theory account they support the story would look more plausible because in order to accomplish this task a "subject would have put himself in each of the imaginary travelers' 'shoes' and imagined how he would have felt in their place." In this case 'simulation' pretty much means to imagine oneself in the situation of the other and then to imagine *one's own (emotional) reaction in that situation* and then attribute it to the other. Thus on a neurological level I presume it might involve a visualization of oneself as experiencing two different *external scenarios*, which help form the judgment of whether one would have felt differently in the two situations. What is simulated is an emotional response given specific prior information in a situation. In the second example the task is one of *prediction of decisions*, and it is even more reliant on the language of 'pretense simulation' that Goldman normally uses:

> To predict White's next move in a chess match ST suggests that you try to simulate White's thought processes and arrive at a decision, which you then attribute to him. First you create in yourself pretend desires, preferences, and beliefs of the sort you take White to have; for example, preferences among chess strategies. These pretend preferences and beliefs are fed into your own decision-making mechanism, which outputs a (pretend) decision (see Fig. 2). Instead of acting on that decision, it is taken 'off-line' and used to predict White's decision. (Gallese & Goldman, 1998, p.497)

The success of our prediction judgment is thought to depend on a *similarity* between the thought processes that we go through by way of simulation and the actual thought process of White. The authors describe how this thought process simulation could take place as an attempt to 'pretend' to have similar motivations and knowledge and then use one's own 'decision-making mechanism' 'off line' to predict action rather than performing it. (See Figure 4.1) The prediction is thus not rationally inferred but rather 'read off' our own simulated decision process and then attributed to the other. Compared to the first example, the idea of using one's own cognitive system to understand the other is now taken to another level. We are now not just trying to imagine being in the other's shoes or situation, but rather to be in their flesh and thought. Importantly, one is asked to first think of the other as different from oneself, in the second place pretend to be the other, and then to simulate what one would do if one is the other. They write that simulation theory "…depicts mindreading as incorporating an attempt to replicate, mimic, or impersonate the mental life of the target agent." (p. 497) With this model Gallese and Goldman make the interesting assumption that our cognitive system can - and regularly does - operate independently of our actual flesh and thought; that we have the same kind of imaginary process and imaginary powers in the realm of 'beliefs' and 'desires' as we do in regards to external events or situations, and thus that we can pretend to be in someone's motivational and epistemic situation and then treat that as content 'inputs' for rather impersonal modular cognitive mechanisms.

Figure 4A. Simulation 'Boxology' reproduced from Gallese & Goldman (1998)



The 'boxology' model described in the quote and reproduced above is implausible on many levels and it probably should not be taken too literally. However, the elements of this model seem to sneak into the mirror mechanism theory, so for this reason alone it must be taken seriously. The central idea is that we have a cognitive input-output mechanism – here the 'decision-making mechanism' the function of which is independent of its precise contents and their origin or destination. Thus, the mechanism can be taken 'off-line', and seemingly under our voluntary control we can exchange its normal inputs with pretense input to get a simulated output that can be used for epistemological purposes rather than as determining overt action choices. The model thus embodies many of the – in my eyes - problematic and misleading traditional assumptions of mechanistic information processing models of cognition. For now I will just underline the implicit assumptions that 1) the mechanism has an exact function that remains a constant independently of its concrete contents and that 2) the nature of its inputs as well as the use of its outputs are under our full voluntary control. Hence, there is a strong dichotomy of form-content but also an implicit idea of a neatly defined and circumscribed set of inputs and outputs that can be 'handled' by the mind like 4-dimensional objects between our hands. I do think that we have abilities of imagination and pretense and also that we can use these to 'simulate' and further understand hypothetical scenarios, but on my view such pretense involves primarily a reorientation of our ongoing sensorimotor based cognitive engagement, and not an empty machinery (form) waiting to be fed real or pretense beliefs and desires (content). In other words, I see pretending as a dynamic cognitive process in which form and content are interdependent, and which can be voluntarily 'directed' but not totally controlled. Further, 'inputs' are not well-defined entities and the process not 'all or none' kind of phenomena. On my theory we can construe and manipulate hypothetical situations by imagining the situation and/or our knowledge in regards to it to be different. But I think of it in terms of an imaginary 'affordance space,' such that the pretense situation (just as the online perceptual situations) depends not just on external or objective 'inputs' but also on how our cognitive processes are shaped by pragmatic and sensorimotor experiences. In other words, we can pretend to be in a situation with different affordances –different action invitations and possibilities – but not that the same situation has affordances that we would not ourselves perceive or understand in any situation. Thus, I find it implausible that we can "simulate White's thought process" and thereby predict his action as different from what we ourselves would have done in his shoes as Gallese and Goldman propose. The assumption is exactly that we can simply replace our beliefs and desires with the beliefs and desires we take White to have and then we can go through the motions of an action decision process from his perspective. But the chess example used clearly demonstrates the absurdity of this idea: if I could simulate the decision process of a chess master I would be a chess master! The truth is that the positions on the chessboard afford wildly different moves to a novice like me than it does to the chess master. Therefore I cannot predict the chess master's moves but only that she will probably do something that I would not expect. It is not just that I have different beliefs in my belief box and desires in my desire box – or that I would be able to pretend to have other beliefs/desires at will. Further, it is assumed that I can simply add these beliefs and desires to some kind of objective non-personal perception of the chessboard. The idea of a simulative anticipation of the concrete action choices of others does not make sense unless we from our own knowledge and skill-set can understand how the goal is or at least could be afforded by the situation. In other words, unless we can see that a goal is attainable in a concrete or hypothetical situation, how would we be able to predict the action? If we do not ourselves have the right knowledge or skill set I do not see how we can simulate or anticipate concrete actions. Rather we can then just have a diffuse expectation that other's actions will happen in accordance with their goals, knowledge and strategic preferences. The point that I am trying to make is firstly that our particular knowledge and skills in many ways *are* our cognitive machinery – and not something that is fed into it. In other words, form and content might not here be separable in the neat way Gallese and Goldman suggest. And, secondly, I think that we cannot simply jump into the thought process of the other and simulate their action choice based on belief and desires without a sense of their pragmatic skill-set and thereby their affordance relation to the context in which the action is chosen.[208]

My objection to the simulation theory is not that we never understand others by way of some sort of simulation, but rather that I do not think that the modularization of our cognitive processes into modular mechanisms and separate content inputs and outputs, perception and decision processes is empirically plausible. I do think we to some degree can 'simulate' others' perspective and situation as in the first example mentioned. Further, it also appears that we can use the prior knowledge about the others' abilities, knowledge and perspective to anticipate how their choices might be different than ours –

[208] In the *Philosophical Investigation*s (1953) Wittgenstein famously writes "If a lion could talk, we could not understand him." (Pt II, p. 223 of the 1968 English edition) I understand this in parallel to the point I am making here. The idea being that communication - and in this case understanding and predicting the decisions of others - is not simply about 'information' but about sharing purposes, practices and skills. Thus the parallel here is 'if a chess novice could see the beliefs and desires of a chess master, he still could not understand her decisions.' One might say that knowing White's personal strategies is exactly to know something about his affordance perception and might help us predict his move – but would this prediction involve a simulation of his inner thought process or rather an attempt to see the situation 'through his eyes'? The latter involves an attempt to change one's own outward attention process in the situation to see other possibilities rather than a change in the inward response to the same features of the experience.

but I think that this latter process is mostly negative, in the sense that we might rule out some of our affordances as not relevant for them, or imagine them as having affordances that are exactly not available to us. In this way, one might anticipate others to make moves not afforded us – but I do not think this can be construed as a positive simulation process. The 'as if' of simulation, the idea of actually using one's own cognitive apparatus to understand and further one's insight into the perspective of others, exactly relies on a real similarity between our cognitive abilities and not a pretended similarity.[209] I suggest that that our cognitive processes cannot easily be divided into form and content, machinery and beliefs/desire representations and perception and decision mechanisms. If this suggestion is right and if these processes are to some degree rooted in sensorimotor representations, which at the same time inform form and content, i.e. the process and the information they carry, then simulation processes are very different from what Gallese and Goldman suggest here.

B. The idea of retrodictive pretense simulation

After having set up their framework of a positive pretense simulation process by way of the examples critically analyzed above, Gallese and Goldman move to the idea of a "retrodictive use of simulation." As presented it seems to be a truly far-fetched idea, but it is later hypothesized as an important function of mirror neurons and as essential to the claim that mirroring can yield intention understanding. They write:

> The attributor starts with the question, 'What goal did the target have that led him to perform action m?' He conjectures that it was goal g, and tries out this conjecture by pretending to have g as well as certain beliefs about the effectiveness or ineffectiveness of the action m vis-à-vis goal g. This simulation leads him to form a (pretend) decision to do m. He therefore uses this result to conclude that the target did indeed have goal g. In this fashion, the attributor ultimately makes a 'backward' inference from the observed action to a hypothesized goal state. (p. 497)

Well, one way of paraphrasing the proposal is that we can test a hypothesis of the goal of an observed action (that we have 'conjectured' by inexplicable means!) by a simulation process where we check if we would have done the observed action given the hypothesized goal. It is thus rather misleading to suggest that the simulation process

[209] In cases of less radical cognitive differences, one might be able to 'simulate' the preferences or information levels of others by way of a different attentional emphasis on one's own perceived affordances. This might in some ways be similar to Goldman's idea that we can simulate the beliefs and desires of others but within an affordance space framework rather than a cognitive 'decision making' module.

produces a 'retrodiction,' i.e. "a 'backward' inference from the observed action to a hypothesized goal state. Rather than providing a retrodiction or an inference to the goal, this presumably time-consuming process would just confirm the plausibility of a hypothesis we already had. Given the temporal constraints of real time interactions this 'trial and error' simulation seems rather implausible as a psychological process used in our understanding of the intentions of others. All this combined with the fact that it is completely left in the dark how we come to 'conjecture' the hypothetical goal in the first place. This is to say that the core mindreading question of how we would get into other people's hidden mental states is not actually explained by the simulation process as described here – only how we could simulatively *test* possible goal candidates. As we shall see in the next chapter this is exactly what Csibra's criticism latches on. He reasons that if the action simulation indeed goes from 'goal to action' as Gallese and Goldman suggest here, then the goal attribution cannot be a product of the action simulation put rather the input that gets the simulation process started in the first place. Csibra hypothesizes that there is a separate extra-motor or 'purely perceptual' goal attribution mechanism. This purely perceptual mechanism in my eyes seems to be yet another mysterious black box allocation of the problem rather than actually a step towards providing an explanation. However, given a narrow idea of motor cognition and the role of such processes in goal understanding, Csibra argues that mirror neurons are irrelevant to the goal attribution process and thus to the social cognitive process of 'reading' the mental states of others that are hidden behind the overt action. I will discuss Csibra's argument and its problematic assumptions in detail in Chapter 5. Here, I want to question whether Gallese and Goldman – in spite of the rhetoric - truly meant to propose this sort of 'goal-to-action' simulation process as a model for the actual mirror neuron activity.[210]

---

[210] This discussion should hopefully also lead to an understanding of a broader problem with the idea of '1st via 3 person' mindreading. Namely, if what I use to understand the other is not simply my innate or 'naked' mental *apparatus* plus various perceptual and informational *content,* then how do we make sense of the simulation process and of it as yielding access to the other's first person perspective or hidden mental states? I suggest that we rather understand the behavior and the information about the beliefs and desires of others in a relational way – and via our own relational engagements and experiences with the world and others in it. On this story, I can understand others' thought processes via shared experience based relations to and understanding of the relational sensorimotor and meaningful engagements. This relational understanding is what I root in an affordance space understanding and imagination of hypothetical affordance spaces not presently perceived. The point is that I do not understand the other's hidden mental life via an introspection of my own hidden experience during a simulated behavior like the one observed by the other now. Rather something relational is at the core not only of my understanding (or failure to understand) the mental life of the other but even at the core of the possibility of my own mental life.

C. Mirroring as retrodictive simulation

Gallese and Goldman do indeed move from the formulation of the mechanics of the 'retrodictive' simulation process to a suggestion about mirror neuron activity. They write: "In a similar fashion, it is conceivable that externally-generated MN activity serves the purpose of 'retrodicting' the target's mental state, moving backwards from the observed action."(p.498) However, one should note that they suggest not that they follow the same process but that they serve the same purpose of retrodicting. I have already pointed out that the described simulation routine actually did not retrodict the goal but simply tested an already conjectured goal. Thus, one might say that if mirror neurons actually work differently than the 'goal-to-action' simulation it might be *more promising* for their role in 'retrodicting' others' intentions and goals. Interestingly, if one reads the following sections carefully it turns out that there is not much overlap between the simulation mechanism of trial and error retrodiction and the suggested function of mirror neurons. Gallese and Goldman's idea is that the response of mirror neurons 'represents' or 'retrieves' an action plan, which we during action observation due to our visual perception can attribute to the other agent and then proceed to inhibit its overt execution. Gallese and Goldman are aware that this process is not very much like the simulation process that they have just described in such great detail. But they defend their insistence on a terminology of simulation and retrodiction for mirror neuron activity as follows:[211]

> Is it clear that anything similar to simulation occurs in externally generated MN activity? The point is that MN activity is not mere theoretical inference. It creates in the observer a state that matches that of the target. This is how it resembles the simulation heuristic. (Gallese & Goldman, 1998, p.498)

What they seem to find to be essential for something to be a simulation is 1) that it is not a process of theoretical inferences and 2) that there is some kind of 'matching' of the state of the observer and the agent or 'target,' to use their mindreading terminology. They suggest that whether this matching state is brought about by pretense or is perceptually evoked does not matter as long as it *results* in a matching state. Further, it is by way of this matched experience – and not a theoretical inference – that the observer gets some

[211] As mentioned, I see this dissimilarity as a good surprise for their mirror neuron theory because the cumbersome simulation process seems very implausible for the purposes of basic action perception and understanding. It does however obscure how exactly we are supposed to understand the idea of mirror neuron activity as a simulative process. Furthermore, it hints at why this attempt to force a marriage between simulation theory a la Goldman and mirror neurons might have led to a number oflots of theoretical confusions.

kind of experiential knowledge of the other's actions and their typical consequences and precursors. Given this very thin notion of mirror neuron simulation as 'matching,' it does seem rather forced to talk about a simulation *process*, or a 'retrodiction', a 'moving back' to the action plan from the action observation. One obvious reason for using the terminology of Goldman's explicit pretend-based simulation process for mirror neuron activity is that the two authors want to suggest that we can get some kind of mindreading from the latter.[212] Accordingly, what is important for Gallese and Goldman is not simply a partial resonance in the observer of the observed goal-directed action but some kind of qualitative or neurological matching that can be interpreted as a mental state matching. They refer to findings of 'mirroring' muscle facilitation in support of the idea that "when one is observing the action of another, one undergoes a neural event that is qualitatively the same as an event that triggers actual movement in the observed agent." (p.498) I have already mentioned both behavioral and fMRI evidence for activation overlaps between execution and observation of actions, but the reference to a 'qualitative similarity' proposes a similarity in the experience from a 1st person point of view. Thus, a move is made from an overlap in action 'representation' or 'retrieval' to an overlap in qualitative experience. This is a very important move if social cognition is thought to be about 3rd person mindreading and mental state attribution. The idea of the simulation theory is that we via simulation can engage a 1st person pretense representation that can yield some sort of 1st person insight into the broader mental states of the 3rd person. Now this introspective element of the simulation theory seems to be exported to the realm of mirror neurons. Further, the evidence of motor facilitation is used to move beyond the sub-personal activity of single mirror neurons to more of a temporally extended and experienced simulation process.

D. Shared representations versus input-output simulation

Gallese and Goldman suggest that mirror neuron activity supports a simulation of some action related mental states of others because they create a qualitatively similar or shared representation of an action plan and thus a link between the observer and the observed

[212] An additional reason to insist on the simulation framework could be due to its connection to the widely accepted – but in my view widely misleading - theoretical framework of computational action theory and the idea of inverse models working to produce action commands from action goals. More on that another time.

agent. At one level they are cautious not to say that the process happens as suggested by the simulation theory. More precisely they do not want to say that it is a relatively voluntary or explicit process by which action decision mechanisms are taken off-line and used to channel the assumed and/or pretend beliefs and desires of the other to reach an action decision or mental experience that can then be attributed to the other. Mirror neuron simulation is proposed as *implicit* and more like a *direct 'mirror'* of the other's action that provides some kind of broader *embodied shared representation* insight into the action of the other.[213] Thus, Gallese and Goldman link mirror neurons to the simulation not by way of the *procedural details* but rather by way of the *resulting shared representation*. One can of course question whether it is plausible that mirror neurons yield such a shared representation or neurological matching. I have already mentioned Pierre Jacob's critique of this very idea of mirror neurons as acting as a mental 'tuning fork' between agent and observer. In the next chapter we shall see that not only Jacob but also Jeannerod and Csibra vehemently argue that action mirroring does not yield understanding of goals or prior intentions. I think that it is right to question the idea of mental state mirroring and the idea that we via mirror neurons come to have 'qualitatively similar' agent neutral shared representations of intentions and goals. But, my main question is whether a lack of exact neurological or experiential matching would preclude mirror neurons from playing a role in our understanding of the goals and intentions of others? I suggest that it is a good thing if mirror neurons do not follow the procedural elements of simulation. And, if this serial and modular model of simulation is bracketed then it might be that the relationship between mirror neurons and goals and action predictions are not easily conceived of in terms of inputs and outputs in the way we shall see that Csibra assumes. To put it more acutely, if sensorimotor processes are not easily conceived of in terms of a serial simulation process, then there might be a way in which such integrations play a role in the 'conjecture' of goals and intentions that is excluded by the serial and un-dynamic 'boxology' model of simulation. I suggest that action mirroring might not form an encapsulated mechanism, but rather that the action sensitive neurons

---

[213] Goldman and Gallese do not actually say a lot about the function of neuronal matching beyond that it recreates a qualitatively similar experience in the observer. But how does that experience yield understanding, and how does one establish the cause-effect relationship between mirror neuron activity and understanding of the action? I will take up this question via my discussion of Jacob and Csibra's in Chapter 5.

are dynamically integrated with other affordance space relevant information. Action goals of others might be perceived not simply based on movements but via the overall temporal environmental engagements. Thanks to sensorimotor integrations we see many potential goals in our affordance space, and we might judge the goals of others by way of the relation between their actions over time and teleological properties and affordances of the present social and physical environment. Thus, I do think that sensorimotor processes could play a role in both predicting the action goals and 'retrodicting' the intentional action choices of others. More on my proposal later, but the narrow point here is that the dissimilarity between the pretense simulation and mirror neuron activity is a good thing, because the pretense simulation process not only takes the goal as the input - it also seems very slow, cumbersome and implausible for the purposes of basic action perception and understanding. However, this is my story - not that of Gallese and Goldman. While they do point out that mirror neuron responses are direct and automatic rather than intentional and explicit like pretense simulation, they do borrow several other elements from the pretense simulation process. I have already mentioned the idea of an internal mental state matching between observer and 'target.' Another notion exported from pretense simulation into their conception of mirroring processes is the idea of an 'off-line' process. This idea serves to encapsulate the mirroring process from our ongoing pragmatic engagements and to narrow its function to a separate epistemological use in 3rd person mindreading.

E. Simulation as 'off-line'

As we have seen, Gallese and Goldman propose that their rather forced marriage between simulation theory and mirror neurons simply rides on the notion of a 'similarity' or 'matching' between the observer and observed. I, however, claim that much more of the pretense simulation model is imported into the mirror neuron interpretations and is still causing various confusions and derailed arguments in the present debate. One key element of the simulation model that I find particularly problematic in this regard is the notion of the process being 'off-line.'

What does the 'off-line' metaphor imply? A common interpretation might be that an off-line process is temporarily decoupled from external connections and interactions. The primary role that the term appears to serve in explanations of mirror neuron simulation is to say that the action that is 'produced by' the simulation process is inhibited from overt execution. But given Goldman's simulation theory of mindreading maybe the off-line metaphor imports more into the mirror neuron theory than simply action inhibition? On Goldman's simulation model, we use our practical reasoning or action decision-making system that normally is guided by our own beliefs and desires as a simulator for the decision process of the other. This is supposedly done by creating relevant pretend beliefs and desires that we take the other person to have.

> Cognitive steps in predicting or explaining someone's decision by means of simulation...A dedicated pretend-state generator generates pretend beliefs and desires suited to the target agent. These pretend beliefs and desires are fed into the attributor's decision-making system (the same system that normally operates on natural, non-pretend beliefs and desires). The output of the decision-making system is taken 'off-line'. That is, instead of being fed into the action control system, the output decision is sent to the behavior-predicting and -explaining system, which outputs a prediction that the target will make that very decision. (Gallese & Goldman, 1998, p. 498)

The term 'off-line' is explained as pertaining to action inhibition and simulation in the sense that the process serves the purpose of knowledge of the other rather than action initiation. In Goldman's own description the inhibition keeping the process off-line happens prior to the engagement of 'the action control system.'[214] One question is thus how we are to understand the mirror neuron activity, which exactly is part of the action control system, as being 'off-line.' Gallese and Goldman's idea is here that one can make a division within the motor system between action representation and execution. The idea is that we covertly simulate the action as if it was to be executed and then it is inhibited just prior to execution. Thus, they maintain a serial non-dynamic process where the ultimate purpose or nature of the task does not influence the nature of the process. This notion of the motor system is discussed more in Chapter 5. I here want to discuss a different issue, namely that when one looks closer at the description of pretense simulation above it looks like the system is taken off-line much earlier than just before action execution, because the normal non-pretend decision process must have been put on hold. Gallese and Goldman talk about a 'dedicated pretend-state generator' that steps in to 'feed' the decision mechanism, but they never talk about the process by which we inhibit to stop our own beliefs and desires from invading the decision simulation. It is

[214] See already Goldman, A. (1993). *Philosophical Applications of Cognitive Science*, Westview Press. p.90.

difficult to empirically envision the proposed mental simulation model, but it seems that if we are to pretend to be in the 'other's mental shoes' we must to some extent leave our own shoes by way of an inhibition or decoupling from our own perspective to give room for the pretense. However, this need for a preliminary decoupling from one's own engagements is not discussed. One reason might be that Goldman's mechanistic information processing framework assumes that the decision-making mechanism normally stands around empty and idle, and given this idea no early inhibition is needed. However, this sort of framework of empty machinery waiting to be fed inputs makes very little psychological or biological sense, and I will therefore now turn to my alternative interpretations of decoupled simulation on the one hand and online mirror neuron and sensorimotor activity on the other.

4.2.6. An 'affordance space' view on mirror neurons and decoupled simulation

My claim is that if one does attempt to make psychological sense of Goldman's model it seems that the decision-making system would have to be taken 'off-line' in a broader sense. Namely, in order to use our cognitive mechanism to pretend to be the other, it seems to me that our own engagement with and perspective on the world must be disengaged before the simulation process can get off the ground. This sort of disengagement might not pose a problem for Goldman's rather explicit pretense simulation routine. But the question is what such disengagement would mean when turning to mirror neuron simulation. Do we also need a kind of general decoupling from our own concrete engagement with the world to channel the action plan of the other? Mirror neuron activity or other perceptually induced sensorimotor activations do not seem to depend on such decoupling. These activations are exactly caused by our concrete perceptual experience of the social and physical environment around us. I therefore propose that the activity of such sensorimotor neurons during perception should be understood as integral to our own current engagements and as playing a role in monitoring of the general 'affordance space' we are in. Thus, some mirror neurons might track the actions and intentional engagements of others, but I argue that this tracking process is typically not done by an 'off-line' motor system as the other's action exactly is related to my own action plans and possibilities. In other words, the perception of the other's action is part of my affordance structure. Furthermore, the second part of my proposal is that as long as the actions of others are monitored in this way integral to our own actively engaged perspective, then they do not constitute fully spatio-temporal action simulations but merely sensorimotor *sketches of teleological relations* between agent and context at various levels of abstraction. I see the rudimentary mirror neuron integration in many ways as similar to the rudimentary affordance integration of the canonical neurons, where many neurons and many different action affordances are sketched in parallel. When we choose an afforded action it gets specified motorically in concert with further perceptual explorations and thus more sensorimotor sketch activations. This being said, I do think that when we want to simulate the actions of others in greater detail, we do have to *disengage* from many of our concrete affordances and action plans to go through the load constrained process of action *specification.* More precisely, we have to abstract from our own perspective and goals to attempt to focus and 'simulate' the perspective and perceptual affordance structure of the other and thereby specify and 'fill-in' the sensorimotor details of the others' anticipated/perceived action. Thus, action simulation seems to come in degrees. Most importantly a distinction must be made between rudimentary stored teleological sensorimotor integrations, which can be used in parallel to represent the affordance space and to guide choices, and then the full specification and spatiotemporal simulation of actions, which have already been chosen and engaged. I suggest that only the latter are under the temporal and load constraints of overt actions. Furthermore, I hypothesize that mirror neuron activity alone like canonical neuron activity falls short of such full action specification. Lastly, I suggest that our normal engaged action choices are made by evaluating multiple affordances. Accordingly, to perform a more fully specified covert action simulation of imagined action or another's action, we must inhibit, disengage, decouple from our own concrete perspective and broader affordance structure. The point is that action specification exactly is a process of evaluating and choosing contextually appropriate sub-goals and motor act/means and thus specification happens in a real or imagined field of affordances. To pretend to have another perspective and another affordance structure I must *disengage* from my actual concrete affordance structure – because the pretense exactly involves engaging a pretense affordance structure.

This affordance structure hypothesis of mine is at the core of my reinterpretation of the role of mirror neurons in social cognition and also, more broadly, of my attempt to do away with the lingering mechanistic information processing input-output frameworks that I blame for many misguided cognitive theories well beyond mirroring. More specifically the distinction between spatiotemporal specification of actions and mere sensorimotor sketches will be essential to my critique of the traditional assumptions of the limitations of motor cognition that are rampant in most mirror neuron literature, and which we will see drive most of the arguments against the role of mirror neurons in social cognition. Here, however, I want to continue the discussion of the idea of decoupling and how it is ignored by Gallese and Goldman and, also, in the literature of mindreading and social cognition in general.

A. Decoupling and the 'false belief test'

Inspired by the traditional theory theory of mind-reading many cognitive psychologists have postulated that mindreading abilities depend on the normal development of a 'Theory of Mind,' or ToM for short. Some have suggested that rather than a 'theory' in propositional terms one should think of ToM as a cognitive 'module' (ToMm), possibly depending on highly innate and localizable neurological structures.[215] But however the implementation the main idea is that only by way of a well-developed 'Theory of Mind' can we attribute and reason about the mental states of others. Much research has been done to try to tap into and distinguish such hidden mental state understandings from mere understanding and prediction of overt behavior. Inspired by ideas by Daniel Dennett, so-called 'false belief' paradigms were proposed to test whether children, chimps etc. really could reason about mental states or merely predicted actions based on overtly observable features.[216] Thus, importantly, the false belief itself is thought to be an unobservable hidden mental state, and the test is thought to determine whether the subject has a well functioning TOM and thereby the ability to attribute and reason about such mental states or not. As we have seen, the proponents of theory theory and simulation theory share the view that the key to social cognition lies in high-level conceptual mental state attributions and they therefore also share the idea that the false belief test is testing something about our core social skills. But if mental state attribution depends on such a fancy late

[215] See: Fodor (1992), Leslie (1992) and Scholl & Leslie (1999).
[216] Dennett (1978). The task was first developed in an experimental form I believe by Wimmer and Perner. See Wimmer, H. & Perner, J. (1983).

developing module – how is it again that mirror neurons are supposed to support mindreading?

Here, I would like to pause and explain what I think that the typical 'false belief tests' are really testing – and how this puts the role of mirror neurons in social cognition in a new light. I propose that passing the false belief test – in its typical form - does not simply depend on belief recognition and attribution but, importantly, also on the ability to decouple from one's own perspective.[217] Further, I suggest that this disengagement is a rather cognitively advanced and late developmental skill and that it might be an important reason why the test is so notoriously difficult for children under around 3 ½-4 ½ years of age[218] – not to mention people with autism and chimps.[219] I also want to open the door for a kind of 'mental state' understanding that does not rely on such advanced disengagement or decoupling processes.[220] And then finally I will argue that mirror neurons and other fronto-parietal sensorimotor neurons might be important for the engaged action and intention understanding but do not functionally suffice for the decoupled kind. To make this argument, I will look a little closer at the false belief test.

The famous test has many incarnations[221] by now, but one frequently used is a cartoon or a puppet display version of the Sally-Anne test. The 'bare bones' of the test is summarized as follows in a non-scholarly entry on Wikipedia:

---

[217] There is some overlap but also important differences between how I use the idea of decoupling and how Alan Leslie uses it in his famous article about pretense play as a precursor for a theory of mind. The key point is that Leslie uses decoupling to describe pretense hypothetical additions/exchanges within the affordance structure, whereas I use it to signal more general disregard for more elements within one's personal pragmatic perspective. Thus, in regards to a developmental story Leslie's 'pretense-decoupling' comes earlier than my affordance perspective decoupling. See Leslie (1987).

[218] From now on for simplicity I will just write 'children under 4-years', but there is significant variability between children

[219] For 'TOM deficit' and false belief difficulties in autism, see Baron-Cohen, Leslie & Frith (1985). For the same on chimps, see the article that originally spurred the debate, Premack & Woodruff (1978).

[220] My story could thereby accommodate various earlier findings that problematize the idea of a sort of 'all or none' TOM module that kicks into gear around the age of 3 ½ or 4 ½. Leslie and Baron-Cohen and others often characterize social cognitive skill prior to 'full-blown' mind reading as 'precursors to TOM,' suggesting that these might be important for mindreading, but do not really involve recognition or attribution of mental states. See for example Baron-Cohen (1991). But Meltzoff's findings of intention understanding and the ability to infer intended goal from clumsy failed action suggest that the early social cognitive skills cannot simply be seen as operant learning. See Meltzoff (1995). Further, Gagliardi and colleagues have shown that even adults have a problem totally leaving their own perspective – thus problematizing the idea of the automaticity of a TOM module and ease of mental reasoning based on what I would call 'decoupling' form one's own affordances. See Gagliardi et al. (1995).

[221] I focus on the 'Sally-Anne' version of the false belief task, but the original Wimmer and Perner (1983) version was with Maxi and a chocolate, and another frequently used versions include a pirate and a cheese sandwich – however the structure of these other versions is similar to the Sally-Anne vignette described

> The experimenter uses two dolls, "Sally" and "Anne". Sally has a basket; Anne has a box. Experimenters show their subjects (usually children) a simple skit, in which Sally puts a marble in her basket and then leaves the scene. While Sally is away and cannot watch, Anne takes the marble out of Sally's basket and puts it into her box. Sally then returns and the children are asked where they think she will look for her marble. Children are said to "pass" the test if they understand that Sally will most likely look inside her basket before realizing that her marble isn't there.[222]

The typical interpretation of why children under the age of 4 – and older children with autism - are systematically prone to fail this test is that they are not able to represent and reason about the mental states of others. In other words, that they do not have a 'theory of mind.' And, I am the first to readily admit that from a parent's perspective the change that happens around this age is truly amazing. And, in the experimental setting it is particularly astonishing to see how the little test subjects are so sure of themselves as they quickly volunteer the true location of the marble and thus totally fail to take Sally's perspective or maybe even to understand that it is different from their own. However, I want to hypothesize that it might not be because they fail to understand that Sally has a perspective or thoughts and perceptual experiences of her own – or even that they fail to somehow understand what Sally's perspective is. Spending time with 2-3 year olds makes it apparent that they are not particularly good at being considerate of others or at inhibiting their impulses, but their overly socially engaged behavior in my view far from suggests that they are 'mind-blind.'[223] My hypothesis here is that there is a pivotal difference between on the one hand the ability to understand that others have a different perspective and on the other being able to reason and choose based on that other perspective, i.e. letting it override one's own action affordances, goals and engagements.

---

below. Other often used 'Theory of Mind' test are 'appearance-reality' tests such as the so-called 'Smarties test.' This one is different from typical false belief tests in that it provides knowledge to alter typical expectations (there are typically Smarties in a Smarties box but this one has buttons), and then asks the child what someone else, who has not opened the box, would think is in the box (or retrodicting their *own prior beliefs)*).

[222] See http://en.wikipedia.org/wiki/Sally-Anne_test. There are some variation in how different researchers actually carry out the test that might be of great importance such as whether Anne is presented as "naughty" whether the child is ask a memory control question (do you remember where Sally first put the marble?) before or after the test question (Where will Sally look for the marble?) etc. As we shall see I think all these contextual differences are immensely important. But I have chosen this 'bare bones' description precisely to captures the general commonality between various false belief test.

[223] Scientific documentation of this general observation of pro-social behavior is given by, for example, Warneken & Tomasello (2007). And in regard to children's understanding of intentional actions, see Call & Tomasello (1998) and, also as I will discuss later Buttelmann and colleagues as they report that 18-month olds can perform helpful actions portraying their understanding of the false beliefs of others. (Buttelmann, Carpenter & Tomasello 2009)

In sum, it might be that little children can understand minds and false beliefs but have a problem disengaging or 'decoupling' from their own pragmatic perspective of goals and affordances, and due to the latter difficulty 'fail' the ToM tests.

A big challenge to the Sally-Anne-style protocol and its typical interpretation comes from Kristine Onishi and Rene Baillargeon have done some interesting experiments with 15-month olds based on a non-linguistic version of the false belief test that yield rather different results.[224] Onishi and her colleagues use an experimental paradigm based on the idea that the infants look longer at unexpected rather than expected actions.[225] Thus, they attempted to test whether 15-month old infants expect others to act according to their true/false beliefs or according to their own knowledge of the scenario. They found that 15-month olds seem to expect others to act according to their false beliefs:

> …we propose that the present results suggest that 15-month-old infants expect an actor to search for a toy where she believes, rightly or wrongly, that it is hidden. Such an interpretation calls into question the notion that preschoolers undergo a fundamental change from a nonrepresentational to a representational theory of mind. (Onishi & Baillargeon 2005)

Given the robust findings of the failure of 3-year olds to pass the traditional Sally-Anne test these new findings are rather surprising. The question is this: If Onishi and Baillargeon are right in suggesting that 15-month olds do expect others to act according to their own knowledge/perspective even when it is false, how come 3-year olds fail to give the right answer to the action expectation questions of the Sally-Anne experiment? Many have suggested that the problem might be connected to linguistic challenges – this might also be Onishi and Baillargeon's preferred explanation. They suggest that their studies call into question the fundamental TOM/mindreading module idea that 'preschoolers undergo a fundamental change from a nonrepresentational to a representational theory of mind.' I agree. But, still, I also do share the view of many TOM proponents that some kind of radical cognitive development does seem to happen around four years of age, which does not have to do only with linguistic abilities.

---

[224] In connection with Csibra's critique of mirroring in Chapter 5, I discuss an interesting experiment by Onishi and her group showing that young infants are able to follow and reason in a pretense scenario.

[225] The 'looking time' paradigm is a groundbreaking paradigm that has become enormously popular in infant studies as it provides a way of gauging the mental life of prelinguistic children. However, it is not without complications, and it seems tricky to always interpret the reason why an infant looks longer at a stimulus as generally due to surprise. Could not there be a series of other reasons for longer looking times? Well, this is naturally not to be discussed here, but I just want to underline that it would make the Onishi et al. findings more convincing if one could get similar results with different paradigms.

Rebecca Saxe, a prominent TOM and mentalizing researcher, rightly points out the sticking fact that the 3-year olds are not hesitant or reluctant but confidently and eagerly volunteer their consistently wrong answers.[226] However, my proposal is that the important cognitive development that happens around 4 years of age is not a sudden introduction to hidden mental states, but rather has to do with our ability to disengage or decouple from our own present concrete social and perspectival affordance structure. The idea is that the child cannot inhibit the urge to blurt out their knowledge of where the marble is - maybe the social urge to volunteer the true location of the marble is even enforced by their knowledge of Sally's lack of knowledge! They cannot reign in on themselves and their own desires and affordances and therefore fail to do what the experimenter asks them to, namely to passively report where Sally would look – *as* if she was in her own world and the child was not right here to help her in her search. I am suggesting that the problem in the Sally-Anne task might have to do with its uncooperative and anti-social 'fly on the wall' structure and with the requirement that the child has to inhibit a strong social affordance of their own to answer the question correctly. In brief, besides linguistic difficulties children might fail the Sally-Ann test not so much because of what they don't understand but because of what they can't do.

This speculation is backed up empirically by an interesting recent study by David Buttelmann, Malinda Carpenter and Michael Tomasello where they tested exactly false belief understanding of small children in what one might call a 'socially and pragmatically coherent' situation.[227] They set up a scenario where a child could help an experimenter by unlocking one of two boxes, either the one where the toy originally was put and which the experimenter tried to open, or the one where the child knew that toy currently was. They tested respectively false belief scenarios (where the experimenter had

[226] Accordingly, Saxe's objection to my interpretation might be that if the child really did know that Sally thought it was in the bucket, then the *tension* between this knowledge and the child's knowledge that the marble is in the box should result in some confused uncertain answer (which to some extent is seen in contradictory eye movements to the bucket in some 3-year olds, see Clements & Perner (1994)). But, as Saxe rightly points out, most 3-year old children seem very assertive when pointing to the box. One reply one could give to this challenge is that the affordance to blurt out the true location overrides the knowledge of Sally's false belief completely – i.e. that there is no tension. However, I would suggest something different, namely, that the knowledge of Sally's false belief might exactly *feed* the impulse to *correct* Sally's knowledge and therefore *reinforces* the urge to state the true location and might lead to an actual misunderstanding by small children of the odd uncooperative *task*. *Given* Sally's false belief, the most natural *social task* in the situation would be to correct her knowledge. See Saxe (2006).

[227] Buttelmann et al. (2009).

not seen where the toy was moved to) and true belief instances (where the experimenter did see the toy being moved). They found a significant difference between the actions of not only 2 ½ year olds but also 18th-month olds, such that they would predominantly open the box presently containing the toy in the false belief case (and show the experimenter the new location), and on the contrary help the experimenter open the 'old' box in the true belief (assuming that they must have another reason than finding the toy to try that box). This study indicates that young children do use their knowledge of other's beliefs, and certainly stresses the need to precisely analyze the social affordances of false belief scenarios in order to appropriately interpret the responses that they elicit in preverbal children.

But if mirroring, simulation, mentalizing and you name it are thought of as being *off-line* and *epistemological* processes, then these important differences in affordance structures and between false belief understanding and pragmatic and social inhibition are simply not seen, because the *affordance structures* of the experiments are not given any attention whatsoever. The Buttelmann et al. and the Onishi and Baillargeon experiments - if robust - suggest that the children at some level do understand that Sally thinks the marble is in the bucket and also that given that information she should be *expected* to look in the bucket. However, the Sally-Anne task includes a strong *parallel challenge* to the child that is not present in the Onishi looking time experiment, i.e. the child is not just watching Sally, but they are asked to involve themselves in relation to the scenario. However, the twist is that the kind of involvement they are asked into goes counter to the natural social affordance of joining the marble search. In regard to the quest to find the marble, their own perspective represents the marble as being in the box, and therefore they have valuable information – information that they *ought to* share with Sally! The Buttelmann et al. (2009) study precisely explores this social affordance and how it varies according to their knowledge about the beliefs of the other. However, in the traditional false belief tasks the child is precisely asked to totally neglect not only their own representation of reality but also the tension between this knowledge and Sally's knowledge and the affordance to share it. Rather, they must pretend to be a fly on the wall and simply – in a truly detached way – report where Sally will look, given that the tested child – most deceptively - lets her stay misinformed.[228] In brief, I hypothesize that small children cannot decouple from their immediate affordance structure regarding the marble location and the strong social affordance to cooperate in the search for it. It seems that the study by Buttelmann and colleagues lends direct support to this hypothesis. The Onishi and Baillargeon experiment can be interpreted as being set up in away where the infants do not have strong competing affordances – they are simply monitoring and predicting the action of others – and that they can do. The Sally-Anne test on the other hand seems to rely on an ability to decouple and inhibit strong affordances to successfully firstly understand the experimenter's full question and secondly perform the uncooperative disengaged task required. Thus, I hypothesize that it is the decoupling and inhibition ability that is not matured until around the age of 4 – not the ability to understand others as having minds of their own.[229]

B. Autism and affordance structures

These ideas of affordance space understanding and decoupling might also provide a useful background for understanding some of the differences in social skills between on the one hand typically developing children and some of the characteristic symptoms of children with autism and on the other hand between human social cognition and that of non-human primates. The clinical labels of Asperger's syndrome and autism represent an enormously broad spectrum of symptoms and degrees of impairment and my thoughts here are merely meant as hypothetical gestures as I do not have the required clinical expertise. However, since mirror neuron problems have been implied as underlying many

[228] In regard to deception, some experimenters present the task as one where Anne purposefully hides the marble from Sally and others simply suggest that she happen to put it in the box after playing with it. There might be some additional element of the child being invited to be naughty and to conspire with Anne and the experimenter in the deception case. But all this questioning to me makes it obvious that the child's affordance structure and its variations must be taken into account and be empirically tested before one can really know what is going on cognitively in the children that fails or passes the test respectively. I suggest decoupling from one's own affordance structure is important, but to test this proposal we need to work out the social factors at stake in each variation of the task.

[229] My distinction between the ability to decouple and the social understanding within our affordance structure might also fit well with the extensive findings by Rebecca Saxe's lab and others of the importance of the right temporal parietal junction (rTPJ) for mentalizing skill such as the false belief task that involves taking a perspective different from one's own; Saxe and Kanwisher (2003) and Saxe & Powell (2006). The question is whether this area really pertains to mental states and the understanding of others' perspectives as Saxe and others suggest or whether it rather, as suggested by for example Decety and Lamm (2007) and Mitchell (2008), supports attentional selection processes more generally and thereby our ability to disengage from our own concrete perspective. This is to say that the question is whether rTPJ might only become crucial to our understanding of others when the task demands us to leave our own present pragmatic perspective.

autistic symptoms I find it important to situate my view in regard to the typical 'broken-mirror' story.

Children with autism generally have notorious difficulties with many social, emotional, linguistic and overall communicative skills and, hereunder, also the Sally-Anne task. Or, to put it the other way around, tests and questionnaires used as parameters to diagnose autism are often primarily trying to measure these social skills. Accordingly, it has been proposed that one of the key underlying problems in autism is a dysfunctional TOM.[230] Gallese and other mirror neuron enthusiasts have in extension of their mirror theory of mindreading alternatively proposed the 'broken mirror' theory of autism.[231]. The idea is here, as the name suggests, that a mirror neuron system dysfunction is at the core of the notorious social difficulties typical of people with autism. As I do not share the view of mirror neurons as instantiating a relatively modular mechanism yielding agent-neutral shared representations, I naturally do not think that the idea of the dysfunction of this - in my eyes implausible - mechanism offers a very satisfying answer to the riddle of autism. As mentioned, children with autism have a difficult time with the Sally-Ann task long after the cognitive age of 4, but they do not seem to have equivalent problems with supposedly parallel, equally demanding non-mentalizing tasks such as the 'false-photograph' test.[232] Given the analysis presented in the previous section, it could be tremendously interesting to see how children with autism would do on the non-linguistic action expectation task of Onishi and Baillargeon or the helping paradigm of Buttelmann et al. I hypothesized that typically developing children under 4-years are able to understand false beliefs of others and predict their actions accordingly, as long as they are not required to decouple from their own immediate affordance structure. Accordingly,

---

[230] Baron-Cohen, Leslie & Frith (1985).

[231] See Gallese (2006), Iacoboni & Dapretto (2006), Oberman & Ramachandran (2007), Rizzolatti & Fabbri-Destro (2009). For an interesting critique of the 'broken mirror theory' focusing on issues of imitation that I shall not discuss here, see Southgate & Hamilton (2008).

[232] Often the social specificity of the core cognitive symptoms of autism is based on reference to experiments that show that the failure on the Sally-Anne type tasks is paralleled with preserved skills in regards to 'similar tasks' that do not involve mental states reasoning. An example of such a non-mental task is the so-called 'false photograph task' developed by Debbie Zaitchik. This task involves a puppet or cartoon series like the Sally-Anne task, where a cat is on a bed, a picture is taken, and then the cat is moved to the chair and the child is asked where the cat is in the picture. (Zaitchik (1990)) Thus like in the Sally-Anne task they have to report not what they see or where things are now, but include prior information about a different earlier circumstance. 3-year olds apparently also generally fail this picture task but older children with autism do very well on the picture task as opposed to the Sally-Anne task. The interpretations of the results are challenged in a footnote at the end of this section.

typically developing children under 4 pass the 'expectation' false belief test but fail the Sally-Ann test. Given the wide ranging social difficulties and the many other behavioral abnormalities of children with autism, I would expect that their difficulties would not only have to do with disengaging their own affordance structure. Instead I would hypothesize that they will have quite different and probably compromised hierarchical and attentional flexibility and connectivity already within their own perspective. Children with autism often engage in highly repetitive action engagements and do not seem to evaluate and choose between affordances within their own perspective in the same way typically developing children do. Thus, they often seem to have a great deal of problems already within their own perspective, and in how they represent and move their attention within the affordance space that they share with others. It seems to me that their affordance space might be more concretely object-centered and possibly limited to a more '4-dimensional' representation of individualized particulars over time. This should be understood in contrast to typically developing children, who seem to a higher degree to generalize and simultaneously represent an affordance space that meaningfully points beyond itself in multiple hypothetical, anticipatory and social directions.[233] In other words, I speculate that some core social and linguistic problems of children with autism might be related already to the way they hierarchically represent and relate to *their* affordance space – even before they are asked to disengage from it. Hence, in regard to false belief testing, my prediction would be that many autistic children would have trouble not only with the Sally-Anne test but also the more simple action expectation test of Onishi's group and also with the 'helping task' of Buttelmann and colleagues. As mentioned, within the mirror neuron debate autism is repeatedly proposed to be due to a mirroring problem. I actually agree that visuomotor neurons such as mirror neurons might be involved, but I would deny that the symptoms of autism are narrowly linked to mirroring or shared action representations. Instead, I think that the symptoms are more broadly linked to a difficulty with representing a multi-layered and massively parallel affordance space and in orchestrating and emotionally/motivationally navigating multiple

[233] In this regard it would be fascinating also to look at the many recent developments in the research on counterfactual thinking and mental time-traveling as uniting typical memory and future action planning abilities and somatic and motivational aspects of action choices, and compare these findings to motivational abnormalities and the atypical – but sometimes savant-like – memory and future planning thought processes of people in the autism spectrum. See Gerrans, P. (2007) and Pacherie & Haggard (In press).

simultaneously perceived teleological sensorimotor integrations. People with autism often seem to have a difficulty with the hierarchical 'zooming' in and out that is typically used to navigate and chose actions in situations with complex social and linguistic meanings.[234]

In sum, I hypothesize that some of the difficulties that children with autism have of representing other people's perspectives, emotions and presumed goals appear already *within* their own affordance structure representation – and not only when asked to reason from the perspective of others. And if their perception is of a sort of rich 'bloc universe' but the representation of the counter-factuals and teleology of their affordance space are impoverished, then it might make sense that autistic children do not seek social engagements and have a hard time with the notorious parallel and complex affordance spaces of linguistic communication, not to mention metaphors, pretense etc.[235]

But, this alternative story should be seen as the sketchy speculative proposal it is. There is not much reason to believe that the complex and heterogeneous spectrum of symptoms and characteristics of autism can be understood or accounted for with a simple idea like a limited fronto-parietal representation and anticipation of the shared environment and intentional actions. My proposal is meant to counter respectively the 'broken mirror' and the 'broken TOM module' theories of autism, rather than to pose as a full-fledged theory itself.[236]

---

[234] The point is that we precisely need to simultaneously represent different affordances at different scales and contextual orders to feel the need for such attentional shifts and 'zooming' moves. It is interesting that many people with autism are able to apply 'top-down' strategies of attentional shifts but unable to intuitively reach this behavior. This could indeed be a sign that there is an affordances space problem.

[235] The findings of the 'false photograph task' (Zaitchik, 1990) also seem to make sense based on my broader interpretation of the affordance space of children with autism as being more concrete, in that they only need to be able to remember and represent the prior scenario in which the picture was taken not to represent multidimensional meanings within one structure. (That the typically developing kids under 4 often fail the false- photograph task demands a different explanation). Or, as Perner & Leekam (2008) argue in a very interesting article, the false-photograph task does not really involve a 'false' photograph and is actually not structurally that similar to the false belief task where one needs to entertain two different perspectives on the same scenario simultaneously.

[236] A recent study by Cattaneo and colleagues suggested that children with autism had a radical lack of temporal integration and anticipation in 'mirror' neurons of the fronto-parietal areas. Rather than having action sequences hierarchically and temporally coordinated such that each upcoming motor act is anticipated and begun already during the preceding motor acts, the children with autism seemingly prepared and performed each sub-goal action independently. (Cattaneo et al. 2007) This along with evidence of an abnormally small cerebellum in most children with autism presents very interesting findings that need to be addressed in any attempt to theorize the social cognitive problems in autism.

C. Cognitive continuity

I now turn to the issue of how a social affordance model as opposed to the traditional mirroring theory would make sense of cognitive continuities and discontinuities. As we only have direct evidence of fronto-parietal mirror neurons from single cell studies of macaque monkeys, it is extremely important for the argument of Gallese and Goldman and other proponents of the idea that mirror neurons might underlie basic human mind-reading abilities, to establish some kind of cognitive continuity between the social skills of monkeys and those of humans. Thus, Gallese and other proponents of the mirror theory of autism might suggest that a 'broken mirror' is at the core of the autistic symptoms. But, they also want to suggest that there is a cognitive continuity between the social abilities of macaques and humans and between children and adults. In regard to the performance of typically developing children on the false belief test, one would not expect them to suggest that children all of a sudden at the age of 4 acquire a mirror neuron mechanism that was not there before. Hence, to even make sense of the proposal that a mirror mechanism underlies social cognitive abilities there is a story to be told about what it is that accounts for the rather sudden developmental change in typically developing children. Or differently put, if small children already have a mirror neuron system, and if it is so essential to social cognition, why is it that they just cannot seem to pass the Sally-Anne test? As we have seen, I suggest that the Sally-Anne task taps into a radical development not in mental state attribution as such but rather in children's ability to decouple from their own perspective. As I see it this idea seems more compatible with the findings of cognitive continuity and discontinuity than Gallese and Goldman's own story. Their argue that evidence from non-human primates suggests that these animals are not simply associating stimuli and responses in non-mentalistic ways like proposed by for example Cecilia Heyes.[237] They write:

> Being a 'cognizer', nevertheless, does not necessarily imply being a mind-reader, or a possessor of the ability to detect intentional states in others. The argument that seems to suggest the presence, in non-human primates, of elementary forms of mind-reading abilities comes from the discovery of deceptive behavior. In a series of field experiments, Hauser showed that rhesus monkeys can withhold information about food location in order to deceive conspecifics and obtain more food for themselves. Deception is particularly relevant here, since deceptive behavior calls for the existence

[237] Heyes (1998).

> of second-order intentionality, and therefore for the capability to attribute mental states to conspecifics. (Gallese & Goldman 1998, p. 499)

I shall not here engage the debate over these various anecdotal findings of mentalizing and deception in non-human primates. I simply want to voice a question that the positing of such deception skills raises: If it is possible to have the ability to deceive in the absence of the ability to pass the Sally-Anne test and other such highly advanced mentalizing tasks, then what are the neurological underpinnings of respectively the more basic and more advanced mentalizing skills? Gallese and Goldman seem to suggest that mirror neuron activity – present in new world rhesus monkeys – could be underlying the core of the complex ability to deceive. They also want to suggest that a dysfunctional mirror system is underlying the notorious social problems that children with autism face and which makes them unable to pass the false belief test. Yet typically developing 3-year olds should have a well functioning mirror neuron system, so how do Gallese and others explain their problems with the Sally-Anne test? These claims and findings seem to present a problem for Gallese and Goldman's simplistic mirror mechanism view. On my account mirror neurons and the premotor-parietal circuits more generally play a role in predicting and monitoring the affordance space we are in, and hereunder also the social affordances that are tied to the presence of other agents in this affordance space. In regard to rhesus monkeys, it might be that some clever (typically enculturated) specimen can understand the social affordance well enough to deceive a competitor present in the affordance space. The social affordance to deceive would here be based on something like the understanding of another monkey's competitive interest in a given affordance or goal-opportunity in the shared space – let us say some food. The observing monkey might in such scenarios develop the further goal and complementary affordance of hiding the food to avoid competition. In this way one could construe the deception as based not on *taking the other monkey's perspective*, but on understanding the teleological affordance relation between the other monkey/agent and a shared object of desire – the food. Thus in parallel to the discussion of the false belief test, the deceiving monkey acts within its own affordance structure.[238] On such an account of deception the key cognitive mechanism that needs to be in place is not simply a sensorimotor affordance space

[238] Note that the other monkey can be understood via affordance and intentional *relations* to a real and a hypothetical environment. It is therefore not mainly a question of ascribing *hidden* mental states to others, but rather about predicting potential affordance relations between the other monkey and the environment.

monitoring/anticipating system, but also the ability to have a social goal – in this case to understand 'deception' and action manipulation as a possible goal. Thus, I suggest that a radical difference between humans and lower primates is to be found in the sorts of things that can count as action goals.[239]

Going back to Gallese and Goldman and the simulation theory of mirror neurons, the importance of this discussion is partly the challenge that the findings pose to the modular information processing view of mindreading. But this discussion of affordance space, perspective and decoupling is also meant to throw light on my alternative view of the role of mirror neurons as involved in monitoring our concrete affordance space. I suggest that we can understand others' perspectives and predict their actions in a way integral to or own affordance structure. In other words, we monitor and predict the actions of others such that we can also monitor and predict what these actions afford us. If we are asked to isolate our knowledge of others' perspective from this broader affordance space and focus our concern and action choice on their particular point of view we must disengage from our normal action choice process which is based on balancing competing affordances from our own perspective.

4.2.7. **Clinical evidence – imitation behavior**

Gallese and Goldman suggest that various clinical phenomena provide evidence for an automatic ubiquitous mirroring process. More specifically they point to the syndrome known as imitation behavior, also mentioned in the previous chapter, where patients with frontal lesions often inappropriately imitate the actions of others. The appeal to imitation behavior is informative in that it reveals that their simulative mirroring theory tacitly is conceived of much like covert action imitation. I argue that this might not be the best representation of what is going on and point to a rather different characterization of the syndrome, which fits much better with my account of mirror neuron function in a broader sensorimotor context. Gallese and Goldman write that "a mind-reader represents an actor's behavior by recreating in himself the plans or movement intentions of the

[239] One should also note the much more narrow range of actions that seem effective in monkey mirror neuron experiments as opposed to human data. Further, it is interesting also to think about the studies done e.g. by Tomasello's lab suggesting that social cooperation and the intrinsic drive towards social relations might be key to understanding the extraordinary cognitive achievements of humans as opposed to other primates. I.e. our ability to entertain abstract social goals might be pivotal. See Tomasello (2008).

actor."(p. 498) Taken alone this claim might fall short of the explicit claim that mirroring is a covert imitation process. As discussed earlier, the issue and the debate over the imitative abilities in macaque monkeys is itself a very thorny issue. Gallese and Goldman are clearly aware of the imitation conundrum. But they seem to forget this problem as they use the clinical syndrome of imitation behavior as evidence for an automatic mirroring process. The driving force behind their argument is namely that mirror neuron activity yields an automatic 'action plan' that must be inhibited so as to not result in overt imitation. They write:

> Clinical evidence of a similar phenomenon [automatic action plan recreation during action observation] is found in so-called 'imitation behavior'. A group of patients with prefrontal lesions compulsively imitate gestures or even complex actions performed in front of them by an experimenter. This behavior is explained as an impairment of the inhibitory control normally governing motor schemas, or plans. It may be inferred from this that normal humans, when observing someone else perform an action, generate a plan to do the same action, or an image of doing it, themselves. Normally this plan is inhibited so that it does not yield motor output, but such inhibition is impaired in the patient population in question. (pp. 498-499)

Gallese and Goldman here claim that action perception causes the observer to generate a plan to do the same action. In so far as this 'plan' is seen as a rather minimalist sketch I do not disagree. However, they leave room for confusion as they characterize imitation behavior simply in terms of an 'impairment of inhibitory control.' One then gets the idea that all action observation would lead to overt imitation unless each such covert action 'plan' or 'image' is specifically inhibited.

Imitation behavior is often observed in connection with another related syndrome called utilization behavior. The combination is interesting in regard to the findings of visuomotor neurons in premotor areas, as they respectively seem to reflect the imitative properties of mirror neurons and the object affordances of canonical neurons. I do not want to deny that imitation behavior and utilization behavior could be linked to activity in respective groups of mirror and canonical neurons. But the question that I want to raise is whether we have reason to believe that this mirror and canonical neuron activity in normal agents would automatically lead to overt actions if not actively controlled and inhibited. It is possible that action plans are not normally explicitly inhibited but simply *out competed*. In other words, imitation behavior might not primarily be a 'release' syndrome but rather one rooted in a *broader problem* of loss of action drive. This is exactly what is proposed by Erik Rietveld in his discussion of utilization behavior in his doctoral thesis on unreflective action.[240] I have already in the previous chapters mentioned Rietveld's theory of unreflective intentional action as normally shaped and chosen in a 'field of affordances.' He is concerned with how our intentional actions are dynamically shaped relative to a horizon of multiple action affordances. Appropriate action choices are exactly appropriate if they are sensitive to the constraints and alternatives of the present and carefully weigh the choice in a sort of cost-benefit analysis between endogenous drives, emotional alerts and the current context. He characterizes utilization behavior as a syndrome where patients with frontal lesions are highly inclined to perform instrumentally appropriate actions in inappropriate personal and social contexts, without any signs of distress over their compulsive and inappropriate behavior. (Rietveld, 2008:172). Rietveld is a phenomenologically inspired philosopher, who seeks a detailed analysis of the phenomena in question over simplified models of underlying mechanisms. Looking for example at utilization behavior, he criticizes the common 'medial-system hypothesis' that narrowly characterizes the syndrome as one of disinhibition of stimulus driven lateral systems due to a lesion in medial areas of internally driven motor control. (Rietveld, 2008:172) This is important because it is exactly this characterization of the syndromes of 'frontal patients' that is referred to when imitation behavior is taken as evidence for the automatic mirroring. The question is whether in characterizing utilization and imitation behavior, it is right to focus so much on the dis-inhibition of stimulus driven areas, rather than on the lack of internal motivation and ideas that normally *compete* with but also direct our stimulus driven behavior. Rietveld draws attention to the fact that utilization behavior patients are apathetic rather than disinhibited in the sense of restless or impulsive. (Rietveld, 2008:173) He quotes an example from Lhermitte of a woman that would be apathetic all day but upon request would go and prepare a meal. Rietveld writes:

> Notice that preparing a meal is an action that requires focus. This is another indication that, notwithstanding the stimulus-driven nature of their utilization behavior, these patients should not primarily be characterized by impulsiveness or distractibility. (Rietveld, 2008:173, footnote)

He thus goes on to ask whether "apathy in the specific form of both lack of initiative (reduced inner-driven action) and reduced interest and distress, characterizes not only the daily-life of patients with utilization behavior, but should also be understood as a

[240] Erik Rietveld. 2008. *Unreflective Action.* Doctoral thesis, University of Amsterdam.

psychological feature 'producing' or 'determinant in' utilization behavior." Thus, he suggests that rather than merely being an accompanying symptom of frontal lesions, apathy could actually be a cause of the stimulus-driven behavior observed in such patients.

To spell out this contrast between apathy and disinhibition Rietveld discusses the analogy often drawn between anarchic hand syndrome found mostly in connection with parietal lesions and the imitation and utilization behavior of frontal patients[241]. Both are normally characterized as *dis*-inhibition symptoms; i.e. as failures to hold back an action impulse. In the case of anarchic hand or 'alien hand' syndrome,[242] the hand contra-lateral to the lesion suddenly has become disobedient and performs all kinds of socially inappropriate and 'naughty' actions to the patient's great distress. At the surface this syndrome might look like the similarly inappropriate actions produced by people with frontal lesions. Rietveld however argues that imitation and utilization behavior patients should rather be compared to clinical cases of for example akinetic mutism. Akinetic mutism is a syndrome where patients slowly lose their action initiative and eventually even lose their response to externally afforded actions. These patients totally lack endogenous action impulses and the last actions they seem to lose are strongly externally afforded actions such as picking up a ringing phone or a handshake invitation.[243] If imitation and utilization behavior are thought of as having apathy, a lack of concern and internal drive as an underlying cause for the random inappropriate imitative and affordance-utilizing responses, then the parallel to akinetic mutism is clear. Further, it is evident that anarchic hand syndrome, which is highly disturbing for the patient him/herself, is rather different from akinetic mutism and utilization and imitation behavior. It is not that the patient with the anarchic hand does not have any other action plans or preferences but rather that the disobedient hand does not seem to listen! There is some kind of coordination or local inhibition problem – and the hand really seems to have gained its own life and will – in addition to the patient's 'own' will and action plans. The anarchic hand acts against the

[241] See Rietveld, 2007, p. 178.

[242] Most refernces suggest an equvalence between anarchic and alien hand syndrome, but other researchers suggest that these are two different syndromes (Marchetti & Della Salla, 1998). To prevent a terminological mix up, I am referring to the phenomenon of as a hand that is 'disobedient,' or 'out of one's own control' as due to either parietal lesions or damage to the corpus callosum.

[243] The finding that these basic social affordances are the strongest inducers of reactions in people with akinetic mutism coheres with my hypothesis of visuomotor neurons as modulated by social affordances just like traditional canonical neurons respond to object affordances.

general will, concerns and knowledge of the patient and his or her desire to act socially appropriate, and this results in a conflict that naturally is very disturbing to the patient. In contrast, utilization and imitation behavior patients are generally totally undisturbed by the inappropriateness of their actions and do not seem to experience any conflict. If Rietveld is right in connecting these syndromes with akinetic mutism and seeing apathy and loss of action drive and concern as a possible underlying factor, then it makes a lot of sense that the imitation and utilization behavior patients could not care less – not only about their inappropriate imitation behavior and/or utilization behavior – but about much in general.

Hence, to sum up, Gallese and Goldman see imitation behavior as a product of disinhibited mirroring processes – and as evidence for something like a ubiquitous context independent mirror mechanism. However, the actions of imitation behavior patients might not reflect action plans that under normal circumstances would have gotten fully specified. The syndrome seems to reflect a loss not so much of inhibition as of initiative and drive. There seems to be a loss of competing action intentions and drive to act appropriately, which leaves a vacuum that then lets the externally driven action observations and object affordances drive overt actions in ways they normally do not. My point is that the theoretical problem arises from seeing the mirror mechanism as an encapsulated module rather than as dynamic processes interrelated with other cognitive processes. Further, Gallese and Goldman's descriptions leave room for the ill-conceived idea of motor activations where covert actions normally happen one by one, and generally is are all-or-none phenomena. In other words, the possibility of multiple simultaneous more or less specified action activations is left out. In the next chapter I will argue that such influential assumptions about motor cognition, as all-or-none occurrences that are always fully specified under great spatio-temporal constraints, are indeed very problematic for our understanding not only of syndromes like imitation behavior but also of the cognitive role of motor processes more generally. But, for now, I just conclude that imitation behavior provides little evidence for a mirror mechanism of automatic covert simulations of fully specified actions.

4.2.8. Mirror neurons and language

The last issue brought up by Gallese and Goldman that I want to mention is the proposed relation between mirror neurons and language, and between the mirror theory of action understanding and Liberman's motor theory of speech. Gallese and Goldman briefly refer to the fact that the premotor area in humans, which is homologue to area F5 in macaque where mirror neurons were first found, overlaps on the left side with Broca's area. Broca's area is an area famous for its important role in regard to language, particularly speech production and grammar. They do not further develop the details of the connection between mirror neurons and language in this present piece, but Gallese and various other authors have written extensively about the hypothesized evolutionary and functional connection between mirror neurons and language.[244] I will later in relation to my discussion of Jeannerod and his view of communicative intention say a bit more about the functional relations between such linguistic communicative actions and the affordance space functions I hypothesize these relevant parietal and premotor areas as having. What I want to attend to here is that Gallese and Goldman in their 1998 article also mention a more theoretical link to language. In connection with their interpretation of mirror neurons as simulators, in the sense of yielding a shared representation, they mention the similarity between this proposal and Liberman's famous motor theory of speech.[245] They write:

> It [the theory of mirror neuron based simulation] also bears a resemblance to the motor theory of speech perception advocated by Liberman, in which the common link between the sender and the receiver is not sound but the neural mechanism, shared by both, allowing the production of phonetic gestures. (Gallese & Goldman 1998, p. 498)

The problem with the parallel to Liberman's theory might not so much be Liberman's ideas per se, but more the fact that most interpretations of this theory stress the motor aspect as opposed to the *sensory* or *perceptual* aspect, and thus the motor command rather than the common sensorimotor basis of phonetic gestures. Gallese and Goldman's description of it is no exception. They write that it is the neural mechanism underlying phonetic production that establishes a common link between sender and receiver. It might

[244] In very brief terms Gallese and Arbib suggests that mirror neurons evolutionarily could have supported the development of language skills as the shared meaning of hand and mouth gestures could have served as foundation for more abstract shared symbols and meaning areas. See for example Arbib (2005), Fogassi and Ferrari (2007), Gentilucci and Corballis (2006).
[245] See Liberman et al. (1967) and Liberman and Mattingly (1985).

be that this mechanism is sensorimotor but by only using the terminology of production the implication is that Liberman's speech theory, as well as Gallese and Goldman's mirror neuron theory, are distinctly motor theories. In other words, it is implied that these theories about the role of the motor system in various central cognitive and perceptual processes. To me these implications are misleading, as I want to suggest that the idea of a self-contained motor system is misleading. As I see them, the mirror neuron and other sensorimotor findings such as Liberman's are - some might say ironically enough - a testament to the need for a radically new theory of motor cognition rather than simply the concession of a larger functional jurisdiction to the old notion of the motor system. In other words, Gallese and Goldman's reference to Liberman's theory as a 'motor theory' is indirectly feeding the fire of arguments *against* the role of mirror neurons in social cognition. As we shall see, Jacob and Jeannerod have written an article particularly devoted to refuting the social functions mirror neurons and the 'motor theory of social cognition.'[246] It is by way of linking mirror neurons to a particularly motor simulation process under a traditional interpretation of the motor system that Jacob and Jeannerod argue that such narrow motor functions cannot account for social cognition. Chapter 5 will take up this issue of how the traditional assumption of motor and social cognition that feeds into this argument. Here, I would just like to underscore that Gallese and Goldman, by simply focusing on neural similarity and the idea of mirror neurons instantiating a shared representation and the off-line use of these representations, lose sight of the possibly broader pragmatic social functions of mirror neurons both in regard to linguistic abilities and to other social interactions.

4.2.9. Consequences of interpreting mirror neurons as action simulators

In many ways I appreciate the intuitions of Gallese and Goldman, for example in regard to the importance of detection of others' goals for our social cognitive abilities, and to their observation that the cognition and behavior of non-human primates should not be seen as just operant conditioning, but rather that there is an important cognitive continuity and a grounding of higher cognition in sensorimotor processes. I am supportive of these general ideas. But my point is that one can accommodate these intuitions better without claiming an automatic agent-neural action matching mechanism of hidden mental states

[246] Jacob and Jeannerod (2005).

inhibited in normal behavior. I have already in the prior chapter discussed the inherent problems with the mirroring *metaphor*. Now I suggest that the marriage of the mirror neuron inspired observation-execution matching theory and the *simulation theory of mindreading*, takes the theory even further away from accommodating the heterogeneous visuomotor findings. The theory seems to come out even more intertwined with the traditional frameworks of modular input-output information processing than it was before. The matching process becomes an 'off-line' action simulation, a process by which one disengages one's own pragmatic goals and plans to channel someone else, 'as if' their action was one's own. It might be that Gallese and Goldman never explicitly formulated a strong mirror neuron based theory of mindreading, but their joint-venture article surely implies such a theory. Thus, the flirtation with the disengaged off-line model of the simulation theory further removes and isolates the mirror mechanism from the overall functions of the cortical motor areas in which the mirror neurons are found, and suggests an encapsulated process with an epistemic rather than pragmatic function. Ultimately, the overall idea of social cognition as depending on 3rd person mind-reading and more specifically the idea of simulation as providing a sort of epistemic 1st person glance into the minds of others carries over into the later versions of the mirroring theory to which I shall now turn.

4.3. <u>Embodied simulation & the 'action chains' version of mirroring</u>

We saw in Chapters 2 and 3 how the mirror mechanisms were originally mainly conceived on the basis of the idealized properties of individual strictly congruent mirror neurons. Then we saw via the flirt with Goldman's simulation theory of mindreading how mirroring came to be conceptualized more like a temporal process of action *simulation*, i.e. not just as a retrieval of some abstract action representation but more as a process of covert imitation, of 'as if' action supposedly leading to a broader mirroring and shared representation of the mental states behind the action. Accordingly, the connection with Goldman's simulation theory of mindreading has turned the focus of the mirror theory in the following ways:

1. From single cell mirror matching to a temporally extended simulation process
2. From mirroring of subpersonal action representations to sharing of internal possibly experienced mental states, and

3. From low-level sub-personal processes to personal level sharing.

Although Gallese and other neuroscientists in the Parma group turn away from Goldman's theory of explicit pretense simulation, I suggest that *all* of these three new simulation inspired elements of the mirror theory to some extent make it into the theory of implicit 'embodied simulation.'

All along there was an attraction towards a broader simulation process than simply a single cell observation/execution matching mechanism. In connection with the discussion of the possible role of mirror neurons in action understanding I referred to the suggestion of Gallese, Rizzolatti and their colleagues that it is our broader action knowledge from our first person action experience that can be extended to others' actions via externally triggered mirror neurons (Chapter 3, Section 3.3.2). However, on an empirical level it was still the activation of single strictly congruent mirror neurons that drove the interpretations. This was significantly changed by a ground-breaking set of studies first done on humans with the help of fMRI by Iacoboni's lab at UCLA and then replicated by the Parma group with single cell recordings in monkeys. These studies ignited a new version of the mirroring theory that attempted to combine the idea of classical mirroring with a broader action knowledge as both being part of an externally triggered 'embodied' action simulation process depending not only on strictly congruent mirror neurons, but also on some of the less symmetrical properties found in many action sensitive visuomotor neurons.

4.3.1. Action understanding beyond mirroring – the role of intentional context

I will start by discussing the empirical findings that have to a large extent provoked the change of the mirroring theory. Iacoboni and colleagues published in 2005 an attention-sparking article entitled 'Grasping the intentions of others with one's own mirror neuron system.'[247] The article reported studies showing that an action in an intentional context activates the mirror system to a greater degree than the action or context themselves.[248] Similarly Fogassi and his colleagues in Parma later that year showed with single cell recordings in monkeys, that many parietal mirror neurons were modulated by the final goal/target of an action rather than just the simple motor act. They argued that these

---

[247] Iacoboni et al. (2005).

[248] See Iacoboni et al. 2005 for the 'intentional context' experiment.

neurons were modulated by what they called the 'why' rather than the 'what' of the action.[249]

The key finding that Iacoboni and his coauthors report is that the mirror neuron system is differently modulated by different intentional contexts, i.e. different suggestive clues as to the ultimate purpose of the action. They found that the perception of video-clips of the action of lifting a cup more strongly modulated the mirror neuron system (i.e. lateral fronto-parietal circuits) when perceived in an intentional context suggesting 'drinking' versus 'cleaning' being the purpose of the action, but that the mirror neuron system was similarly modulated by the two contexts without action. To explain these results they hypothesized firstly that individual mirror neurons are differently modulated by the same actions given the implied ultimate goal or visually available target. This was the finding that Fogassi and his colleagues seemed to be able to replicate in monkeys, where very similar actions often activated different sets of parietal mirror neurons depending on whether the action was part of an 'eating' versus 'placing' action context.[250] Secondly, they also hypothesized that the stronger intensity of the modulation in the drinking context could be because this action yielded a more clear action prediction and a more dominant following action – namely of bringing the cup to the mouth. Thus, the idea is that the drinking context sparks a clearer sequential action simulation process than that of cleaning. Iacoboni and his colleagues make the theoretical suggestion that activity is due to a combination of strictly congruent mirror neurons and more broadly congruent or logically related neurons. Since this idea becomes the cornerstone in the 'chains of action' version of the mirroring theory I will allow myself to quote their reasoning at length.

> Our results suggest that a subset of mirror neurons in the inferior frontal cortex discharge in response to the motor acts that are most likely to follow the observed one.

[249] Fogassi et al. (2005). There are a number of tricky issues with the sort of fMRI study that Iacoboni and his colleagues did, as the data interpretations and theoretical conclusions are based on a string of assumptions. The single cell findings in monkeys reported by Fogassi and colleagues eases at least some of these concerns. I have already discussed the issue that some of the early studies report examples of single cell recording of mirror neurons that are modulated in an inhibitory way by relevant action stimuli – this cast a grim shadow on this sort of 'subtraction findings' of mirror neuron activity. Or to phrase it in a more general way; how do we track mirror neuron modulation by the sum of oxygen used when some mirror neuron use more and some less when modulated? Then there is of course the broader problem of how – given the vast heterogeneity of the single cell visuomotor findings – one controls that a task actually taps into mirror neuron activity and not a much broader range of visuomotor neurons.

[250] This study is not without complications of its own, since the two actions were not strictly motorically identical. See Fogassi et al. 2005.

> In other words, in the Intention condition, there is activation of classical mirror neurons, plus activation of another set of neurons coding other potential actions sequentially related to the observed one. This interpretation of our findings implies that, in addition to the classically described mirror neurons that fire during the execution and observation of the same motor act (e.g., observed and executed grasping), there are neurons that are visually triggered by a given motor act (e.g., grasping observation), but discharge during the execution not of the same motor act, but of another act, functionally related to the observed act (e.g., bringing to the mouth). Neurons of this type have indeed been previously reported in F5 and referred to as ''logically related'' neurons. In that previous study, however, the role of these ''logically related'' mirror neurons was never theoretically discussed and their functions remained unclear. The present findings not only allow one to attribute a functional role to these ''logically related'' mirror neurons, but also suggest that they may be part of a chain of neurons coding the intentions of other people's actions. What are the possible factors that selectively trigger these ''logically related'' mirror neurons? The most straightforward interpretation of our results is that the selection of these neurons is due to the observation of an action, also coded by classical mirror neurons, in a context in which that action is typically followed by a subsequent specific motor act. In other words, observing an action carried out in a specific context recalls the chain of motor acts that typically is carried out in that context to actively achieve a goal. (Iacoboni et al. 2005. p. 533)

It is interesting to see an interpretation that does not simply refer to abstract mirroring, matching or strictly congruent mirror neurons but also includes more of the heterogeneous properties of action sensitive visuomotor neurons. Here Iacoboni and his colleagues suggest that specifically the earlier mentioned (but largely neglected) 'logically related' neurons in which there seem to be a visual triggering not of the *same* (mirror) but of a *temporally following action* could be involved in the broader modulation in the context suggesting an action intention of drinking rather than the (more ambiguous or simply less appealing) action intention of cleaning up.

But how is this process of context sensitive action anticipation supposed to take place? They write here that the 'logically related mirror neurons' may be part of a *chain of neurons* coding the intentions of other people's actions. And they propose that "observing an action carried out in a specific context recalls the *chain of motor acts* that typically is carried out in that context to actively achieve a goal." It sounds like they want to suggest that different intentions are linked to different action chains and that the role of the context and action observation is to simply *trigger* the previously learned action chain. This is indeed the idea that also seems to drive the interpretations of Gallese and others in their later articles as they often refer explicitly to 'prewired chains.' However, the question is why we should assume that the action sequences are 'prewired' rather than dynamically and creatively specified in relation to the concrete perceptual context? Why not suggest that we could have a spectrum of sensorimotor integrations that ground a process of action prediction and initiation based on action and context perception without the need for prewired chains or even serial action-by-action simulations?

4.3.2. From mirroring to action chain simulation

The classical idea of mirroring as involved in 'mindreading' and action intention understanding was that action observation via mirror neurons gave rise to shared action representation and then a broader understanding of that action. Via the affair with Goldman's simulation theory it was suggested that action mirroring can give us access to the 'qualitative' aspects of the observed action and somehow 'retrodictively' to the intentions that would be associated with such action. However, Gallese and Goldman were already at the inception of the mirror-simulation theory aware that the metaphor of 'retrodiction' should be taken rather loosely in regard to action mirroring. Even within the 'higher-order' pretense version of the simulation theory the idea of retrodiction seems wrongheaded as the model suggests more of a 'trial and error' prediction and comparison model. As I discussed in the previous part of this chapter, given the boxology of the simulation model it remains unclear how we come to hypothesize the intentions or mental states (i.e. the input to the simulation process) to begin with.

Enter the hypothesis of 'pre-wired chains.' If we somehow stored intentionally organized chains of action, then we might have a wedge into how low-level mirroring of motor acts could contribute to an insight into the broader intention understanding. The idea is that via action mirror triggering of pre-wired chains we can anticipate the upcoming actions and thereby *predict* the goal and thus the local intentional outcome:

> The mechanism of intention understanding just described appears to be rather simple, i.e. depending on which motor chain is activated, the observer is going to activate the motor schema of what, most probably, the agent is going to do. (Gallese, 2007, p.662)

The 'simple' mechanism of intention understanding is thus that action chains lead us to anticipate action goals and by way of the goal and the idea of the observed action as intentional we understand the intention by way of the predicted goal. Thus, it is not a 'retrodictive' process as such but rather an anticipatory process. I shall later discuss this assimilation between intentions and anticipated action goals, but first I must address the big gorilla Gallese's quote places in the room. The pressing issue here is how the 'right' action chain gets activated, given that the same motor acts can have many goals. I would think that we generally anticipate action goals not as a product of *simulating prewired* action chains but by seeing and understanding the observed action in its teleological context. And, it is exactly the context that seems to modulate the 'why' mirror neurons described by Iacoboni. Gallese also appears to realize the essential importance of the context for our understanding of the action and for triggering the right chain. However, he wants to propose a mirroring mechanism that brings us from the action observation to the action understanding via an automatic mirror triggered action simulation. By defending the idea of pre-wired action chains he might be attempting to ensure the automaticity and encapsulated idea of action simulation as bringing us from low-level mirroring to intention understanding. However, as usual the cost of automaticity seems to be inflexibility and possibly impotence in regard to higher cognitive functions. Gallese himself suggests that the action chains are based on simple statistical associations:

> How can such a mechanism be formed? The statistical frequency of act sequences, as they are habitually performed or observed in the social environment, could constrain preferential paths of act inferences/predictions. This could be accomplished by chaining together different motor schemata. At the neural level, this would be equivalent to the chaining of different populations of mirror neurons coding not only the observed motor act, but also those that would normally follow in a given context. (Gallese, 2007, p.662)

In regard to the statistical issue, one should note that categorizations and associations are not necessarily based on frequency but could be rooted in more meaningful pragmatic interpretation. Accordingly, I would presume that we do not with similar ease associate any action or any event with each other. Rather, the neurological association processes might depend on some level of teleological or causal 'judgment.' As to sensorimotor integrations, it seems to be the case that judgments of intentional connections are very effective in associating actions and perceptual outcomes as opposed to randomly co-occurring events.[251] Figuring out the neurological workings and establishment of intentional, teleological and causal organizations of sensorimotor connections is a big project in its own right. Here, I just want to point out that if sensorimotor integrations and motor resonance[252] are organized and modulated by intentional factors then Gallese is

[251] Here it is interesting to think about findings of, for example, 'intentional binding effects' in which intended effects of actions are experienced as happening temporally nearer the action than unintentional effects. See for example Haggard and Clark (2003).

[252] I pointed to several motor resonance findings in the previous chapter that indicate that the way action

leaving out an essential part of the story of intentional action understanding when simply referring to statistical associations. I think that we need a more complex picture of sensorimotor integration, which includes our natural tendency to interpret events causally and actions intentionally and thus pick out elements of both our actions and of the perceptual experience.[253] As Gibson has pointed out there is no such thing as objectively given stimuli, rather the stimuli are picked out by our intentional and pragmatic engagements. The problem is that this causal/intentional dimension is mostly ignored. Like Gallese does here, many researchers simply refer to our sensorimotor integrations as the product of a statistical calculation. This turns sensorimotor connections into non-intentional reflexes rather than pre-intentional and causal predictors, which meaningfully can be used both for action choice and for understanding the pragmatic relations we perceive around us. This is for me an absolutely central point, as I think that Gallese here buys automaticity and encapsulation at the high price of not really being able to give a story of how low-level sensorimotor processes contribute to high-level social cognition. Further, there is the issue of the context of the actions. In the above quote Gallese includes the importance the social environment, but it is a little ambiguous how exactly the context is supposed to assist the action prediction and the establishment of action chaining. It seems that he simply wants to say that the context *triggers* the most probable action simulation. As mentioned, Gallese elsewhere also explicitly proposes that the motor schemata chains are pre-wired.[254] And it is for him exactly this pre-wiring that ensures that a simple action observation can activate a chain and the embodied simulation that then yields the intention understanding. The question is if one could tell a different and more flexible story. It seems unlikely that there is one distinct pre-wiring simply

perception yields motor resonance is highly influenced by the intentions of the observer. A further claim would be that the influence of the observer's intention enters the picture because these sensorimotor integrations are themselves organized intentionally and as such might ground our understanding of the particular goals or intentions represented.

[253] As I discuss in Chapter 5, the issue of teleological perception is an important insights that is empirically documented in Csibra's work. However, he interprets this teleological understanding in purely perceptual terms rather than as sensorimotor as I suggest.

[254] For example, earlier in the same 2007 article he writes: "Single motor acts are dependent on each other, as they participate in the overarching distal goal of an action thus forming pre-wired intentional chains, in which each subsequent motor act is facilitated by the previously executed one. This suggests that in addition to recognizing the goal of the observed motor act, mirror neurons allow the observing monkey to predict the agent's next act, henceforth the overall intention." (Gallese 2007, p. 662)

based on 'action kind' and 'context kind' parameters.[255] Furthermore, it seems unlikely that we would need to go through a detailed simulation process of one distinct 'action chain' rather than another to anticipate action outcomes – given by contextual factors. The broader evidence of sensorimotor activity in fronto-parietal circuits indicates that the perception of another person's action in a pragmatic context leads to multiple parallel sensorimotor representations. And I suggest that we understand the goal of the other's actions not by going through a spatiotemporally specified action chain but by a *relational* insight. This relational insight is based not only on the presently observed action and the social and physical context of 'potential goals' (i.e. affordances), but also on prior knowledge of the relational features of the agent and her context. In other words, we see the other's action in an affordance structure, and as we perceive actions and affordances we do not need to go through all the motions of the other movement to anticipate their action goal or intention. The anticipation of upcoming 'logically connected' actions is not on this account based on 'statistical frequency of observation' but on hierarchical pragmatic action goal representations. The idea is that the precise action anticipation activation pattern is not 'pre-wired' but adaptable and sensitive to the details of the present context. Goals are recognized relationally by way of parallel activations of sensorimotor representations, but not necessarily temporal simulations. Temporal simulations, as a consequence of my view, are neither something that we ubiquitously engage in and nor are they necessary for action goal understanding. Temporal simulations of the actions of others rather serve purposes of filling in the spatiotemporal details of the sensorimotor engagement between the observed agent and environment. Such simulations might tell us something about the causal properties of the target or movement constraints on the agent, and thus help to make exact temporal and motor predictions.[256] But again,

---

[255] Remember here for example the finding of Schütz-Bosbach et al. (2006) that the same 'action types' cause facilitation from perception to execution but not from execution to observation, discussed in Chapter 3, Section 3.5.2.C.

[256] There are several experiments that speak to this point. Salvatore Aglioti and colleagues have published a very interesting study showing that professional basketball players can predict the accuracy of free throws before the ball even leaves the hands of the shooting player. This should be juxtaposed with professional coaches who are visually familiar but do not possess the sensorimotor skill of the players and who only at a later stage can predict the outcome of the throw (See Aglioti et al.(2008)). The point in regard to my present discussion is that a detailed sensori-motor prediction does depend on a simulation and that this simulation depends on exact motor skills (See also Calvo-Merino et al. (2005)). But to *understand* that the *action intention* is hitting the hoop, one does not need to simulate the action in any detail. Sports is interesting in the dynamics between simulations and relational affordance understanding, because an important part of the good player's game is the purposeful subversion of the opponent's sensorimotor

even in such cases I still presume that we do not simulate by way of 'pre-wired chains' but via a spatiotemporal action specification process that relies on *current* dynamic relations to multiple parallel affordances and teleological relations of a specific context and our knowledge of it.

Consequently, if the action sequencing and simulation is thought of in dynamic terms and goal understanding as based on parallel sensorimotor representations of the affordance space, as I suggest here, then the traditional mirroring and simulation story must be profoundly revised. The process can no longer be thought of as a modular or insulated process where one is automatically brought from the inputs of a non-contextual low-level action perception plus a context perception to the output of action intention understanding. However, in contrast to my proposed dynamic and parallel interpretation of the sensorimotor basis of goal anticipation and action simulation, it seems that Gallese tacitly is clinging to this idea of automatic and modular mirroring simulation. We shall see in the next chapter that it is exactly on the basis of this assumption of modular simulation that Csibra, Jacob and Jeannerod vehemently argue that the intention rather than being a product of mirror neuron activity and action simulation must be *recognized* first based on the contextual information and then used as an input to pick out the right 'chain' motor simulation. My suggestion is that the intention is neither an input nor an output of a serial action simulation process, but rather that the intention is understood by virtue of relations between many parallel goal-organized sensorimotor integrations or 'schemas.' Thus, in my view the intention is neither purely perceptually understood prior to the action simulation (as we shall see Csibra, Jacob and Jeannerod argue), nor does it depend on an automatic serial spatiotemporal action simulation process (as proposed here by Iacoboni and Gallese).

---

predictions through deceptive affordance messages (look away, body fake etc.). Another interesting study here is the box lifting experiment by Hamilton and colleagues showing problems in assessing specifics of observed actions – I would suggest due to an interference from doing similar type action with different kinetic sensorimotor specifics box. Thus, as I interpret it this experiment shows that detailed simulation can be used to judge sensorimotor specifics of observed actions – such as a lifted box's weight – but that these details are irrelevant to the understanding of the intention of the other to lift the box. (See Hamilton et al. (2007)) A similar story can be given with studies finding temporal overlap between covert simulation and observed action such as slower action simulation while watching others' constrained movements etc.

4.3.3. Relations between action goals and intentions

We saw Gallese suggest that the action intention can pretty much be seen as equivalent to the goal anticipation. He writes:

> Ascribing intentions would therefore consist in predicting a forthcoming new goal. According to this perspective, action prediction and the ascription of intentions are related phenomena, underpinned by the same functional mechanism, i.e. embodied simulation. (Gallese, 2007)

I agree that action prediction and intention understanding are related phenomena, but if 'embodied simulation' is supposed to constitute an insulated automatic process of full covert action simulation, then my hunch is that neither goal prediction nor intention understanding depend on such a process. Full covert simulation in terms of sequentially reenacted actions probably supports detailed kinetic sensorimotor predictions like Csibra and others suggest. Csibra as we shall see infers from this that intentions are conjectured via a special purely perceptual mystery goal attribution mechanism. I however think that Gallese and others' are right in pointing to the importance of premotor and parietal sensorimotor areas for intentional and teleological predictions. But in my view these judgments and predictions are based on many parallel sensorimotor 'motor schemata' activations monitoring and predicting the overall affordance space, not as Gallese suggests on a serial chain of mirror neurons activated narrowly by actions. Furthermore, under my interpretation it is pivotal to distinguish between multiple degrees of action activation. Most importantly I insist that a distinction must be made between full action simulation that is specified and created specifically for each individual occasion on the one hand and stored sensorimotor integrations that can be automatically and directly triggered in parallel by multiple features of the perceived or the environment on the other. To me it looks like Gallese with his new notion of 'embodied simulation' focuses on sequential and rather specified action activations as triggered by mirror neuron activity, and that seems to undermine many of his otherwise often solidly post-cognitivist intuitions.

The difference I sketch here between my affordance space account and Gallese's embodied simulation account might look subtle, but it is significantly different which can be exemplified in the case of action intention and goal understanding. To put if roughly, Gallese suggests that we simulate actions and thereby understand goals and intentions. Thus, he assumes that we simulate one action at a time and simply introduce contextual information to disambiguate and predict the goal of the observed action. My story it different in that I suggest that we continuously monitor and predict not just *actual* but also *potential actions*, and that the intention and goal understanding does not result from sequential action simulation but from parallel, relational and dynamically updated features of the experience. Via sensorimotor integrations we monitor many potential goals in the shared affordance space, and our understanding of others' action intentions is guided by their overt motor actions in regard to this context. Accordingly, on my account we do not understand the goal by mirroring the observed motor action or by starting a cascade of motor chains predicting the 'why' of the observed action. Rather, the observed motor acts of the other yield important relational teleological information about the intentional relations between the other agent and the potential goals that are perceivable in a shared context. This relationally based intention understanding is obviously not just driven by hands grasping objects, but by all sorts of broader behavioral information such as gaze direction and past and multimodal contextual information (as shown by the occlusion study by Umilta et al. 2001, and the auditory-visuomotor mapping found by Kohler et al. 2002). Furthermore, naturally a lot of prior knowledge depending on areas beyond the fronto-parietal circuits is important for intention understanding. An interesting question in regard to higher cognitive resources and the affordance space model is whether this 'extra sensorimotor' knowledge is used to dynamically modulate our understanding of the concrete affordance space representations of fronto-parietal circuits. In other words, whether the visuomotor integrations in fronto-parietal circuits can be modulated by prior information not pertaining only to the objects and agents present but possibly also to character evaluations and such. Another big – and empirical - question for the affordance model is whether it can be seen as a template and broadened to a more general theory of social and pragmatic understanding that would include neurological processes well beyond the fronto-parietal circuits. I have developed the theory specifically in regard to the evidence of concrete affordance space monitoring in fronto-parietal circuits, which primarily seems to pertain to basic goal directed actions in our concrete space. However, most social and communicative affordances or action invitations happen in a much more abstract, hypothetical and metaphorical 'space of reasons' – or as one might say – 'space of hypothetical teleological representation.' I am obviously interested in suggesting that the social affordance model is important for social cognition well beyond fronto-parietal circuits, but for present purposes this is purely speculative.

The takeaway point is that the way I develop the affordance story within fronto-parietal circuits suggests a more abstract and relational role of sensorimotor representation than detailed overt action simulation, and thereby also a very different story about the role of such sensorimotor representations in action intention understanding. And, as I shall argue at greater length in the next chapter, this difference crucially reconfigures the debate about what role 'motor cognition' can play in social cognition. Further still, the idea that the action understanding is based on relational features of agents in a shared affordance space blurs the line between the understanding of observable actions and 'hidden' intentions. I claim that in so far as we can understand the intentions of others it is via relational features. This constitutes an attempt on my part to stick a behavioral and affordance based wedge also into 'higher' social cognitive processes. In regard to the discussion of 'embodied simulation' and action understanding, I have argued that Gallese's account focuses on serial and detailed action simulation and that 1) this simulation process is implausible as our normal process of action perception and prediction and that 2) this conception of 'simulation' seems to describe a process that might support detailed motor learning, recognition and prediction, but does not seem to give us much in term of intention understanding and social cognition. Gallese, however, thinks differently – and sees the theory of embodied simulation as a "unified theory of social cognition."

4.3.4. The 'embodied simulation' unified theory of social cognition

I have now introduced the new mirroring simulation 'machinery' of action chains. I have also discussed how this part of Gallese's model of 'embodied simulation' keeps the sequential, modal and input-output framework of the traditional simulation models, and discussed some reasons why this framework is problematic and implausible as a model of action understanding. However, Gallese proposes that 'embodied simulation' can provide basis *not only* for action understanding but serve as a *unified basis for social cognition*. In this section I will discuss this broader claim and further show the exact differences and overlaps between the 'embodied simulation' account and my account of the role of mirror neurons and fronto-parietal circuits in social cognition. Here, the focus will not be so much on the process of simulation and the functional elements of the simulation model but on the *product* of the process. More specifically, I will question the thought that embodied simulation provides some level of 1st person access to the 3rd person mental states by way of 'as if' shared representations – and it is these shared representations that are thought to serve as a unified basis for social cognition.

Gallese, Keysers and Rizzolatti first proposed the idea of 'a unifying view of the basis of social cognition' in a 2004 opinion article under that very title. This article is interesting not only because of the ambitious title but also, more specifically, because of the explicit suggestion that mirror mechanisms can provide a "neurophysiological account of the experiential dimension of both action and emotion understanding."[257] The authors thus move from the idea of sub-personal mirroring of action representations to a personal level experiential simulation process. This idea clearly carries the baggage of the flirt with Goldman's simulation theory, in addition to the new notions that Gallese and coauthors attempt to bring to the table. They suggest that social perception is special because of the way we share actions and emotions. They write:

> There is something shared between our first- and third-person experience of these phenomena: the observer and the observed are both individuals endowed with a similar brain–body system. A crucial element of social cognition is the brain's capacity to directly link the first- and third person experiences of these phenomena (i.e. link 'I do and I feel' with 'he does and he feels'). We will define this mechanism 'simulation.' (Gallese et al. 2004)

They here stress a link created between the 1st and 3rd person and also imply that the process has to do with a simulation of action that leads to 1st person experiential content – as they suggests that we via mirroring action simulation can use our link between 'doing' and 'feeling' to understand the feelings of others. This looks somewhat like the traditional simulative mindreading focus on getting to the hidden mental states of others, but with a less 'propositional' focus. The thought seems to be not that we can 'read' hidden mental states, but that we can get a 1st person access to the experiential aspect of an observed action, emotional expression or undergone sensation. I do not necessarily disagree with the possibility that such experiential aspect to some extent can be had when the other agent's circumstances are allowed to drive our sensorimotor and emotional-

[257] Gallese, Keysers & Rizzolatti (2004).

visceral responses. However, the question is how much of social cognition this sort of 'contagion' experience explains.

As I have tried to argue in the earlier sections of this chapter I doubt that we should see mindreading of hidden mental states as being at the core or as the 'basis for social cognition.' Instead, I have tried to emphasize our relational understanding of actions and their teleological structures and contexts.[258] As I see it, it is how others relate to us and the general shared affordance space that shapes our basic social understanding. Other people's individual 1st person experiential perspective is theirs to keep, it is exactly beyond our reach. But the 1st person 'qualia' of experience, the 'what it is like' to use Nagel's time-worn words, might not be very relevant for our basic social understanding of others. It is via the intentional and emotional relations to our shared world that we understand them. Interestingly, Gallese and his coauthors do sometimes seem to go beyond their simulation model and touch on the importance of the way we share relations to the world and to each other. One could actually see the 'embodied simulation' experiential sharing as precisely based on overtly available relational features – and, in so far as this is the focus, I agree that they have their finger on something important. For the most part, though, they venture right back to their 'action simulation to inner feeling' framework and suggest that it is the conscious sharing that drives the understanding, and here our theories clearly part ways.

A. Beyond the fronto-parietal action simulation

It is by way of the 'action to inner feeling' simulation model that Gallese and his coauthors in this and other article propose to have found the cornerstone of social cognition. Most prominently, the mirror mechanism is expanded from the area of intentional action to that of emotions and viscero-motor resonance. There have been various behavioral and fMRI findings of activation overlap in for example the insula and the cingulate cortex.[259] I cannot here get into either the empirical details or the theoretical conundrums of emotional mirroring. I just want to give an idea of the broader proposal that 'embodied simulation' could serve as a unifying mechanism of low-level

[258] Action understanding is the focus here, but I do think that emotions and the understanding of emotions are essential not only to social cognition but also to intentional action choices.
[259] See for example Dapretto et al. (2006), Wicker et al. (2003), Keysers et al.(2004).

mindreading and mental state sharing well beyond the premotor-parietal mirror circuits. While action mirroring has been proposed as important for functions of action and intention understanding and predictions, emotional mirror mechanisms are heralded as functionally essential to emotional sharing and empathy. Thus, the idea is that mirror mechanisms more broadly can be seen as the "neural basis for social cognition."[260] I have already raised fundamental concerns to this proposal already within the realm of action based social cognition. I shall not exclude that there are some islands of for example emotional social cognition that includes mirroring elements – but the point is that even in these areas the more basic 'mechanism' is not likely to be one of symmetric context independent or simulative mirroring. But this is an argument that is beyond my present project.

B. The shared manifold versus the shared affordance space

When Gallese and his colleagues speak about embodied simulation narrowly in terms of mirroring and an 'action to inner feeling' simulation process, it constitutes a major departure from the parallel and relational dynamics between agents and a broader social affordance space which I stress as important for basic action understanding and inter-subjective relations. However, as mentioned before, sometimes Gallese's intuitions seem to take him towards such a more relational and less sequential and encapsulated mirroring/simulation process. One place where these broader intuitions are apparent is his 2001 article titled 'The 'shared manifold' hypothesis.'[261] In it he writes:

> The aim of my arguments will be to show that, far from being exclusively dependent on mentalistic/linguistic abilities, the capacity for understanding others as intentional agents is deeply grounded in the relational nature of action. Action is relational, and the relation holds both between the agent and the object target of the action (see Gallese 2000), as between the agent of the action and his/her observer. (p.33-34)

I could not agree more, as I have repeatedly argued, it seems to be the relational nature of intentional actions which makes it possible that fronto-parietal sensorimotor representations can ground social action understanding. More generally, these relational, causally and teleologically organized sensorimotor representations allow us not only to see how others relate to 'object targets,' but also to perceive these objects and the rest of

[260] Gallese et al. 2004, p. 401.
[261] Gallese (2001).

the perceptual and social environment as affording various actions to both oneself and others with whom we share the affordance space. Thus at first encounter with this article, I excitedly thought Gallese with idea of a 'shared manifold' was going to present something close to my idea of a 'shared affordance space.' Gallese writes that his 'shared manifold' is supposed to be "a conceptual tool to capture the richness of the experiences that we share with others." (p.44) But the manifold for Gallese refers not primarily to the complex social, pragmatic and historical relations of a given situation, but rather to a multitude of cognitive elements of importance that we can share with others; he mentions actions, emotions and body schema. Thus, it seems that Gallese's idea is of a manifold of inner experiential mirroring and not a contextualization of mirroring in a manifold as I would suggest. It is almost as if the manifold is additive rather than relative and dynamic. Accordingly, Gallese suggests that the shared manifold is constituted by a series of simulation routines instantiated on a sub-personal level by "a series of mirror matching neural circuits." (p.45) Thus, he is still talking of mirror mechanisms that are simply occupied with reproducing the same action, expression or sentiment of the other in me. He continues to reiterate that these circuits have a "dual mode of operation, an expressive mode and a receptive mode." (p.45) This idea of two neatly distinguished modes of one and the same mechanism seems again to hinge on a lingering cognitivist conceptual framework. The underlying idea is that we have one self-identical mechanism that then can be fed an input of perceptual experience of another's action/emotion or some kind of endogenously formed intention to perform this action/express this emotion. Seemingly the immediate mirror 'output' is thought to be similar for perception and action inputs given that the mechanism is exactly thought to 'match' them. This is why the two modes are distinguished primarily by way of the source of the input and use of the output. The typical mirror story is that in the case of expressive mode i.e. endogenous input the mirror mechanism leads to overt action/expression, whereas in the receptive mode the output is inhibited from overt execution and used for epistemological purposes. I think this theory shows that the encapsulated mechanism framework used is far more misleading than helpful. I would argue that we have no reason to assume that any 'mechanism' or 'circuit' is behaving in exactly the same way during action perception and execution. Yes, there are many meaningful and important overlaps of activity for action observation and execution and also emotion perception and expression. But certainly it seems like there are many differences as well and, as long as these differences are not accounted for, we have little reason to isolate and reify self-identical and non-dynamic input-output circuits and mechanisms.

Kant and also Merleau-Ponty talk about a 'sensible manifold.' As I read Gallese, he might with the use of the term 'manifold' want to allude to a complexity of intersubjective relations and experiences, but as the theory is spelled out he only refers to various mirroring mechanisms and does not develop an account of how we are supposed to understand the complex interrelations and affordance structures. This understanding does seem to depend on not just a mapping of the other on me, or on simulating their perspective, but rather on the *concrete relations* and differences between me and the other. My point is that there is a social manifold or affordance space that we share into without necessarily becoming a medium channeling the actions and emotions of the other 'as if' one was in their flesh.

4.4. <u>Summary and conclusion</u>

The project of this chapter was to analyze and discuss the ways mirror theories have been transformed under the influence of the idea of simulative mindreading. I introduced the idea of social cognition as based on 3rd person mindreading and suggested that the focus on recognition and ascription of hidden mental states to others might be problematic. This sort of framework assumes a dualism between overt behavior and our inner 'truly' mental states or experiences. A dualism I argue is implausible, but has persistently been taken for granted or tacitly assumed, and which has limited both the empirical and theoretical developments in the field of social cognition. The research of what I would see as more basic inter-subjective and 2nd person social cognition is ignored or theoretically downplayed. However, without an account of this more basic building block of social cognition, the outcome seems to be that mindreading remains an obscure ability – largely relegated to some kind of cognitive and neurological black box called the 'theory of mind' ability. Goldman introduced his simulation theory of mindreading to give an account of 3rd person mindreading, which was not based on abstract theoretical inferences about the minds of others. But he retained the idea that social cognitive abilities are about 3rd person mind-reading abilities and his theory stays neatly with the inner-outer dichotomy and also within a sequential and modular input-output framework of cognition in general. I suggest throughout that the sensorimotor integration findings precisely suggest that this framework is implausible and misleading in important ways. Gallese and Goldman's joint article on mirror neurons and simulation theory tries to remedy on the one hand the lack of neurological specificity of Goldman's theory and on the other hand the credence given to the role of mirror neurons in 'true' social cognition - namely mindreading, i.e. the ascription of hidden mental states to others. They thus join forces to suggest that mirror neurons 'simulate' the actions of others and therefore could be a neurological precursor or a basis for high-level mindreading. Both Goldman and Gallese seem to be aware of the tensions between their respectively sub-personal data and high-level mind-reading theory, which is probably why they attempt cautious formulations of a mirror-simulation theory. However, I show that between the lines they suggest a much stronger mirroring-simulation theory – and it is this theory that since has become enormously influential in the mirror neuron literature. This section is also used to further develop how my 'social affordance model' and discuss how core social cognitive issues brought up by Goldman and Gallese such as "Theory of Mind' testing, cognitive continuity, autism and imitation behavior would be understood under this new proposed alternative model.

In the last part of the chapter I traced how Gallese and others – after moving away from the notion of an explicit high-level version of simulative mindreading - still retain core features of the simulation framework in their new 'embodied simulation' proposal. More precisely, what is retained is the idea of simulation as a temporal process yielding some sort of access to the experiential aspects of the 3rd person's mental states, i.e. we are still talking about '3rd via 1st person mind-reading' and the 2nd person social interaction is still ignored as contributing to the 'basis of social cognition.' I have also argued that the relational information of an agent in a context or in relation to the perceiver is pivotal to our understanding of others, and not because it can be simulated to give us access to inner states, but because mental engagements themselves often are grounded and revealed in overt social engagements. However, the relational aspects of the mental are – possibly due to the tacit dualism of overt behavior and inner minds - left out of the core story both in the case of high-level 'pretense simulation' and in the case of 'embodied simulation.'

We shall see in the upcoming chapter that it is exactly the ‘3rd via 1st person simulation’ version of the mirror theory that has come under hostile fire from various researchers and philosophers over the last years.

Chapter 5

Alternative interpretations of mirror neurons

5.1. Introduction

In this chapter I will discuss a few of the most vocal and influential critics of the traditional interpretations of the mirror mechanism and its functions. My aim of engaging these critiques is naturally to position my view of mirror neurons in relation to these, but as importantly also to show the strong cognitivist assumptions, which run through some of the major alternative interpretations of the process and function of mirror neurons. Thus, on the one hand it feels to me as if the voices of Csibra, Jacob and Jeannerod by way of the assumptions of their frameworks ignore the real promise of post-cognitivist progress inherent in the groundbreaking sensorimotor findings of the Parma lab and beyond. On the other hand, one can also see their respective analyses as simply taking many of the tacit implications of the mirror metaphor and action simulation that I have been pointing out in the previous chapters to their logical conclusion. In any event these critical discussions are important to my project as they show that a radically new interpretation of motor cognition, sensorimotor activity and social cognition is needed, and that mirror neuron research looked at through a radically new and possibly 'post-cognitivist' lens indeed supports such a re-interpretation of the broader cognitive processes and abilities in question.[262]

Most of the criticism takes the claim of mirroring proponents to be something like what I in the preceding chapter referred to as the simulation version of the caricature mirroring theory, namely, a theory that posits a ubiquitous process by which an observed

[262] I have already briefly discussed the findings and alternative mirror neuron interpretations of Cecilia Heyes and Caroline Catmur. Their view deserves a more full presentation but as my main concern in the present project is not developmental issues I will not discuss their work in further detail in this chapter. See Heyes et al. (2005), Catmur et al (2007) and Catmur et al (2008). Another recent outspoken critic and mirror neuron skeptic that I will not be discussing is Greg Hickok. I think that his 2009 article "Eight problems for the mirror neuron theory of action understanding in monkeys and humans," brings up a lot of worthwhile issues, but most of them and also Hickok's own underlying assumptions are touched upon in various chapters and I, in order to avoid redundancy, I have find it unnecessary to devote an autonomous section to this article. See Hickok (2009).

action/action goal is generating a covert imitation/simulation in the observer. Mirroring is conceived of as an agent-neutral process where the action observed is simulated 'as if' to be executed, and thus the action simulation is thought to provide the agent and observer with a shared representation of the action and possibly its broader intentional structure. The standard functional hypothesis is that mirroring, via the shared representation that it yields between agent and observer, has multiple important social cognitive functions. Action understanding is often the primary social function discussed, but mirroring is also thought to provide a broader basis for emotional and conceptual sharing needed for skills of empathy and language, and normal social interactions in general. As shown in the previous chapters, the caricature story does not do justice to either the empirical findings nor to all versions of the mirroring hypothesis, but the key figures of the Parma lab themselves and also the other prominent researchers of the human mirror system such as for example Iacoboni and Ramachandran continuously summarize the mirroring and its function as a process of covert simulation yielding shared representation. This simplistic version of mirroring is still enormously influential and therefore deserves all kinds of critical attention. In other words, even though nobody explicitly holds all tenets of this theory, it still seems to mesmerize and monopolize the debate. My hope is exactly to try and inspire a debate on the more complex spectrum of mirror neurons, broader sensorimotor findings and more inventive theoretical proposals, but in order to get there we must first chase the caricature off the stage.

As we have seen in Chapters 2 and 3, the empirical evidence shows a much more complex picture of mirror neurons and motor activation during action perception than the caricature model of mirroring lets on. The heterogeneity of the findings, in the face of the idea of a simple automatic observation-execution matching mechanism, is fueling some critical arguments against the suggested mirror mechanism. However, often the focal objection of critical re-interpretations is directed against the proposed functions and outputs of mirroring and whether they can be said to provide:

1. *Mindreading*; i.e. goal/intention understanding and thereby low level 'mind-reading', and
2. *Shared representations*, i.e. whether they support a direct matching process yielding shared representations between self and other.

Many of the further hypothesized roles of mirroring in various social cognitive processes such as empathy and linguistic ability can be seen as derived from the 'mindreading' and/or 'shared representation' functions and are thus also indirect targets of any criticism of the basic mechanism and its functions. I will discuss some of the key insights of some of the prominent reinterpretations – this will hopefully elucidate the mirror neuron research but also more broadly make explicit the theories of social and motor cognition that this research and also the critiques of it are inherently intertwined with.

Many of the prominent critics of classical mirroring, including Csibra, Jeannerod and Jacob, the views of whom I will be discussing in detail, often seem to be caught up in models of motor cognition conceptualized in terms of covert action, serial input/output information processing and in tacit assumptions about the difference between action and perception. These assumptions, which are often shared by proponents and critics of the standard mirroring hypothesis, lead to an understanding of mirror neurons as instantiating a rather specific motor simulation process. *Given* this narrow idea of motor cognition as covert simulation the critics make many convincing arguments that this specific process of action mirroring does not do much for social cognition. However, I hope to show that this view of motor cognition is wrong or at least incomplete. I shall argue that the use of traditional cognitivist frameworks of motor cognition obscure important insights about sensorimotor integration and action organization, which can be gained precisely from the fascinating studies of mirror neurons and fronto-parietal circuits in general. The reason why I have analyzed the details of the empirical data and questionable interpretive choices presented in the previous chapters is that I have a different story to tell. I do not want to argue against Csibra, Jacob or Jeannerod within their frameworks. Rather, I argue that these frameworks, albeit widely accepted, are misconceived and that if one looks at the empirical data through a different theoretical lense the findings of mirror neurons and parallel sensorimotor integrations are convincingly fertile with evidence that points to motor cognitive processes beyond covert action simulation, and social cognition beyond 3rd person mindreading.

5.2. Jeannerod – motor cognition as action simulation

Marc Jeannerod is a prominent neuroscientist, prolific both as a researcher and in questioning and developing the theoretical frameworks of cognitive science. His work is on multiple levels enormously relevant in regard to my project. Firstly, simultaneously with the Parma lab he was in the 80's and 90's playing an influential role in parallel groundbreaking research on motor and sensorimotor functions. His interpretative theories of this motor research in terms of, for example, motor schemas have been very influential and I think pushed the field forward in very constructive ways. Thus, he has played a role in the process that led to the discovery of mirror neurons in the first place and also in the advancement of our understanding of motor cognition in general. Secondly, he is *explicitly* defending the idea that motor cognition in general is action simulation, the very idea that I think is so problematic and a hindrance to the advancement of our understanding of the motor role in social cognition.[263] Thirdly, he is a prominent critic of the Parma Group interpretations of mirroring as yielding shared representations that can serve as a cornerstone for various social cognitive skills. In total, Jeannerod is a fascinating figure by way of whom many of my themes come together by exactly, so to speak, falling apart.

In this section I will present Jeannerod's general view of motor cognition as based on action simulation processes. I will focus on his article from 2001 titled 'Neural simulation of action: A unifying mechanism for motor cognition,'[264] and his 2006 book on a similar topic but broader in its reach. *Motor cognition – what actions tell the self*. I will subsequently discuss what I call his perceptual and motor representation dualism, and will thereafter return to his more specific criticism of the standard interpretation of the mechanism and function of mirroring based on an article co-authored with Pierre Jacob.

5.2.1. Motor cognition as covert action simulation.

As the title suggests, in his 2001 article Jeannerod argues that motor simulation can be seen as the unifying mechanism of motor cognition. The first question is thus what he takes motor or neural simulation of action to be. His basic idea is that we can think of actions as composed of a covert and an overt stage, and that motor simulations are such covert 'pre-executed' actions. He writes:

> Covert and overt stages thus represent a continuum, such that every overtly executed action implies the existence of a covert stage, whereas a covert action does not necessarily turn into an overt action. The simulation theory to be developed in this

[263] I will point to a tension between this idea and the above-mentioned idea of motor schemas.
[264] Jeannerod (2001).

> paper postulates *that covert actions are in fact actions, except for the fact that they are not executed*. The theory therefore predicts a similarity, in neural terms, between the state where an action is simulated and the state of execution of that action. (Jeannerod, 2001, p.s103, my italics)

Jeannerod here describes action simulation as what I would call 'pre-executed' actions. In other words, he does not see any relevant structural or neurological differences between covert action simulations and overt actions except for the actual execution and the neural processes that specifically pertain to the execution of overt movements. Jeannerod does however also sometimes talk about the action simulation in more stationary terms as a motor or action *representation*. As I understand him, he is in these instances thinking of the simulation more like a stored *motor schema* that is not specified or actively 'run' or simulated at the moment. I argue this sort of stored rough sensorimotor relation is necessary as a category under which to understand mirror neurons and the multiple parallel automatic sensorimotor activations, which are part and parcel of both regular perception and regular intentional action choices, planning and specification. The important conflict in regard to Jeannerod's overall project is that if we include the notion of a stored motor schema then we are already operating with *3 categories of motor activation*: overt action process, covert simulated action process and action schema/representation. However, when talking explicitly about motor cognition Jeannerod mentions only overt and covert action – thus lumps the two covert categories together. It is the failure to keep a distinction between the different levels of action specifications and kinds of motor cognition that I insist derails the understanding not only of motor processes but by implication also of the sort of functional roles these processes can have in other cognitive tasks such as social cognition.

Specifying the informational content of covert action simulations he writes: "This covert stage is a representation of the future which includes the goal of the action, the means to reach it, and its consequences on the organism and the external world." (Jeannerod, 2001, p.s103) Notice that he now presents the action simulation in abstraction from the concrete environment to be acted into, and away from the motivational and emotional factors important for actual action choices. Accordingly the description of action simulation sounds here more like a stored action schema and not like a spatiotemporally specified process. Thus, Jeannerod seem to talk about action schemas/representations under the label of action simulations. But I argue that more needs to be added to the action representations to get actual overt actions than simply the overt movements of limbs. More specifically, there must be a difference between a covert action in the sense of a non-executed action simulation that is procedurally specified in its concrete environmental and motivational context on the one hand and this above mentioned static 'action simulation' as a representation of goals, means and consequences on the other. In short, Jeannerod seems to switch back and forth between a procedural and a static notion of action simulation that probably should not be seen as the same *neurological* event. But, Jeannerod hypothesizes that his notion of covert action simulation can be seen as the unifying mechanism of ALL motor cognition implicated in a great variety of processes from regular intentional action to imaginary action, action and object perceptual resonance processes and also as implied in action predictions, judgments and decisions. Of all these motor processes he writes: "Some of them are accompanied by conscious experience, some are not. In spite of these differences, they all bear the same relationship to action, both at the behavioral and the neural levels." (p.s103) It is here reiterated that action simulations are pre-executed covert actions. Thus, the idea that all action simulations can be seen as belonging to one category is based on the idea that they are all actions minus execution, and consequently all covert actions should be considered behavioral and neurological just like overt action but without the execution processes. This claim is by Jeannerod backed by behavioral findings of how various forms of action simulation in many ways closely resemble the temporal and biomechanical features of overt actions, and can cause action facilitation and an anticipation of effects.[265] This empirically based argument is crucial. I think that the findings are very robust, but the question is whether they tap into a universal feature pertaining to ALL motor cognition or only to a specific a sub-category thereof. I argue for the latter. As I see it, these findings more *narrowly* show that spatiotemporally specified action simulation - be it endogenously generated imaginary action or exogenously triggered action simulation – have similar temporal, biomechanical and load constraints to the ones overt actions do. I do not see what it is about these behavioral findings of for example elaborate motor

[265] The latter point has already been discussed at length in Chapter 3, when I discussed the difference between the sub-personal single cell findings of mirror neurons as opposed to the personal level behavioral findings. The temporal, biomechanical and load constraints on imagined actions are found by way of behavioral studies of global motor facilitation and reaction times. See additionally Decety, Jeannerod & Prablanc (1989), and also Schott & Munzert (2002), Stelmach, Castiello & Jeannerod (1994), and Johnson (2000).

imaginations that can possibly show that all processes of motor cognition are constrained in the same way as overt actions are. Given the discussed anatomical and functional findings that question the idea of the motor system as a unified output system this claim is empirically questionable. For example, the massively parallel perceptually induced sensorimotor activations in the fronto-parietal circuits found via single cell recordings seem to provide evidence against this idea. However, from the basis of the behavioral load constraint findings Jeannerod seems simply to infer that they apply universally to motor simulation and thereby to motor cognition. He does present a detailed analysis of the overlaps and also the many differences between the focal cortical and sub-cortical areas of neurological modulation during the various kinds of covert and overt actions. And one is left wondering why he is not more interested in giving a structured explanation of the differences between overt and covert actions, and of course between various sub-personal motor processes that might not easily be seen as fully specified covert action simulations. Rather, he simply concludes: “The above pattern of results on the mechanisms of covert action corresponds to the central stages of action organization, uncontaminated by the effects of execution.” (Jeannerod, 2001)

As we saw in regards to the mirror hypothesis, if covert action simulation is just like overt action minus execution one might want to ask why the former does not in fact end up in execution. Jeannerod brings this question up himself:

> …how come that covert actions, in spite of activation of the motor system, do not result in muscular activity and overt movements. There are two possible explanations for this absence of motor output. The first one is that motor activation during S-states [covert motor processes] is subliminal, and therefore, insufficient to fire spinal motor neurons. The other, complementary one, postulates that motor output would be blocked before it reaches the motor neuron level, by an inhibitory mechanism generated in parallel to the motor activation. (Jeannerod 2001, S106)

He thus proposes two complementary theories for the absence of motor output in covert actions: 1) subliminal processing, i.e. that the simulation activation might have the same neural underpinnings as the action but ‘weaker,’ and 2) active inhibition, i.e. an inhibitory process outside of and additional to the motor simulation that inhibits its execution. We See here the clear parallel to the discussion in Chapter 3 of the ‘sub-threshold hypothesis’ versus the active ‘extra mirror inhibition’ process. Here I discussed how the idea of symmetric mirroring logically favored the latter process, even if many empirical findings challenge this symmetry to begin with. Jeannerod who is not wedded to the symmetric mirroring idea seems to support a combination of the two. However he sees motor cognitive processes as fully specified actions that would always lead to overt action if sufficiently strong and non-inhibited. In other words no extra perceptual or other cognitive ingredients are needed for the overt action process.

I argue that we use our motor system to track and prepare multiple actions and relations simultaneously, and that static (sensori)-motor representations are significantly different from temporally specified actions. For such covertly activated schematic sensorimotor representations to become an overt action, they need first of all to 'out-compete' other sensorimotor activations, which could be seen as compatible with the sub-threshold idea. But I claim that such schematic sensorimotor representations then further needs to be specified and controlled under perceptual (or visual imaginary) guidance. And this latter point is a odd with both the view of symmetric mirroring and Jeannerod's view of motor cognition as simulation. Jeannerod writes that the differences between the overt and the covert action lie simply in the 'contaminating effects of execution,' but I suggest that not all covert motor activity is created equal. It is important to note that I agree that sensorimotor activations can lead directly to overt action and sometimes do need to be actively inhibited. Accordingly, it should be emphasized that I am not trying to install an extra insulated processing stage of action representation that then needs some kind of additional explicit act of the will to be specified for overt action. Rather, my idea is very much along the previously mentioned lines of traditional ideomotor and Bergsonian approaches, namely that the motor system does not naturally run idle.[266] Along similar lines, Paul Cisek gives a contemporary and empirically based interpretation of competitive and affordance based action selection.[267] Cisek's argument in many ways concurs with the ideomotor insights I have been pushing as he points out that if motor activations are left alone and not *competitively* inhibited by other stronger parallel action

[266] For James the idea of ideomotor action similarly suggest that the movement from idea or percept to action is automatic, after the establishment of a sensorimotor connection the action follows ''unhesitatingly and immediately." James. W. 1890. *Principles of Psychology* II, p. 522.
The Bergsonian idea is that in nature there is 'existence rather than nothingness' and that we should think of brain processes in similar ways, as if given room automatically filling and growing into it. Changing dynamics is thought of as the norm and, rather than physical objects in need of an external impulse or additional mechanism to change, Bergson thinks of mental processes more like forces. Bergson extensively criticizes the assumption of nothingness as being logically prior to anything in *Creative Evolution* (Bergson,1907), but develops a similar idea in regard to mental processes even in his earlier *Matter and Memory* (Bergson, 1896).

[267] See Cisek (2007) and (2008).

impulses they will lead to overt action if not actively inhibited. The key is that we can have many sensorimotor activations that exist in parallel and never need to be actively inhibited simply because we as agents have an alternative hierarchical action plan that never lets these sensorimotor activations become fully specified actions. It is a process with many simultaneous activations at the early stages of motor processing, but in accordance with the limited spatiotemporal nature of our embodied agency in one concrete environment the performance and detailed kinetic specification of actions are under strict load constraints. Oftentimes when Jeannerod talks about load constraints and subliminal activity, I think his ideas could very easily be understood as consistent with this idea of dynamic ideomotor competition. However, again my claim is that we then exactly also must operate with a category of early sensorimotor activity that is *not* constrained by the same requirements as overt actions. In this context, more specifically, a process of parallel competitive inhibition is hard to make sense of in a motor system that is thought to be limited to action simulation under the strict load constraints of full spatiotemporal specification all the way through. Consequently Jeannerod's view is highly problematic. I argue more specifically that without – and only without - such constraints can we make sense of not only competitive inhibition but also more generally of multiple simultaneous motor representations and, as we shall see later, of any kind of motor representations beyond our specifically practiced motor repertoire. Plenty of findings suggest the existence of both multiple simultaneous action goal representations and action goal understanding beyond our motor repertoire, but as we shall see Csibra as well as Jacob and Jeannerod are so caught up in their narrow motor cognition assumptions that rather than revising these, they infer that such findings must depend on non-motor processes. The narrow conception of motor cognition thus has considerable influence on their interpretations of empirical findings regarding mirror neurons and other sensorimotor processes.

5.2.2. Problematic dualism between perceptual & motor representations

Another important assumption that seems to run through many both mirror neuron theories and critiques is the idea that perceptual and motor/action representations are two distinct kinds of representations. On the one hand we have perceptual representations that are about the present external environment and on the other hand we have action representations that portray a possible future scenario by way of a goal representation along with the motor actions that effect this outcome. Many mirror neuron theorists use the terminology of mirror neurons 'transforming' a perceptual representation into a 'motor format.' As previously mentioned I think this terminology is problematic in that it posits a sensorimotor 'interface' of one to one translation or mapping, rather than integrative processes, which actually inform and organize the mental 're-presenting' that these circuits support. I don't know if many proponents or critics of the traditional mirroring hypothesis would actually when pressed deny that 'perceptual' and 'action' representations respectively rely on some elements of sensorimotor integration. In other words, I am not sure that they actually mean to support a neat dichotomy between perceptual and motor representation and process. However, I want to draw attention to the fact that the language and arguments both for and against mirroring often *imply* a dualistic dichotomy. Ironically this dichotomy was exactly one of the main theoretical issues that the early Parma group's fronto-parietal research was aiming to change. But it never fully happened. I tried in Chapter 2 to show that the distinction was inherent already in the very early idea of F5 as representing a 'motor vocabulary,' in that it was implied that actions were organized independently from the sensory integrations. I argued that it is difficult to see how this motor vocabulary could obtain its goal-centered organization without being thought of as inherently sensorimotor rather than primarily motor with, so to speak, supplementary sensory services. What is at stake is whether the idea of a purely motor goal representation is even meaningful – much less empirically plausible.

Whichever way the sensorimotor organization turns out to be, it is clear that the dualistic metaphor of separate motor and perceptual representations is setting the tone and shaping the terminology and schematics of the debate. We shall see for example how Csibra, Jacob and Jeannerod all contrast motor mirror action understanding with a 'purely perceptual' mechanism, and how they along with most people in the field think of mirror theories as exactly *motor* theories of action perception, social cognition etc. One might ask whether action representations really are thought of as 'purely motor'. Sometimes, and sometimes not, it seems. By Jeannerod mirroring is clearly thought of as a process by which a certain action perception is translated into the covert motor performance of that very action, and that it is the supposedly 'strictly motor' part of this process that is thought to be respectively either enormously important or largely irrelevant to social cognition.

In his 2006 book on motor cognition[268] he explains the difference between perceptual and action representations by way of two opposite 'directions of fit' and causal directions between world and mind. Perceptual representations are thought of as descriptive of the world, molded of the world or as he puts it, perceptual representations have a 'mind to world' fit and a 'world to mind' direction of causation. Motor or action representations on the other hand are prescriptive, molding the world to fit the mind and thus exhibiting a 'mind to world' causal direction. Jeannerod refers to Searle and his distinction between perceptual and action representations as pertaining to respectively present and future events.[269] Action representations are 'anticipatory' and 'proactive rather than reactive.' A further element stressed by Jeannerod is the temporal preexistence of the action representation in regard to the action it is, so to speak, "about." He sees this temporal distinction between the action representation and the action itself as suggestive of also some degree of existential independence such that the action representation can exist in the absence of the later overt execution of such action, but not the other way around. Jeannerod's distinction as described here, does not exclude sensorimotor integrations from either perceptual or action representations, in that he proposes a difference not in terms of mutually exclusive neurological underpinnings but in terms of causal and intentional directionality.

So far so good. Jeannerod might have his finger on something true when he draws these temporal and causal – rather than anatomical distinctions between perception and action. But the question is whether even this distinction might not be too neat in its directional dichotomy and in suggesting that mental representations can be clearly classified as either stemming from perception or moving towards action. It might seem sensible to distinguish between the 're-presentations' that are used towards and alerting an overt action in the future and those that are a product of a more 'purely perceptual' process. My point is that these perceptual and action representations are normally both dependent on processes of sensorimotor integration, and also often overlap. In other words, it is not just that sensory and motor representations rely on sensorimotor integrations, it is that they

---

[268] Jeannerod (2006) *Motor Cognition – what actions tell the self*. Oxford University Press.
[269] Searle (1983).

often seem to be inherently sensorimotor – directed simultaneously both against future action and from present perception. Perceived affordances and the sensorimotor integrations that support such affordance understanding are exactly examples of what one might call 'perceptions from the world towards the future,' and therefore not easily classified as perceptual or motor even under a 'direction of fit' definition. Accordingly, in regard to for example fronto-parietal sensorimotor circuits it might be more useful to talk about sensory and motor *aspects* of inherently sensorimotor representations rather than to reify two sets of mental representations.[270]

5.2.3. A Sensorimotor grounding of mental representations

The issue of mental representations is of course a big can of worms that I will treat somewhat tangentially in this dissertation. There has been much debate over the existence of mental representations. My interest is not so much in this debate but rather in what mental representations can be conceptualized as, if one moves away from the re-constructionist model of perception and the idea of serial information processing of representational content.[271] Behaviorists were notoriously skeptical of the possibility of a scientific study of processes between overtly observable stimulus and response. For me the issue is not about abandoning such scientific endeavors or the meaningfulness of cognitive categories and organizing processes beyond the overt behavior. I thus share the traditional cognitivist motivation and the critique of the idea that intentional mental agents can be understood solely on the basis of statistical couplings of rewards and isolated stimuli and responses. However, I am highly skeptical about the typical input-output constructivist conceptualization of mental representations as information carrying units being serially processed between stimulus and response. Both traditional behaviorists and cognitivists assume that one can un-problematically talk about a definite 'stimulus' and 'response.' Against this backdrop I would rather agree with critical voices like that of Gibson who argue that the individuation of a stimulus or a response does not make sense in abstraction from the larger spatiotemporal pragmatic experiential and

[270] Again, my view and terminology here is anticipated by Bergson already in Matter and memory (1996).
[271] See Akins (1996) for an important analysis and critique of this constructionist view of sensory systems and representations. Another big issue that I shall cowardly ignore in the present discussion is the issue of how the category mental representations relates to consciousness and more broadly how it is one can talk about representations at a sub-personal level, as I clearly do in this thesis.

motivational context of the perceiving agent.[272] For me the idea of mental representation must exactly help answer not just how we get from an individual stimulus to a response, but also how we actively analyze and categorize such stimuli in the first place. Another way of getting at the central issue here, is to say that I am interested in a notion of mental representation that is not meant as a personal level but as a sub-personal level category. In other words rather than thinking of ‘a’ mental representation as equivalent to an experience or a thought experience, I instead focus on composite aspects of the present mental activity – be it conscious or non-conscious. I think one could treat many traditional representational episodes of perception or thought along these lines[273] – but my narrow present focus is on how sensorimotor integration processes like those in fronto-parietal areas can be understood as representational in this composite/aspect sense. I am aware that my view here is rather unique and speculative, but I think that some comments and stage-setting might suffice to make the points I need for present purposes.

I suggest that one could ground an alternative view of mental representation in sensorimotor integrations, and thereby in new ways inform 1) our view of what sensorimotor processes are, 2) what such processes might contribute to mental functions, 3) the different respects in which sensorimotor processes might be thought of as both ‘inner’ and ‘outer’ and, lastly, 4) how this consequently would open the door for some social cognitive processes of mental state recognition and ascription that can be based on 2nd person social perception and interaction rather than 3rd person mindreading.
This would obviously be a big endeavor and the goal in this section is therefore merely to motivate these ideas by showing that some sort of alternative view of mental representation is not only possible, but is also needed to account for the neurological findings of fronto-parietal sensorimotor integrations.

---

[272] See Gibson’s early and excellent analysis of the uncritical use of the concept of the stimulus in empirical psychology (Gibson, 1960).

[273] If one thinks of a tree and characterizes the experience as a mental representation of a tree, I would here rather say that the thought process *involves* the representation of a tree. In neurological terms I think that one could find sub-personal processes that – based on earlier engagements - supports the tree aspects of the thought. However I would hypothesize that this would be merely one partial aspect of the actual present thought. In other words under my hypothesis even the ‘content’ of a thought or a perception is not a mental representation or a sum of a number of representations but rather present engagements beyond these evoked representations. See also the following section.

A. Representation, 'aboutness' and empirical content

The typical story of mental representation in analytic philosophy of mind focuses on a combination of some sort of personal level mental attitude such as a perception, belief, desire, recollection of a corresponding 'representational' content that the mental attitude is 'about.' Hence, the idea is that the mark of consciousness or mental processes is the 'aboutness' of intentional mental states. The content is mostly thought of in terms of some sort of factual or imaginary state of affairs. 'Emma believes that it is raining' and 'Emma hopes that it is raining' thus suggest that Emma could have two different mental states with the same content but different attitudes – they are both about the same possible state of affairs in the world but indicate a different relation between Emma and this content – namely either a 'belief-relation' or a 'hope/desire-relation.' The word 'representation' is thought of in terms of something representing i.e. being about something else, and the word mental suggests that the representation is couched in terms of a special mental relation or 'attitude' between a mental agent and the content.[274] Turning to neurological evidence, one question is whether this personal level story of mental representations and the neat distinction between form and content is very helpful in actually moving beyond behaviorism to characterize the processes behind the overt behavior that we can observe. A basic question is how we can represent content, how it is that we can come to categorize the real and imagined world in meaningful categories and feel this or that way in relation to it. The standard 'representationalist' story has here focused on a 'constructivist' theory of perception as construing an inner model of the world, which can then be stored in memory.[275] Given such a theory of perception and memory, the 'aboutness' theory of mental representation makes perfect sense: first we have our model of the world and then we can break it into compositional pieces and combine this content in an infinite number of ways and then think or feel various things about it. The problem of course is that the constructivist story of perception is problematic and empirically implausible in numerous ways.[276] I will not here attempt to

[274] See Davis (1995) and Lycan (2000) for an informative overview.
[275] The representationalist-constructivist theory of perception has deep roots in the history of philosophy, but is most directly rooted in Locke (1690). For classic cognitivist versions see: Marr (1982), Rock (1983) and Rock (ed.) (1997). As mentioned, what Akins (1996) provides is great for critical analysis.
[276] Beyond Akins (1996), I think that Bergson (1896) already points to a series of inherent problems with the representaionalist-construtivist view of perception. Then of course there is Ryle and his famous intellectionist argument against indirect perception and representationalism in Ryle (1949). Gibson is also a

give a new theory of perception but will simply suggest that it looks like our so-called 'attitudes,' our cognitive emotional and also pragmatic relations to the world, seem to play an important role in actually categorizing our perception in the first place. In other words, there does not seem to be a pre-perceptually given stimulus – a sense datum – and neither an inner construction process from such mythological givens. We seem to need a story of our internal cognitive processes that can help us actually define and pick out what counts as stimulus and response. Hence both classical behaviorist and classical cognitivist/constructivist approaches seem unsatisfactory in regard to a range of psychological processes.

One obvious question is whether mental representations always should be thought of as interpretable in terms of a strict form-content attitude-content dichotomy. To understand a non-constructionist story of perception and cognition more broadly we seem to need an understanding of some such non-'attitude-content' mental organizations or 'representations.' In other words, we need a conceptualization of mental processes that is done neither simply in terms of stimulus-response relations nor interpreted in terms of attitudes and contents. I am not arguing against the usefulness of representational accounts or divisions between intentional attitudes and contents that these are about *as such*, but rather pointing out that we seem to need a more basic theory that gets at the mental processes that are *prior to* such divisions.

B. Bergson and mental re-presentations

I suggest that we could broaden and reinterpret the category of mental representations through an analysis of sensorimotor grounded *re-presentations*. Instead of focusing on the 'aboutness,', on what is presented, the stress is here on it being something that is stored and can be repeated, something that is presented again. I suggest that we need a concept of stored mental presentations – stored interpreted and generalized aspects of past experiences through the prism of which we experience, categorize and use the present. This idea is inspired by Bergson's sensorimotor theory of cognition, which I have already referred to multiple times. The basic idea is that mental 'presentations' are

good source - even if one disagrees with his positive theory of perception the challenges that he raises to traditional theories are quite informative. See Gibson (1966), (1976) and (1979/1986). For more recent critiques, see Noë (2004) and Gallagher (2008).

always interpreted through 're-presentations' and thus depend not only on the facts of the perceived environment, but also on internalized and teleologically or causally interpreted relations of previous experience of perceptual change in the face of our own and possibly others' active engagements. The mental event thus becomes a dynamic product of both the ongoing engagement with the environment and the hierarchical re-presentations that one brings from prior experiences and current motives.

Okay, I know that this is a rather obscure description. But one point to notice is that I do not want to narrowly tie sensory inputs to motor outputs and simply claim some sort of internalized behaviorism or automatic associationism. This is because I see all experience as already pragmatically and concernfully interpreted by way of earlier represented experience. Thus, beyond the spinal cord there is no such thing as a simple stimulus-response association, as the perception already is pragmatically interpreted and temporally reaches beyond the present 'input.' The present perception is based both on the actual external features of the world and our historical and biological 're-presented' experiences and drives – of how we have previously engaged the environment. I shall for present purposes bypass the discussion of emotional and episodic memory factors in mental representations, and simply stress the sensorimotor and the temporal aspect as this is what I will directly use in my discussion of mirror neurons and how they might be involved in processes of mental representing. Again, stored sensorimotor 're-presentations' cannot simply be seen as statistic stimulus-response correlations (- what one might call 'internalized behaviorism'). My point is exactly that what counts as a stimulus or response are meaningful categories dependent on purposefully driven and goal directed behavior on the part of the animal, thus there is always an internal 'mental' dimension involved in categorizing the input and outputs.[277] Directly inspired by Bergson's view in his 1896 book *Matter and Memory*, one might want to say that the mental as opposed to the non-mental physical world is characterized by temporal *transcendence.*[278] The idea being that current mental events are anticipatory and that this

[277] I shall return to this point because this is essential to how I see mental processes as always transcending overt behavior in an environment, but also always to some extent revealing themselves in overt engagements.

[278] I do not mean temporal transcendence in the sense of eternity or going beyond time, but instead simply in the sense that more is perceived in the present than a time-slice of the current events would portray. Thus I take perception to show temporal transcendence in the sense that the it includes past experiences and future anticipations.

process rely on re-presentations based on past engagements. In Bergson's parlance: the basic function of cognitive processes is to use the past to envision a future and to drive our motivations and guide our actions to change the present.[279] We perceive the present not as some objective reconstruction, or 'redoubling of the world' to use his terminology, but rather that we perceive the present via the past towards the future.

Under this alternative Bergsonian interpretation of representations as involving 're-presentations,' I would agree with Jeannerod and other vehement critics of behaviorism who insist that understanding mental representations is essential to understanding the mind.[280] However, under this description of re-presentations as transcending the present – as representing the present 'via the past towards the future,' the 'direction of fit' and 'direction of causation' become less neat. The sensorimotor affordance integrations in fronto-parietal circuits can be seen as an example of such re-presentations. Here teleological relations between actions and perceptual changes are stored at some level of abstraction above the actual individual past sensorimotor engagements. And when new affordances of the relevant kind are perceived then such past generalizations are re-presented. My point is that such affordance representations have both world-mind and mind world directions of fit in that they are saying something both about the world and about how one could change it. One might suggest that this notion of re-presentation makes particular sense of processes that involve teleology and that such goal and affordance representations can be expected to involve sensorimotor integration. The general idea being that the perception of a present environment is interpreted through the prism of past experiences and action motivations towards the future. Perceptual experience affords action goals and emotions by way of the past. We perceive future goals and anticipate perceptions via past experience re-presented.[281]

With this story in mind it is interesting to turn back to Jeannerod's view on the specific issue of goal re-presentations. He argues that such representations must be 'centrally

[279] See Bergson, H. (1896) *Matter and Memory.* 1911 translated by Nancy Margaret Paul and W. Scott Palmer, London, Swan Sonnenschein.
[280] The irony of course is that Jeannerod thinks of this re-presentation view as a version of behaviorism.
[281] Endogenously generated re-presentations would carry the same pragmatic and temporally integrated structure I presume, but I will not try to make a general claim here that everything we would call mental representation must be sensorimotor or have a pragmatic 'present via past to future' structure. Instead, and much more modestly, I will suggest that there are many cases of mental representing where such a sensorimotor based structure makes sense.

stored' in that they are accessible independently of specific concrete perceptions or actions. Jeannerod refers to various experiments showing that we can access and initiate actions from stored 'motor representations' without somatosensory or visual feedback. However, my suggestion is that the very stored representation nonetheless seems to involve integrations of some level of perceptual experience and motor anticipation. I think that such representations at least in the context of the fronto-parietal circuits can be thought of as goal, affordance or teleological representations rather than pure motor representations. Furthermore, against Jeannerod's view, I would claim that even if such a sensorimotor representation is engaged endogenously in the absence of online perceptual guidance or feedback, it becomes specified as a spatio-temporal action in relation to a previously perceived or imagined environment. This 'motor' process relies heavily on processes and areas typically thought of as sensory or perceptual – even if the *overtly engaged* perceptual guidance is *presently absent*. One might therefore say that both 'perceptual' and 'action' representations involve moves from present to future over the past and speculate that both kinds of processes generally rely on sensorimotor representations. In short, under this conceptualization whether something is classified as a perceptual or a motor representation is a matter of *degree* rather than a mutually exclusive dichotomy of natural kinds.

C. Can we make sense of sensorimotor 'translations'?

Maybe it is simpler to make my point in regard to distinct perceptual and motor representations via a sort of 'reductio ad absurdum' style strategy. Admittedly, I have a difficult time of really making sense of the idea of a certain neurological point of 'translation' or 'mapping' between two independently organized types of mental representations. The terminology often suggests that perceptual and motor representations each have their particular format and can be two forms expressing similar 'content.' With mirror neurons the mapping process is narrowed down to a single cell integration of perceptual and motor representations of the 'same' action. Thus, the idea is that the effective action type is the content that gets 'translated,' and the metaphors implied by these terms seem to evoke a model of these neurons as little machines, which are fed pictures of actions and output the movements of these actions.

Interlude: The sensorimotor notion

I pause here for a second to make explicit how I use the notions of sensorimotor integration and sensorimotor neurons. Firstly, we must abandon the tempting but meaningless picture of single cells with a perceptual input and motor output. I think it is pretty obvious on a physiological level that individual cortical neurons with their complex web of projections and synapses do not have these sorts of concretely reified and pure sensory inputs and motor outputs. This would only happen if there exactly was an *encapsulated one-way route* from the perceptual receptor areas to these few 'interface' neurons that then again projected in a similar encapsulated one-way fashion to cortico-spinal motor neurons. Though this idea might be obviously wrong, the difficulty and the reason why the metaphor of this interface persists in the face of disqualifying evidence might be that it is not so easy to come up with an alternative notion or metaphor of what it means for a neuron to be sensorimotor. I shall, like many scientists working in the area of sensorimotor integration like e.g. the Parma group, classify something as a sensorimotor neuron if its activity is systematically correlated with (and thus thought to be modulated by) both perceptual phenomena and executed actions. I say systematically because I think that the modulation is not necessarily linked to *specific stimuli*, but could be a correlation based on relational, or more abstract features of the environment, the action goal etc. Further, there is the possibility that the systematic feature derived from the overt perceptual and action engagements could carry over and be used for more abstract cognitive purposes.[282] Given a focus on systematic modulation there is no need for individual projections between neurons to be labeled either sensory or motor, but rather one would expect also projections in such a complex web of sensorimotor neurons to be both sensory and motor. Accordingly, I see sensorimotor neurons as a very broad category encompassing large amounts of cortical neurons with very heterogeneous properties. Some might argue that I make the category so broad that claims about sensorimotor grounded cognition become if not tautological then at least very watered down. In one way it is true that I want to significantly broaden the category and the cognitive reach of the sensorimotor realm to go beyond the traditional notion of what is strictly sensory or motor. But, on the other hand, I think that the systematic modulation

[282] Thus on my definition of sensorimotor processes one can image higher cognitive processes as grounded in and exploiting the more basis systematic sensorimotor organization for less concrete purposes.

demand still upholds the link to the overt perceptual and action engagements and thus warrants the term sensorimotor – as opposed to the idea of, for example, neurons that are modulated by purely pictorial perceptual features.

The point is exactly to stress that if one wants a definition that captures mirror neurons and other fronto-parietal neurons as sensorimotor, then given the rather abstract response modulation of such neurons one *needs* a broad definition of what counts as sensorimotor. More specifically a notion is needed that does not depend on reified uniquely sensory inputs and motor outputs. Thus, the category must exactly be so broad that it also somewhat blurs the distinction between basic sensorimotor processes that are directly modulated by external sensory and overt motor phenomena and higher cognitive processes in which the connection is much more complex and less reifiable. Or to put it differently, if one adopts a narrow notion of sensorimotor activity as closely linked to input and outputs then one might not empirically find any cortical sensorimotor processes or at least limit these to a very few primary sensory and motor areas. Most importantly under such a definition the fronto-parietal integration would fall out of the sensorimotor category, which I think would not help our overall understanding of these areas or mental processes but would instead just introduce another empirically mysterious black box of mental functioning. In sum, as I see it, the choice seems to be between either dropping the notion of advanced sensorimotor cognition or admitting that such sensorimotor cognition might be involved in cognitive tasks well beyond concrete actions and concrete perceptions. I of cause advocate the latter and challenge anyone advocating the former to present a new positive theory of fronto-parietal circuits that explains the basic role of these areas in goal directed actions that does not make use of the idea of sensorimotor integration.[283]

[283] Further, one would need an explanation of why these areas are active in social perception and possibly linguistic tasks, short-term memory and the perception of non-biological movement. One might expect a story about semantic action meaning (Jacob, 2008), but then a new mysterious translational interface is created between basic sensory and motor processes and higher semantic cognition. I have already hinted at the problems of such a theory in simply explaining how perception and action are possible. On my view sensorimotor processes naturally work at some level of pragmatically meaningful and anticipatory abstraction form overt actions and perception as part of their primary function. Therefore the continuity between low-level and high-level cognition becomes one of degree of involvement of memory and other 'offline' processes involved in creating, monitoring and anticipating hypothetical imaginary affordance structures.

Returning now to the sensorimotor 'interface' metaphor, the idea implied by this metaphor is that we have two parallel representational systems for action and perception and thus the content of the representations is 'understood' or at least categorized independently on both sides of this sensorimotor translation point.[284] In other words, the perceptual system and the motor system have each their independent recognition and classification system of the various contents, i.e. for example in the typical mirror neuron the content would maybe be the 'grasping' action goal. As argued above, the problem is that this dualism makes the generalization and classification process equally obscure and mysterious on each side of the sensory and motor divide. Bergson, in my eyes, contributes an important discussion on the age-old riddle of how we can generalize from individual cases, i.e. on how we can go from unique perceptions to general concepts. His key insight is that he takes a biological/evolutionary approach and by way of that turns the question upside down. He argues that it is a mistake to think that we necessarily understand the particular before the general. He suggests that biological organisms by way of their limited set of responses to infinitely new environmental exposures in their very sensorimotor engagements are automatically generalizing.[285] His point is that we by way of sensorimotor responses perceive the general categories first, and then only on later evolutionary stages with the help of an extremely developed memory capacity learn to distinguish our perceptions from each other and understand them in their unique and particular aspects. The point in regard to the present discussion is that by separating sensory and motor representations one might make a mystery out of the functional meaningful categorization processes on both sides. Logically everything resembles

[284] The term 'sensorimotor' is obviously somewhat problematic in itself as it also tacitly carries the idea of such a singular point of integration between two independently reified streams of information. Looking at the complex web of referent connections even just within cortical areas normally seen as perceptual and motor, it is not clear how a simple sensory-motor interface is supposed to be understood. The problem is that each 'sensorimotor' neuron or local population of neurons forms synaptic connections with a host of other such 'sensorimotor' neurons. Thus, beyond the sensory receptors and spino-cortical motor neurons it does not seem straightforward of establishing how any inputs or outputs are uniquely sensory or motor. Thus, 'sensorimotor' neurons or areas rarely integrate purely sensory information with direct motor muscle commands. Their division of labor seems to be within an 'always already' inherently sensorimotor mélange. Hence, it is difficult for me to see how the meaningful classifications and organization of effective stimuli for such individual neurons would rely on purely motor or purely perceptual 'representations' or categories. The fault might here be in my powers of imagination as I admittedly am only barely treading water. However, the key point is that regardless of how the brain turns out to organize itself one cannot simply assume specific points of translation between meaningful purely sensory inputs and purely motor outputs.
[285] Bergson (1896) p.202-204.

everything and sameness must therefore either be limited to identity, which is biologically impossible, or admitted to be a very slippery slope. The suggestion here is that it is by way of sensorimotor integrations and perceptually induced resonance of such stored connections that the meaningful categorization first happens. To put it simply, the generalization comes from the fact that animals always covertly respond – via multiple hierarchical generalizing sensorimotor activations - and that it is via these responses that we 'recognize' or categorize the perception. Further, if that is the case, one might suggest that what we would call 'pure motor' and 'pure perceptual' representations should be seen as secondary to a common sensorimotor (and emotional and memory based) provenance and organization.

D. <u>Summing up the sensorimotor representation issue</u>

I am aware that this is a too short, speculative, and somewhat mangled story, but for present purposes I do not aim to present a full positive account but merely argue that the idea of purely motor or purely perceptual representations might be trickier than sometimes assumed. I have pointed out that the neat distinction and the idea of mapping and translation might obscure the role of sensorimotor integration in the very organization and constitution of representations. Further, I have tried to suggest that in the absence of purely motor and purely perceptual representations, it seems that we need a more fundamental re-conceptualization of mental representation, and that an inherently sensorimotor category might be helpful. I have argued that there might be certain empirical and theoretical advantages of utilizing this sensorimotor category to think of some mental representations as 're-presentations' of relations between animal and environment rather than mentally representing non-mental content. In other words, to see the 'content' of such mental representations as relational – as relating emotional drives, behavioral goals and perceptual pre-cursors and consequences. Such mental re-presentations are sub-personal, 'private' and unique to the individual agent, but also to some extent shared with other agents with similar experiences and drives. Further, re-presentations get engaged by the shared present social and physical context and I want to suggest that they therefore are partially accessible to others that simply perceive the agent's engagements. These sorts of mental re-presentations are exactly inherently present in observable contextual behavior and also at all times temporally transcending the present and overtly observable.

5.2.4. A sensorimotor grounding of intentional action?

In extension of the organizational point made above, I want to also suggest that sensorimotor integrations might play various essential roles in making intentional action possible. Firstly – as suggested already by the ideomotor theory of Lotze and James[286] - the organization and representations of intentional goals might very well depend on learned sensorimotor associations between actions and perceptual change. Secondly, it might be that our perceptually evoked sensorimotor affordance responses help create and shape our action intentions.

It is very informative to notice that both Jeannerod and Jacob's thoughts seem to be shaped very much in contrast to behaviorism, which they appear to see ALL sensorimotor-based theories of cognition as being instances of. In the 1980's Jeannerod published a book called Cerveau-Machine, which at its core is a defense of the information-processing model of the mind and its strong reliance on mental representations as the content being processed.[287] Jeannerod wrote this book already a quarter of a century ago in 1983 and the worlds of cognitive science, neurology and philosophy admittedly looked very different then. This being said, it seems that his later writings and thoughts still are conceived through the prism of these earlier ideas of how one should conceptualize the workings of the brain in general and – as the book's subtitle suggests - voluntary action in particular. The issue of voluntary action proves to be essential to Jeannerod's negative view of sensorimotor theories and is an integral part of the classical cognitivist framework in which he interprets mirror neuron research. Conversely, to me voluntary action seems to depend on sensorimotor integration. I suggest that mirror neurons play a role in monitoring the surrounding affordance structure, and that it is exactly the way we 'always already' are localized in a field of affordances that makes voluntary and intentional action choices possible. I do not have the space to discuss the literature on and the competing theories of voluntary action, but in regard to Jeannerod it is obvious that his main argument against sensorimotor theories

[286] See references and discussion of ideomotor action in Chapter 3.5.2 B.
[287] Jeannerod (1983).

of cognition is that he thinks that these theories are behaviorist and he argues that a sensorimotor idea of action representation cannot account for voluntary action.[288] Very roughly, the argument runs along the lines that if motor outputs are determined by sensory inputs then endogenous action is not possible. On the contrary, evidence shows that we can generate actions from internal representations without perceptual guidance and hence the sensorimotor account of action initiation must be wrong. However, the problem with the logic here is that sensorimotor integration only leads to determinism or behaviorism if one ignores what I see as the fact of multiple simultaneous parallel integrations. As I have already introduced it in Chapter 2.2, Erik Rietveld's theory of skillful unreflective intentional action hypothesizes action selection as depending on our being in a 'field of affordances.' Given such a story one can suggest almost the exact opposite of what Jeannerod concludes. Namely, that the sensorimotor integrations are precisely what make informed voluntary action choices possible. As mentioned, I sympathize with the ideomotor line of thought that intentional actions depend on learned sensorimotor action-consequences connections, which can be turned into actual movements via contextual specification. I suggest that the goal is stored as a sensorimotor representation, and that in order to become actually overt it needs sensory guidance, but that this guidance in endogenously generated actions might be internally imagined or 'simulated.' Thus, in regard to the argument that sensorimotor theories cannot account for endogenous action generation, it must be noted that sensorimotor theories do not have to claim that *present* perceptual guidance is necessary for each singular endogenous action initiation, but simply that some perceptual guidance must figure at some stage in the learning of actions and in maintaining their accuracy over time. There is here a temporal dimension that is often overlooked. Another dimension that is mostly ignored is the difference between actual perceptual guidance and specification via an imagined or simulated sensory context. I would say that both these functions might rely on sensorimotor processes – the point is that there does not seem to be a 'purely motor' endogenously generated action initiation.

One further point pertains to the role of sensorimotor guidance in action choice and the generation of intentions. Some might say that the action intention is internally generated and that only the action specification – and thus choice of action means - is supported by

[288] See Jeannerod (2006). p.8.

our continuous sensorimotor engagement with the presently perceived affordance structure. However, the point that Rietveld makes is that also the overall informed action choices themselves are generally made in a 'field of affordances.' It is exactly the evaluation of the available action invitations that makes a given action choice appropriate at a given moment, and one might therefore suggest that intentional choices normally arise in the emotionally invested and active sensorimotor engagement. The current engagements are thus evaluated and chosen in regard to the horizon of sensorimotor affordances but also calibrated to more abstract goals and motivations, such that the here/now is judged based on a broad and rather underdetermined field of affordances in regard to the broader spectrum of endogenous motivations and their respective psychological 'pull.' The overall points are that intentions are easier to understand if they are *not* endogenously created in a Cartesian vacuum without external influence, and further that if one acknowledges the multiplicity of affordances this does not lead to behaviorism or simple determinism. Ultimately, it seems extraordinarily difficult to make sense of voluntary action if one simply posits the need for an endogenous pineal gland-like point of intention creation and choice.

The presented opposing interpretations of the relationship between sensorimotor integration and voluntary action might thus be rooted in assumptions about 1) the possibility of many simultaneous sensorimotor integrations, 2) how they are informed by memory and 3) the possibility that sensorimotor connections can be intentionally evoked. In his 1983 book Jeannerod attacks Bergson's theory of sensorimotor based cognition and claims precisely that it is behaviorist and unable to account for voluntary choice. He focuses on a passage where Bergson describes the brain's activity as a telephone switchboard where perceptions are connected to actions.[289] I share Jeannerod's discontent with Bergson's terribly ill-conceived switchboard metaphor, which exactly suggests a 'one at the time' sensorimotor integration process done by a Cartesian homunculus rather than multiple simultaneous connections as the basis of the overt action choice process.[290]

[289] See pp.113-115 in English version: Jeannerod, M. 1985. *The brain machine*. Harvard University Press.

[290] Bergson at multiple other junctures certainly presumes simultaneous sensorimotor integrations, so it could be that the telephone switchboard is simply a terrible choice of metaphor. He for example writes that all perceptions are prolonged into some nascent movement and that different movement tendencies compete etc. But be that as it may, the point is that *even if* Bergson had an ill-conceived notion of sensorimotor integration, Jeannerod's critique of sensorimotor theories *in general* as unable to account for voluntary action rests on an idea of sensorimotor integration as a simple behaviorist association between

However, Jeannerod's discussion totally ignores the fact that according to Bergson each actual sensorimotor representation is grounded in past engagements and infused with memory that transcends the present sensorimotor engagement, and he precisely thinks of the contribution of memory as the 'mark of the mental' (if one is allowed to use such anachronistic terminology).[291] But, Jeannerod simply disregards the Bergsonian notion of memory and the contribution thereof. I shall not here attempt to defend the Bergsonian idea of memory, but will just note that it is easy to find a theory – like, say, a sensorimotor-based theory of cognition - lacking if one leaves out its very core. Jeannerod clearly reads Bergson, James as well as Gibson as pure behaviorists, as denying any kind of mental intermediary between stimulus and response. It should be clear by now that I do not share this reading at all. As I have argued in the previous section, it is not that sensorimotor theories necessarily deny the existence of representations. In the case of Bergson for example, the difference is that the central representations are thought of not as internal reconstructions of the external world, but as grounded in pragmatic sensorimotor connections that re-present not simply objective or un-interpreted states of affairs but relations between organism and world.

In Jeannerod's own view the processing of action representations is essential to an understanding of the mental, and he further he supports the idea that such representations must be placed between perception and action execution in an adequate story of voluntary action. The logic seems to be that there must be a choice before an action representation is engaged. This is an extremely important theoretical point in regard to the issue of social cognition as mindreading introduced in the previous chapter, because this dichotomy between 'centrally-coded' intentions and motor action seems to provide a foundation for the neat dualism between observable behavior and inner hidden mental states. Jeannerod's analysis of voluntary action depends on a theoretical framework that juxtaposes peripherally and centrally initiated actions. He argues that centrally executed actions are organized in a way such that their individual execution is independent of

---

stimulus and response that ignores the cornerstones of Bergsonian and Jamesian psychology, namely, memory and action choice.

[291] To be exact one would here need to get into Bergson's distinction between two types of memory roughly procedural and semantic/episodic and links consciousness and the 'truly mental' to the latter – but sees all actual mental processes as involving a combination of both kinds of memory.

perceptual and proprioceptive feedback. I think that he might be right that the performance of action tokens is independent of present sensory feedback, but that of course does not have to mean that the organization and continuous shaping of such actions do not involve sensorimotor connections, i.e. that sensory 're-presentations' or 'simulations' on a neurological level are essential for 'centrally initiated' actions.[292] I argue that a 'pure motor' idea of action representation runs counter to both the findings of the Parma lab and Jeannerod's own awareness and insistence on the importance of the 'goal' concept in cortical action organization. In accordance with ideomotor predictions, representations of goals seem to depend on prior action performance and perception, and on the experienced relation and evaluation of perceptual antecedents and consequences. Goal organized actions seem to depend on causal learning of sensorimotor connections and more abstractly of likely action consequences.[293] And if it is the case that our representations and understanding of goals and therefore also our intentions of reaching such goals are grounded in sensorimotor representations of learned relations between actions and perceptual change at various levels of abstraction, then, accordingly, action intentions do not seem per definition to exclude sensorimotor simulations but exactly seem to be grounded in these. This idea of a sensorimotor grounding of intentions is important to my view of basic social cognitive processes and of how we understand others as intentional and mindful agents. In other words, I think that intentions might often transcend our present 'token' engagement with a concrete context, but this does not mean that mental representations or intentions as such exist in a different hidden realm than overt engagements. Rather, it is simply that there is a temporal transcendence such that intentions can relate to goals, perceptions and actions well beyond the concrete present. In regard to the role of mirror neurons and other sensorimotor neurons in social

---

[292] Jeannerod concludes that action representations are centrally coded and independent of sensory feedback, and then without further ado he moves on to talk about various theories of motor representations and 'motor programs' that define motor organization within purely kinematic frames of reference. (Jeannerod (2006) p.8 and (1985) p.112) Hence he implicitly suggests that because sensory feedback and guidance is not necessary for individual token actions it follows that actions are organized and initiated independently of sensorimotor integrative brain processes. The latter of course does not follow.

[293] Voluntary goal directed action seems to depend on the learning of causal relationships since one can only act towards a goal that one has some idea of how to obtain. Thus, to come full circle, the existence of voluntary action in absence of present sensory information does not prove that motor representations do not depend on sensory connections. Rather, insofar as the voluntary actions have perceivable goals they seem very likely to depend on sensory connections and 're-presentations.' This idea and the role of sensory imagery in endogenous and non-perceptual processes can also be likened with Jesse Prinz's contemporary reinterpretation of concept empiricism (Prinz 2002, 2005).

cognition, this idea is important as these sensorimotor integrations do indeed both show themselves in and yet transcend the present action and perception 'tokens.' Thus, one's sensorimotor resonances during action perception could indeed inform our understanding of the intentional choices of others. But, these sensorimotor representations are sub-personal and if consciously perceived only an aspect of an overall experience. Accordingly I do not see mirror neuron activity as guaranteeing any shared experiential first personal experience or '3rd via 1st person' mindreading.

Jeannerod himself refers to both Bastian and James who both thought of action initiation as based on sensory memory traces or re-presentations, but he falls short of endorsing the view that sensorimotor integration is essential to action representations. His focus is on the idea of a basic motor schema, and it seems that he thinks that this schema must be purely motor for it to be activated in the absence of perceptual guidance. However, Jeannerod rightly cautions that even though some motor organization is learned and stored for future use independent of sensory feedback, these schemas should not be thought as static or fully specified: "Schemas should be plastic rather than fixed, in order to adapt the movements to the conditions of each single action."[294] In this way, he allows that actions are in normal conditions specified by the proprioceptive and perceptual context in which they are performed. A tension should be noted here in regard to the issue of whether 'covert actions' should be thought of simply as 'actions minus execution.' The idea of action representations as plastic and unspecified schemas clearly suggests that he distinguishes between such a motor schema and an actual specified action or simulation. We have already seen that Jeannerod in other contexts vehemently argues that covert actions in general are under the same constraints as overt actions, which I think runs counter to this idea of a flexible motor schema. This leads to the problematic notion of motor cognition as full simulation, which we shall see that Jacob and Jeannerod in their discussion of action mirroring 'take very seriously' and is also the notion they accuse Gallese and other mirror neuron researchers of 'tinkering' with.

5.2.5. Summary

Against the backdrop of this discussion of Jeannerod's explicit and implicit views on motor cognition I think it will be easier to understand why he and Jacob insist that action

[294] Jeannerod (2006), p. 12.

mirroring cannot yield goal or intention understanding. According to Jeannerod motor cognition is strictly about actions, whether overtly or covertly produced. A motor representation relates goals uniquely to motor commands not to perceptual situations. Action simulations are just like overt action just without the execution, thus if motor cognition is limited to such full simulations then we can only use motor cognition to reproduce actions already in our motor repertoire. Further, when covert actions are thought of as fully specified then the motor system cannot represent more than one action at a time. Hence the argument against the role of action mirroring in social cognition rides exactly on these self-imposed theoretical limitations of motor cognition. It is argued that if mirroring played a role in action understanding, then social interaction and understanding of simultaneous actions of multiple others would be impossible. Furthermore, examples of action understanding beyond our motor repertoire are used to prove the inessential nature of action mirroring in this understanding.

Hopefully my lengthy discussion of Jeannerod's more general views has shown that these arguments rest on a notion of motor cognition which might be too narrow to accommodate the actual mirror neuron and other sensorimotor findings. Further, I have discussed how the narrow idea of motor cognition and the dualism between sensory and motor representations push the organization and representation of goals and intentions away from the area of sensorimotor integration. It is also essential to Jacob and Jeannerod's view that true intentional mental states cannot be found in the observable actions but are almost per definition hidden behind these actions. I have here tried to show not only why I disagree with their view of the realm and function of motor cognition, but also questioned the consequences of their view of intentions and mental states based on a radical distinction between actions and intentions. I then presented an alternative story of the role of sensorimotor integration in intentional action choice. The disagreement over how to interpret sensorimotor integration will be pivotal in regards to the questions of the nature of social cognition and how we understand ourselves and others as minded creatures. I have already in regard to the Gallese and Goldman simulation theory of mirror neurons criticized the problematic idea of mental states as ultimately severed from and hidden behind actions, and the corresponding view of social cognition as mindreading, i.e. as attribution of such hidden states to others. Given my discussion here of Jeannerod's view of how intentional actions are centrally coded and initiated, I hope to have shown how the narrow view of motor cognition as covert specified actions links up with the view of social cognition as mindreading of hidden states. It is these common assumptions of motor and social cognition that drive Jacob and Jeannerod's critique of mirroring and 'motor theories of action understanding.'

5.3. Jacob and Jeannerod – mirroring is not mindreading

In 2005 Jeannerod co-authored an article with Pierre Jacob offering an alternative interpretation of mirror neurons based on their idea of action simulation.[295] Jacob is a French philosopher, but his work in cognitive science is follows the Anglo-Saxon tradition in the philosophy of mind. He and Jeannerod have collaborated on several books and articles and from where I stand they seem to a large extent to share the theoretical frameworks and convictions that I discussed in the previous section.[296] In what follows I will take a closer look at Jacob and Jeannerod's influential article entitled 'The motor theory of social cognition: a critique.'

5.3.1. Motor theories of cognition are simulation theories

The focus of Jacob and Jeannerod's critique is, as the title indicates, the theory that motor cognitive processes underlie social cognition. The motor theory in question is more

[295] Jacob & Jeannerod (2005).

[296] Lately, Jacob has published two other critical discussions of mirror neuron research and mirroring simulation theories that I have already referred to in previous chapters (Jacob 2008 & 2009). The arguments in these articles seem to follow largely the same lines of critique as pursued in the Jacob and Jeannerod article combined with some of the ideas and issues that Csibra takes on and which I will discuss in greater detail in the next section. Jacob has earlier written about intentionality, visual perception, meaning, and also an early book on logical positivism and the Vienna circle. In regard to the mirror neuron debate, his focus is primarily based on the mind reading question – and he seems to equate the mirroring hypothesis with Goldman's simulation theory of mindreading. We have already seen how Gallese and Goldman's attempt to join forces was burdened by problematic assumptions from the very beginning and we saw that in the end it was not even really proposed that mirror neurons instantiate a simulative mindreading process but rather that they could be the evolutionary 'precursor' in that they produce a similarity between agent and perceiver. Jacob is very critical of what he calls the 'tuning-fork model of social cognition,' which he takes to mean that there is a neuronal matching of the brain activity that causes the overt action in the agents brain and the simulative activity in the perceiver's brain. He does not see a physiological mechanism that can support such an exact brain mirror matching. I agree that a precise neurological equivalence is empirically unfounded. However, his argument is based on the assumption of a radical distinction between overt actions and hidden mental states and thus that the prior intention per definition is underdetermined by action observation. I argue that this strict inner/outer distinction between minds as hidden states versus overt behavior is problematic and ignores the possibility of a less neat distinction based on the temporal transcendence between intentions/ mental representations and active overt engagements. In short, the reason why I do not devote more space to the discussion of Jacob's later arguments is that I think that the key elements are already brought out by my discussion of Csibra and of the Jacob and Jeannerod article.

specifically mirror neuron based action simulation/mirroring theories, but Jacob and Jeannerod purposefully stress that they are offering a critique of a 'motor theory.' Their focus is on 'the motor theory of social cognition' but they tacitly imply that their arguments might have a wider reach in regard to others of the 'ubiquitous' embodied/enactive motor theories of various cognitive processes. In this way, what seems to be at stake is not just pertaining to mirror neurons and social cognition, but something more general about what motor cognition is and what it can and cannot do and contribute to general cognitive functions. I have been arguing that Jeannerod's concept of motor cognition as simulation is too narrow and, if that is the case, one can already see why he thinks it plays a different and particularly a lesser role in other cognitive processes than some other researchers and thinkers have proposed.

A general problem for Jacob and Jeannerod's narrow account of motor cognition is that it is seems incapable of explaining many robustly replicated findings of motor activation. I shall point to three main types of findings that I see as undercutting the empirical plausibility of this narrow account: firstly, the simultaneous activation of multiple sensorimotor integrations, for example in the perception of multiple object affordances and secondly, the level of abstraction in fronto-parietal sensorimotor integration away from specific kinetic motor commands. And thirdly, I point to findings of sensorimotor activation in connection with non-motor repertoire events or actions, for example Ricarda Schubotz's interesting findings regarding the role of motor areas in the prediction of not only non-biological actions but also non- action events.[297] Thus, the first thing to note is that Jacob and Jeannerod's arguments against the importance of motor cognition in various cognitive functions rely on a highly questionable idea of motor cognition as 'covert action minus execution.' It seems that many processes that cannot easily be mapped onto specific movement commands do regardless have a motor aspect and rely on sensorimotor integration. Given the Parma Lab findings discussed in Chapter 2 that many mirror neurons and other premotor-parietal modulations seem to depend on the observed or implied action goal rather than specific movements, it is likely that these neurons should not be seen as simulating specific movements but rather as integrating

[297] See Schubotz (2007) and Schubotz and von Cramon (2004). I shall discuss these findings later in this chapter.

sensorimotor information at a more abstract teleological level. With this big ‘but’ in mind let’s turn to Jacob and Jeannerod’s critique of the role of mirror neurons and motor cognition in social cognitive processes.

5.3.2. The gap between action simulation & mind-reading

Jacob and Jeannerod’s point of departure is the observation that there is a gap between the perception of overt action and the central social cognitive skill of ‘mindreading.’ Mindreading is exactly explained as involving attribution of unobservable mental states to someone:

> …healthy human adults readily explain and predict human actions by representing and attributing to human agents a whole battery of internal unobservable mental states such as goals, intentions, emotions, perceptions, desires, beliefs, many of which are far removed from any observable behavior. (Jacob & Jeannerod 2005, p.21)

Thus, the gap lies in the nature of the action as overt and observable as opposed to mental states that are unobservable and ‘far removed’ from observable actions. One might worry that Jacob and Jeannerod with this description move mental states too far out of reach. I see this notion of mental states as ‘far removed from any observable behavior’ as problematic in that it ignores the possibility that mental processes in their constitution and structure generally depend on the relation between agent and environment as it is shaped over time. Consequently, I think they are right in pointing out that mental representations and other goings-on often transcend actual or present overt action production and perception. But, they might be making a mistake in severing the link to overt engagements as such. Jacob and Jeannerod do not pay such temporal and constituting ties to overt engagements any mind as they construe their argument on the neat dichotomy between overt action and hidden mental states on the one hand and the idea of motor cognition as covert simulations of such overt actions on the other. They argue that such action simulations - i.e. covert re-enactments of overt actions - cannot bridge the inherent gap between actions and mental states.

> Faced with this challenge, the strategy favored by motor theorists of social cognition is to tinker with the concept of motor simulation, as suggested by simulation theorists of mindreading. We disapprove this strategy because it relaxes the fundamental link between simulation and the requirements of the motor system, which we take very seriously. (Jacob & Jeannerod 2005, p.21)

So here it is again: Jacob and Jeannerod take their concept of motor simulation as covert action 'very seriously.' And given their definition of such covert action as simply 'actions minus execution' they see all motor cognition as tightly linked to overt movement constraints. This view entails that mirror neurons – which have a motor aspect - are action simulators that do not give us much in terms of social cognition. My point is that given all these wild assumptions I can draw the logical conclusion that mirror neurons play little role in our understanding of others. But a valid argument does not yield sound reasoning unless the premises are true, and in this case I have tried to point out that the assumptions regarding the nature of motor cognition and social cognition seem rather empirically implausible. I have suggested that the notion of action simulation as equated to fully specified actions the load constrains that these have loses sight of the fact that this is only a small part of what there is to say about motor functions in general and in regard to mirror neurons in specific. Thus, in my view the door is open for a social function of mirror neurons and sensorimotor integration that precisely does not involve fully specified action simulation.

A further point of contention is what social cognition primarily consists in, and whether the idea of mindreading as a process of mental state attribution is a useful description of the basic social abilities of humans. At the most basic level the question is why we should think of minds and other minds as hidden *behind* behavior (and as transparent to introspection)? Why assume an opposition of kind, rather than trying to see the intimate interrelation between mind and behavior? Mindreading theories generally start out with a dualistic notion of minds as opposed to overt behavior. I argue that both basic phenomenology[298] and neurological evidence from sensorimotor integrations such as mirror neurons and their broader context suggest otherwise. The point is not to say that all 'mentalizing' is about overt action or in other ways to reinvent behaviorism, but

[298] After all, only agents *seem* to be minded! This is not meant to deny the *conceivability* of extreme sci-fi or borderline cases of for example conscious experiences in apparently comatose people. My point is rather to state the obvious fact that these are exactly borderline cases, i.e. it is a fact that a total loss of agency undeniably does question the *seeming* mindedness of others. And when we talk about 'mind-reading' and social cognition *appearances* are more important than *conceivability* in giving us a foothold on the *actual* psychological processes by which we actually do understand others. To put it differently, the fact that we also sometimes do fail to understand each other is hardly reason for bringing in these cases as a foundation for our otherwise well functioning social abilities. As I shall also argue in the next chapter, sometimes the attempt to disassociate 'mentalizing' and behavior suggests that we have special para-psychological or telepathic abilities. But we exactly do not – and no amount of TOM modules can make us understand or 'read' minds without some behavioral 'text' or other to read them from.

simply to point to the way our mental life is rooted in engagements with others and the world[299]. Accordingly, I suggest that it is counterproductive to start by theorizing how we can think and reason about our own and others' mental states under relative non-engaged conditions. It seems that we must get a better grasp on the nature of mental aspects and their relation to behavior before we can understand – or understand how we can understand – such mental goings-on in abstraction from behavior. Thus, in presenting my alternative interpretation of mirror neurons and simulative 3rd person mindreading I underline the need for an understanding of basic 2nd person social interactions and the cognitive processes underlying the way we understand others *within our own perspective*. I argue that the typical focus on mental state reasoning in notoriously tricky situations of deception and false belief might overshadow the more rudimentary social cognitive skills that we apply under more cooperative and less Martian circumstances.

In general, Jacob, Jeannerod and Csibra all seem to follow the typical approach of classical cognitive science; i.e. they isolate an advanced skill and then hypothesize an often very modular, rationalistic and non-biological underlying mechanism. In the case of mindreading the hypothesized cognitive apparatus is a 'Theory of Mind' ability or module.' Like in the case of Chomskian theories of innate grammar modules – my view is that just positing a cognitive module or a neurological location, or specific gene for a complex human ability, contributes very little of value to our understanding of the actual processes and developments underlying the cognitive ability in question. As a matter of fact these modules seem to be black boxes postulated in relative abstraction from our overall understanding of the mind/brain. My claim is that more attention must be paid to biology, developmental psychology and also actual 2nd person social interactions to understand social cognition in general and the possible role of mirror neurons in specific. I suggest that social cognition is not primarily based on 'mindreading' of hidden mental states. Instead I hold that many social cognitive skills seem to be grounded in an ontogenetically and phylogenetically earlier understanding of others in their relational engagements. These engagements are simultaneously overt and transcendent of the present behavior in that they are engagements of temporal agents whose action choices bear on both past experiences and goals partially inaccessible to present observers. I have

[299] I suggest that we should see mental states as rooted and inherently related to behavior and outward engagements but always temporally and qualitatively transcending the *present* overt behavior.

already in the previous section discussed this Bergsonian notion of minds as transcending overt action due to their temporality rather than a mystic dualism or hidden soul, and how it is essential to my thinking in many ways.

5.3.3. Mirror neurons as only 'weakly social'

But I will push this issue aside for now and start from the angle that Jacob and Jeannerod take. They bring up an enormously important issue that is too rarely even mentioned in the mirror neuron literature. They ask in what sense the firing of mirror neurons actually can be seen as a social cognitive process. Their answer in turn is that "the firing of MNs is a social cognitive process only in a very weak sense."[300] Jacob and Jeannerod start by noting that when mirror neurons fire during action execution, i.e. in the 'expressive mode,' their activity is not at all social. If mirror neurons during action execution are simply part of the process of initiating object directed hand actions their modulation in such cases does not seem inherently social in any way.[301] In the 'receptive mode' the process is social only in the sense that another individual is being perceived.

> When MNs fire in the brain of a monkey watching another grasp a fruit, the discharge is a weakly social process: the two monkeys are not involved in any kind of non-verbal intentional communication. The agent intends to grasp a fruit, not to impart some information to his conspecific. Nor does the observer's understanding of the action require him to understand the agent's communicative intention (because the agent has none). (Jacob & Jeannerod, 2005)

Thus, Jacob and Jeannerod point out that mirror neuron modulations seem independent of social intentions and what I with Merleau-Ponty would call 'frontal engagements,' and that therefore their activity even during action perception is only 'weakly social' in the basic sense of simply involving other people. I think this is an important point, but clearly for slightly different reasons than those suggested by Jacob and Jeannerod.

---

[300] Jacob & Jeannerod (2005)

[301]This claim might seem intuitively true, but I actually do think that there is a question, not yet empirically determined, of whether mirror neuron activity during action *execution* should be seen as simply producing the action. The time-locked character of many strictly congruent mirror neurons might suggest that they could be having an action *monitoring* function also in regard to *one's own actions*. This could be part of broader action coordination,anticipation and learning processes. One consequence of such monitoring, rather than directly initiating activity, is that the difference between 'receptive' and 'expressive' mode use of mirror neurons seems to shrink drastically. Furthermore, if this monitoring role is seen in terms of affordance structures in the social context, maybe there is then a very weak sense in which mirror neuron activity can be somewhat social even when it is a product of one's own actions. Note however that one's own actions do not become a part of the affordance space in the same way as others'. (Various findings suggest that whereas the perception of others' actions facilitate similar actions, the endogenously produced actions attenuate the perceptually induced motor resonance. See Chapter 3)

They remind us that the findings from monkey mirror neurons suggest that the modulation of these neurons primarily depends on object-directed actions. In other words, these neurons seem to track relations between people and objects rather than interactions and relations between people. The 3rd person modulation of mirror neurons is stressed as exactly that - 3rd person in the sense of a non-engaged observation of the other that could have been done through a one-way mirror. As seen in Chapter 2 & 3 mirror neuron activity is most often described, classified and interpreted as an automatic agent-neutral process of action observation-execution matching. And, given this story it seems right to say that their activity pertains to non-engaged 3rd person action observation. Jacob and Jeannerod's critique in the above quote implies that what is social in a 'strong' sense is socially engaged, 2nd personal and in particular pertaining to frontal communicative interactions. I think this sounds reasonable. I have criticized the traditional classifications of mirror neurons in contra-distinction to canonical neurons based simply on whether the effective stimulus is an object or an action, and argued that this distinction totally ignores the possibility that there would be premotor and parietal neurons that would respond to 'action affordances,' i.e. invitations to perform complementary interactions rather than simply mirror responses. The fact that social affordances have been ignored in the mirror neuron research is however not at all the point Jacob and Jeannerod want to make. They are clearly not interested in pointing out the limitations of the mirroring research but rather limitations in the nature of the motor system as such - or maybe in the motor system as they define it.

Their line of argument also points to a tension in their own position in that they themselves see 3rd person mind-reading as the essential skill of social cognition. Thus, the idea of the social as primarily intentional other-directed interactions is not meant to give primacy to the ability of social interaction, as much as simply to point out that social cognition takes other-directed actions as their object – but still from a 3rd person perspective. Thus, according to Jacob and Jeannerod the primary social cognitive skill is 3rd person mindreading and not 2nd person interaction, and the reason that the modulations of mirror neurons are only weakly social is simply because there is not *much to read* in the case of hand-object interactions. In other words, it seems that they are not saying that it is the wrong kind of task but simply that it is too easy, and that the kind of mindreading that involves solving the oblique hidden riddles of communicative intentions is the true test of the mind-reading mechanism.

As mentioned, I think it is right to challenge mirroring theories on the lack of focus on and research of premotor and parietal responses to 2nd person social relations, but that is a very different line of argumentation than whether mirroring can yield 3rd person action understanding. Mirror neuron theorists, exemplified by Gallese and Goldman's joint venture, want to suggest that mirroring can provide a basis for not 'full-blown' but some 'low level'3rd person mind-reading. The claim is that mirroring can ground social cognition by giving automatic 1st personal access to a mirror version of the 3rd person's goal-directed action. Hence, mirroring under this interpretation never directly says much about 2nd person social relations. What is proposed instead is that mirroring can yield a shared action representation and thereby some level of 1st person access to the 3rd person action experience and goal, such that the other is not entirely beyond our comprehension. Mirroring theories do seem to lose sight of the 2nd person – but the point is here precisely that so do Jacob and Jeannerod! They rightly note that the classical mirror neurons are only weakly social in that they are modulated by object rather than social interactions. But they forget that there is also a clear tension between their own insistence, on the one hand, that 3rd person mind-reading is the basic skill of social cognition and, on the other, that truly social situations are 2nd personal, mutually engaged and communicative.

Returning to the actual mirror neuron findings several other issues also spring up. First of all, one should remember that in the case of monkeys, detached action observation for example via video clips is not very effective in modulating mirror neurons, which clearly 'prefer' live action perception.[302] What does this mean? It could be that mirror neurons actually might not primarily respond to totally detached 3rd person action perception, but monitor actions going on in the monkey's own present environment – within its own present affordance structure as I call it.[303] Only recently is research beginning to try to clarify the precise social context dependency of mirror neuron activity. As discussed in

[302] See Gallese et al. (1996) and Shimada & Hiraki (2006).

[303] By the monkey's 'present environment' I do not just mean its peripersonal space, i.e. the area of possible immediate action impact given effectors/tools available. Evidence suggests that most mirror neurons responded to action outside of the monkey's peri-personal space. What I mean by the affordance structure of the present environment includes more distal action possibilities in both space and time. Thus the idea is that it makes pragmatic sense to monitor and predict not only one own immediate motor affordances, but also pragmatic and causal relations between elements in the environment, such that one's own action choices can be predicted and coordinated.

Chapter 3.2.6, Gallese and his coworkers seemed a little too eager to conclude that the responses of mirror neurons were 'highly consistent' and also context independent and did not then present counterexamples to those conclusions. However, in a very recent study by Parma lab researchers, Caggiano and colleagues found that various mirror neurons are modulated relative to actions taking place inside and outside of the monkey's peri-personal space.[304] These are exciting findings as they suggest that the broader mirror neuron circuits are not totally agent-neutral or context independent. The observed goal-directed actions modulate these neurons in ways that carry important information not just about the action as such but about the actions in the context of the broader affordance space from the perspective of the monkey as an agent itself.[305]

Thus, Jacob and Jeannerod suggest that the modulation of mirror neurons is only weakly social in that they do not respond to 'frontal' or mutually engaged social interaction, but only to the observation of others involved in object-directed actions, doing, so to speak, their own thing. Given the standard account of mirror neurons this claim could be justified. However, I argue that the problem is that little research has been done on a possible motor response to socially directed actions. It is true that F5 monkey mirror neurons seem to respond primarily to object-directed actions or more broadly intentional or goal-directed actions. It has for example been shown repeatedly that monkey mirror neurons are not significantly modulated by threatening gestures towards the monkey. Such threats are emotionally powerful and should motivate strong action affordances - even imperatives - and the lack of response to such actions is as we saw in Chapter 2 taken as strong evidence that mirror neurons are responding to goal-directed hand and mouth action kinds independently of social context. A great problem in regard to this conclusion is that beyond the threatening gestures not many studies have been done on mirror neurons or other visuomotor neurons in response to 2nd rather than 3rd person interactions. And the lack of F5 response to threat gestures might be because this *specific area* mostly pertains to goal-directed hand/mouth actions. The response afforded by threat gestures might be to flee or to reciprocate the violence. In other words, these sorts of non-goal directed intransitive emotional action responses might simply not be under

[304] See the discussion in Chapter 3.2.3 and Caggiano et al. (2009).
[305] For other studies that question the context independence and agent-neutrality of mirror neuron responses see my detailed discussion in Chapter 3.

the F5 jurisdiction. But that of course does not exclude that other areas could contain visuo-motor neurons that respond to the affordances of threatening gestures. And, further, there might be other social affordance actions that do modulate F5 neurons. The conclusion that F5 simply ‘mirror’ hand-object action types independently of social context therefore seems premature. I have already pointed to the recent finding that various mirror neurons are modulated differently depending on whether the observed action is inside or outside the monkeys peri-personal space, i.e. the response depends on the relation between observer and observed rather than simply the kind of actions observed. Further, one could hypothesize that there could be premotor visuo-neurons that would be modulated by socially directed hand/mouth actions. I spent some of section 3.2.5 on how the neat object-action based distinction between canonical neurons and mirror neurons might be problematic. It excludes exactly the possibility of social affordances, i.e. the possibility that some visuomotor neurons respond to actions that afford other complementary actions. This is obviously an important issue in regard to my ‘affordance space’ understanding of the fronto-parietal circuits, but let me now return to Jacob and Jeannerod’s critique of mirroring as the basis for social understanding.

5.3.4. Undercover prior intentions

Jacob and Jeannerod, as mentioned, question whether mirror neuron activity can assist social understanding in cases not just of object-directed actions but also in cases of social interaction. However, their main issue is not the question of social interactions as such, but rather the issue of whether mirror neurons can tell us anything beyond the overt action about potential ‘hidden’ mental states. They propose that social interactions are complicated to understand based simply on action observation as the intentions behind these actions are *further removed* from the overt actions which express them. Their critique focuses on the question of whether mere action mirroring can yield understanding of an agent’s ‘prior intentions.’ They rely on Searle’s distinction between motor intention and the so-called ‘prior intentions.’ Prior intentions are thought of as pertaining to more distal or end goals, and thus as more removed from individual motor acts. In very simple actions the motor intention can be the prior intention, like in the case of simply intending to move one’s finger. The point is that whereas the motor intentions can be seen in the act, the prior intentions are *underdetermined* by the series of observable acts. For the sake of argument they grant that mirroring can yield some 1st person understanding of the 3rd person, namely action mirroring in the observer might give access to what they call the 'motor intention' of the agent, i.e. the local intention to perform the motor act observed. But, Jacob and Jeannerod argue that prior intentions almost as a matter of definition cannot be understood solely on the basis of motor act mirroring.

The argument is built on the core assumption that mirroring is 'motor act simulation,' and given this they argue first that motor mirroring is not sufficient for 'prior intention' understanding and thereafter that it is not necessary. The narrow understanding of motor cognition as simulation then provides the basis for the further positive claim that a 'purely perceptual' social cognition is possible.[306] The following example shows both the strength and the weakness of this line of argument:

> Consider Dr. Jekyll and Mr. Hyde. The former is a renowned surgeon who performs appendectomies on his anesthestized patients. The latter is a dangerous sadist who performs exactly the same hand movements on his nonanesthestized victims. As it turns out, Mr. Hyde is Dr. Jekyll. Suppose that Dr. Watson witnesses both Dr. Jekyll's and Mr. Hyde's actions. Upon perceiving Dr. Jekyll, alias Mr. Hyde, execute the same motor sequence twice, whereby he grasps his scalpel and applies it to the same bodily part of two different persons, presumably the very same MNs produce the same discharge in Dr. Watson's brain. Dr. Jekyll's motor intention is the same as Mr. Hyde's. However, Dr. Jekyll's social intention clearly differs from Mr. Hyde's: whereas Dr. Jekyll intends to improve his patient's medical condition, Mr. Hyde intends to derive pleasure from his victim's agony. (Jacob & Jeannerod, 2005, p. 23)

This is a terrifically bizarre example. Firstly, one should note that the thought experiment is rather difficult to imagine, i.e. that surgically performing a targeted operation on an anaesthetized patient and cutting an alert person for sadistic pleasure could ever involve a series of identical motor actions![307] But given that big concession, the argument is that

[306] "…we shall grant that simulating an agent's movements might be sufficient for understanding his motor intention, but we shall argue that it is not sufficient for understanding the agent's prior intention, his social intention and his communicative intention. Then, we shall argue that motor simulation may not even be necessary for understanding all perceived actions. Finally, we shall argue that a significant part of human social cognition consists of a 'perceptual social' system whose neural basis has perceptual but no motor properties." (p.21-22)

[307] Only a part of the problem lies in it being hard to imagine that the tortured un-anaesthetized victim would lie still (if not Mr. Hyde would have to accommodate his motor acts to the victim's movements). But beyond these issues pertaining to the interaction between Jekyll/Hyde and patient/victim, even if both were anaesthetized and the agony was only meant to appear post-operation, I still don't see how the broader intentional actions would look that same given the two distinct regulatory social/prior intentions of helping versus causing agony? At least one would have to ADD further prior intentions in Mr. Hyde of wanting to cause agony but also of wanting to disguise his actions as an appendectomy – which in the detail of it

we do not understand the prior intentions of *why* someone is cutting another person solely based on the cutting movements. Thus, the example underscores how Jacob and Jeannerod take motor simulation to be an exact covert reproduction not of perceived goal-directed actions in concert with the perceived environment, but as detailed movements and motor commands independently of the goal context. I do think they are right in concluding that such a de-contextualized motor mirroring process could *not* give us 'prior intentions.' But the big problem with this line of argument is the assumption that this exact motor mirroring process is actually what is going on in the premotor-parietal circuits. They suggest that given that Dr. Jekyll and Mr. Hyde perform the same scalpel movements the exact same mirror neurons would fire when Watson observes their respective action sequences. They seem to think that all mirror neurons are modulated simply by the observed movements in total abstraction from the rest of the present scenario and other prior knowledge informing the action perception. However, one should keep in mind that mirror neuron findings from the beginning *never* really supported such a narrow kinematic mirroring process. Rather, the findings have suggested that for example the presence and knowledge of the object and the goal relation indicated by the context of the action is often essential to the modulation.[308] Jacob and Jeannerod are right that exact covert motor action simulation cannot uncover *purposefully* disguised prior intentions. The first problem is that mirror neurons do not seem to simply simulate exact movements and thus one cannot assume that the exact same mirror neurons would fire in Watson's brain in watching respectively Dr. Jekyll and Mr. Hyde. The mirror neurons and overall visuomotor response might not simply depend on Mr. Hyde's ability to replicate a string of surgical cutting movements, but also on whether Watson has other contextual clues - such as for example the lack of concern for the screaming victim - to suggest the presence of sadistic rather than healing action goals. I have earlier discussed Umilta and her colleagues' findings that many mirror neuron activities depend on the perceived and inferred relationship between an action and its target/goal, and that the pantomiming of the action in question is relatively ineffective.

---

would be carried out best if he actually did perform an appendectomy – and of course intended to perform one. Thus, it is just not so straightforward to say that we here have an example of 'same actions-different intentions.'

[308] See discussion of Umilta et al. (2001) in section 3.2.2 and of Fogassi et al. (2005) and Iacoboni (2005) in 4.3.1

The studies by Iacoboni et al. 2005 and Fogassi et al. 2005 provide evidence that many mirror neurons reflect not only the motor actions but are also modulated by contextual information about end goal/intentions. Thus, only if Mr. Hyde miraculously succeeds in *perfectly disguising* his sadistic intentions as appendix operations over time and if the patient/victim is a cooperative Zen-master that can pretend to be anaesthetized, then it does seem reasonable to assume that the mirror neuron response would be exactly the same. But the second problem in this case is that even hardcore mindreading theory theorists will have to grant brilliant actors and Zen-masters that their prior intentions are difficult to decipher! In other words, if the intention is perfectly disguised, then – short of telepathy - there is nothing to 'mind-read' for any hypothesized social cognitive process. Thus, this and other cases of brilliant deception might not do much in terms of arguing that mirror neurons are unimportant to our normal successful instances of social cognition and intention understanding.[309]

However, even if mirror neuron activity observing the doctor and the sadist is not exactly the same, there is of course still a question of whether the sort of goal and immediate action modulation happening in pre-motor and parietal areas can support 'mind-reading', and meaningful reasoning about mental states such as beliefs and desires. Jacob and Jeannerod mention evidence from neuroimaging studies suggesting that reasoning about action intentions and beliefs are supported by different brain areas, and that the action intention understanding is a significantly earlier skill in the developmental process than mind-reading[310]. I have already in section 4.2.6 discussed the classic false belief test that the notorious difficulties it poses to children under 3 1/2-4 years old. I argued in that connection that it is not certain that the children's problems narrowly have to do with social cognition and the understanding of the beliefs of others. Rather, I proposed that an alternative reason could be that small children are unable to decouple from their own perspective and pragmatic affordance structure. Thus, these findings might not be evidence against a role of mirror neurons in social cognition as such. I mentioned, that

[309] In a cases of masterful deception, one might actually suggest that the deception itself is the main prior intentions controlling the overt actions performed – since a good old fashioned sadist presumably would want to *see* the victim's pain…And, maybe one should think of the specific prior intention of attempting to mislead and keep others in the dark about ones true intentions as a rather special case. At least it might be ill advised to case study social deception for the purpose of theorizing social understanding.

[310] They refer here to the following articles: Baron-Cohen (1995). Leslie and Thaiss (1992). Frith and Frith (2003). Saxe et al. (2004).

Gallese and others, who directly engage with the 'mind-reading' literature and the philosophical debate over how we come to understand the mental states of others, never claim that mirror neuron activity is all there is to say about social cognition or even that support 'full-blown' mind reading. Rather, the proposed social cognitive function of mirror neurons is normally described along the lines of a 'low level' process that yields a shared foundation for the understanding of others. Other resent findings support the idea that different social cognitive tasks are supported by different processes, and some have suggested a hybrid theory combining research from the mirror neurons research tradition with evidence for a 'social brain' in more perceptual and medial cortical areas[311]. Overall, it does not seem that many mirroring theorists are trying to say that one can get full-blown mind-reading from fronto-parietal action mirroring. Thus, even if there is a lot hyperbole on the part of proponents of a 'pure motor' theory of social cognition, I do not think that the general conclusion that motor mirroring is not sufficient for all social cognition is very controversial. To me the question is, what *kind* of cognitive and neurological processes can contribute *what exactly* to social cognition, and it is in the failure to pursue the details that I think that both the typical mirror theorists and ToM proponents go wrong. Accordingly, it is clear that Jacob and Jeannerod want to make the stronger argument: Not only is mirroring not sufficient, but it might not even be necessary for many cases of social cognition. Indeed, with the distinction between social intentions and object intentions they imply that mirroring is not necessary for any 'truly' social cases of social cognition. However, as we shall see due to their goal of showing the overt behavior is under-determination by social intentions their examples become rather implausible, unsocial and the truly hidden intentions portrayed appears in need of telepathic interpretations rather than normal humans communicative interpretations.

5.3.5. Communicative intentions and social interactions

Jacob and Jeannerod suggest that mirror neuron activity is only social in a weak sense. They point to the literature where mirror neuron activity indeed seems to be modulated by 3rd person perception of object directed actions. I have argued that it might be

[311] There are various versions of hybrid mind-reading theories. A very interesting article by Gobbini et al. points to the different contributions of for example typical mentalizing areas such as rTPJ and medial frontal areas to the false belief task, and the STS and fronto-parietal areas to intentional understanding of Heider and Simmel style social cartoons with geometrical shapes. Gobbini et al. (2007).

premature to say that it is restricted to either 3rd person disengaged observation or to perception of others' object-directed actions. Given the idea of action mirroring as a full covert action simulation Jacob and Jeannerod assume that we can only simulate or perform one action at the time. And, as mentioned, there is ample evidence from motor imagination, motor interference effect studies and action facilitation studies, that imagined and perceived actions seem to use the same resources as overt actions. However, I argue that only some kinds of action perception lead to such full simulations in the sense of covert action re-enactment. But given Jacob and Jeannerod's idea that all motor cognition is full covert action simulation, we see how it becomes hard for them to understand how we would be able to mirror interactions between more agents, or how we would ourselves interact and respond if we were already busy mirroring the action directed at us. Hence, I want to alert to the fact that irrespective of the empirical findings of mirror neurons responding to object-directed actions, Jacob and Jeannerod have independent *theoretical reasons* to hypothesize that truly social actions must be understood outside of the motor system. In other words, the nature of social interaction as involving more than one person along with the idea that motor cognition is 'one-action-at-a-time full covert re-enactment simulation' makes them think that truly social cognition must take place outside of the motor system.

It is therefore quite natural that they proceed to talk about how we understand socially directed and communicative intentions rather than non-social action intentions. They argue that motor simulations are insufficient in the case of social intentions in general and communicative intentions in particular. Again they give a rather painful example to prove the point. The gist of this example is that at a cocktail party Jill can gesture to her watch with the same overt movements if she wants to non-verbally communicate to her friend Bill that a) she wants to leave or that b) her watch is broken. Well, there is a lot to say here, as the nature of communicative intentions is a big issue in its own right. In way of problematizing Jacob and Jeannerod's example I will allow myself to simply jump directly into my own alternative view.

It seems to me that actions with communicative intentions must take into account what I would call the prior 'shared space of meaning' of the parties of the communication. And further that the agents choose their actions with the expectation that the other/recipient interprets the action given this shared context and prior knowledge. Differently put, for a linguistic or non-verbal communication act to be *appropriately chosen*, it should take one's expectations of the others' knowledge and perspective into account. In Jacob and Jeannerod's example the problem is that using the *same action* of pointing to one's watch with two very different communicative intentions is exactly only plausible as meaningful and appropriate non-verbal action in both these cases if the situations and the 'shared spaces of meaning' *leading up to the action*s are different. Were the prior situations and contexts identical then the *different messages* would have to be reflected in *different actions* of the non-verbal or verbal communication. I am not saying that the whole communicative intention must be in the present action, but I want to suggest that Jacob and Jeannerod's example only achieves a *strict dichotomy* between the overt action and the hidden communicative intention by exactly leaving out the prior context and actions, which make this communicative action meaningful and appropriate in the first place. Communication is not telepathy after all!

Instead of making a strict dichotomy between hidden intentions and overt actions there is a real question about how we via various interactions continuously build and shape a shared communicative space that goes beyond the *present* overt context and overt actions. Thus, in regard to social cognitive communication skills, I think that the real question is not how we understand hidden mental states. Rather, the question to me seems to be what the right tools are for understanding the communicative action of the other as it relates to the particular shared space of meaning, which this communication both builds on and further develops. It might be that action *mirroring* or broader sensorimotor processes are not sufficient or even right for the job, and that the perceptual analysis is assisted by various non-motor emotional, medial and temporal processes. Even if 'folk psychological laws' are what is needed it still seems to me that the example totally misrepresents both the communicative action and its communicative context. The point is that communicative actions are not only acted into a *concrete space* but a *historically*, *culturally* and *socially established space* and without an eye to that contextual space of meaning it is hard for me to see the action as communicative or with social intentions at all. Consequently, Jacob and Jeannerod might be right in saying that individual motor movement simulations do not fully reveal communicative intentions. But I argue that this does not entail that such intentions are 'hidden' or 'far removed' from observable contextual engagements in general.

A further issue is, again, Jacob and Jeannerod's assumption that mirror neurons simply mirror the presently observed action independently of the spatio-temporal context it is acted into. I have pointed to various studies that suggest that many mirror neurons are modulated by contextual and prior information and reflect perceived and inferred relationships between actions, contexts and targets. If that is the case then the Jill and Bill thought-experiment does not settle much. It seems to be an *empirical question* whether mirror neurons and fronto-parietal circuits can differentiate the two communicative intentions and whether these areas play a role in our interpretation, representations and anticipation of communicative intentions as they appear in their communicative and behavioral context. It is an empirical question whether there are motor areas that deal with *communicative actions*, which essentially are actions that *simultaneously* are acted into a concrete space and a linguistic and socially established space of meaning. Broca's area certainly has been implicated as important for appropriate dynamic and creative sequencing of linguistic actions both in regard to production and syntactic comprehension.

In so far as Jacob and Jeannerod want to use the watch gesture example to show that understanding non-verbal communication rides on something else than the simple simulation of present isolated overt motor actions, the point is taken. However, it is very implausible that this sort of simulation exhausts all motor processes or even that it is what mirror neurons do. And I would further argue that if communicative intentions were truly hidden from all observable behavior it is hard to see how any extra-motor cognitive process could find cues to solve the eternal riddle of ambiguity of whether Jill's watch is broken or whether she just wants to get out of here. Thus, as I see it, Jacob and Jeannerod have only succeeded in arguing for the insufficiency of the present observable action, and not for the insufficiency of observable and contextual engagements as such.

5.3.6. The 'one body in the mind' argument

Jacob and Jeannerod move on to try to show that action mirroring is not even necessary for some cases of social cognition. They here refer to the developmental psychological evidence of infants ascribing goals/intentionality to simple social cartoons of geometric shapes moving in relation to each other.[312] I have already mentioned these sorts of Heider and Simmel inspired experiments and will again in the next section as Csibra himself has made such experiments with rather young infants. Whereas Csibra focuses his criticism on the fact that infants do not share bodies or motor repertoires with the geometrical shapes, Jacob and Jeannerod also bring out the issue of how one would mirror an interaction with more than one agent. I mentioned earlier how Jacob and Jeannerod view of the motor simulation as full covert action implies that only one action can be performed or simulated at the time. The idea that motor cognition in general is about action production – be it overt or covert - exemplifies what I would call the 'one body in the mind' view of motor cognition. Motor cognition in general is implied to be under the spatio-temporal constraints of one body acting in one concrete environment at a time. I want to suggest on the contrary that the key to sensorimotor representations and their function in regard to monitoring and predicting affordance structures and simultaneous action possibilities is that stored sensorimotor representations are exactly not constrained in the same way as overt actions. I argue that there is an important difference between actually specified motor actions or mental simulations or re-enactments of such on the one hand and (mental) action representations on the other. Further, I claim that Jacob and Jeannerod's neat distinction between hidden (and non-motor) mental representations and motor simulations is due to a too simplistic view of motor cognition as simply being covert action. The big question to such a view of motor cognition as already fully produced actions in the head is what use these sorts of covert actions are to the process of action selection, planning and initiation. To put it differently, it seems to me that Jacob and Jeannerod have taken the cognition out of motor cognition and ended up with simply actions in the head. As I see it, for motor cognition to play a role in the processes of selecting, planning and orchestrating actions there need to be motor processes that are not yet full actions, and more specifically parallel sensorimotor integrations that make skillful and intentional action choice possible. However, as Jacob and Jeannerod's theory of motor cognition ignores both these theoretical issues and the many empirical findings suggesting parallel and more abstract motor processes, we see how they feed the idea of

[312] I shall repeatedly refer to the studies using the geometrical stimuli protocol based on Heider & Simmel original experiment from 1944 as these problematize the narrow 'one body in the mind' understanding of motor cognition. See Heider and Simmel (1944).

mystical mental representations working – so to speak, behind the scenes - to determine action choice but independently of motor cognition. Thus, we see how they remove the mental not only from the overt behavioral engagements with the environment, but also from internal sensorimotor integrations. Sensorimotor theories are consequently - as we saw in the earlier section of this chapter - thought of as behaviorist and as excluding the possibility of for voluntary action. This logic depends on excluding the possibility that mental representations are sensorimotor based. Now we see that this exclusion is based on the 'one body in the mind' assumption and the thought that sensorimotor integration happens under the same constraints as overt action-perception engagements. Funnily enough, Jeannerod's cognitivist theory, albeit explicitly formulated against traditional behaviorism, now looks like a sort of internalized behaviorism with some additional cognitivist mental representations injected in the middle. In other words, he and Jacob maintain the inflexibility and exclusion of voluntary action in internalized sensorimotor connections that they accused behaviorists of in their stimulus-response pairs. I suggest that we need a different understanding of sensorimotor representations, as working dynamically with multiple simultaneous parallel integrations, shaped by emotion and explicit memory and informing both perceptual interpretations and action selection and orchestration. In brief, we need a new sensorimotor-based understanding of certain mental representations that can account for hypothetical, counter-factual and teleological thinking that goes beyond what is concretely and presently given. We do not need mental representations that are symbolic in a sense unrelated to our current and past sensorimotor experience. Instead, we need to carry the sensorimotor experience of the past into the present perception and action choice. Finally, in relation to the question of prior and communicative intentions as hidden, one can now hopefully see why I claim that they are not hidden from sensorimotor engagements as such but simply transcend the present one.

With this long story in mind let us return to reality – or the scientific reality of mirror neuron research. Often times mirror neuron researchers seem to tacitly assume a view of motor cognition and action simulation that is very similar to the 'one body in the mind' view of Jeannerod. Other times they stay closer to their own findings, which in my view clearly undermines this narrow 'full re-enactment'/'one body in the brain' view of motor cognition.[313] The central problem is that the Parma lab researchers and other mirror neuron theorists do not come out and explicitly profess a different view. This is of course what I am trying to produce with this dissertation. But, given the acceptance of the 'one body in the mind' view, Jacob and Jeannerod's argument that we only can simulate asymmetric object directed actions but not bi-directional social interactions can appear persuasive. They write:

> There is a good reason why the perceptual response to a perceived action directed towards a conspecific would lack motor properties. The inanimate target of an object-oriented action does not act. As a result, the only movements that an observer can automatically match onto his own motor repertoire are the agent's. If, however, the target of a perceived action is a conspecific, then he or she will react. But then, the observer will simply be unable to automatically and simultaneously match onto his own motor repertoire the perceived movements of both agents. Only if he intentionally neglects one of the agent's observed movements will the observer be able to simulate the other's movements. This might be a case of motor simulation, but it is an intentional, not an automatic, process. (Jacob and Jeannerod 2005, p. 25-24)

In this quote we can see how the 'one body in the mind' view of motor cognition leads Jacob and Jeannerod to the idea that social perception should lack motor properties, and further that motor simulation during actual social interactions is based on intentional choice. Given the assumption that motor properties only play a role in full simulation I guess this is the natural conclusion. I of course disagree with this generalized assumption and hence the conclusion. But this does not mean that we do not sometimes use our motor processes in full covert simulations and intentionally so. I do think that in some odd-ball cases we intentionally decide to focus on the actions of only one out of several people, maybe to learn from their movements. Personally I do it sometimes when I watch tennis: I make the choice of not watching the game but only one of the players. It is actually rather difficult and demands special attentional control to ignore the broader affordance space and attentional horizon, i.e. it is tempting to let one's eyes follow the ball. I therefore think that Jacob and Jeannerod are right in claiming that this sort of simulation is based on an intentional choice, i.e. I choose to not focus on the interaction but to engage in a full covert simulation of only one player. I also agree that full covert simulation follows the rules and constraints of overt actions – and can only be done by imagining/perceiving one action in one context. However, the big difference is that I see

[313] I suggested earlier that even Jeannerod's own research and theory of motor schemas are inconsistent with the narrow logic he uses in regard to motor simulation and mirror neuron theories.

this as the rare exceptional case and it seems rather implausible if this sort of simulation process would be the standard motor function during action perception.[314] Based on their ‘one body in the mind’ reasoning Jacob and Jeannerod infer that most action perception is ‘purely perceptual.’ My objection is that the conclusion depends on the exclusion of a variety of motor processes, which do not amount to full simulation. And it seems to me that the very findings of mirror neurons, canonical neurons and other sensorimotor neurons activated in concert by simply perception cannot be explained unless one’s theory allows for more abstract and simultaneous motor activity. As a matter of fact there seems to be rather robust evidence against the claim that all motor cognition consists of full re-enactments or the idea that all motor cognition is constrained in the same way as overt actions.

The full simulation view of motor cognition obviously seems to be too narrow to accommodate many findings of *simultaneous* sensorimotor integration drawn from studies of object affordances.[315] Another big problem for this view is the findings that the premotor mirror areas also seem to be modulated by predictions of external events and non-biological actions,[316] and the very Heider and Simmel cartoons that Jacob and Jeannerod here mention. Based on their assumptions the latter argue that such cartoons cannot be understood via motor processes, since there is no way we could simulate the movements of geometric shapes. The difficulty with this interpretation is that fMRI evidence suggests exactly that the STS and fronto-parietal areas of the ‘mirror neuron system’ are very active during perception of such social but non-biological scenes.[317] This at least should lead Jacob and Jeannerod to revisit their assumptions about the capabilities of motor cognition. And, to me this seems to suggest that various cognitive

[314]As already mentioned, I think that such full simulation helps us create and utilize specific detailed sensorimotor feedback and details that can inform us about the detailed engagements of the other. In other words, the purpose of such simulation is to fine-tune knowledge of affordance and causal properties and not the overall goal or intention of the action. In regard to the tennis example, when I am just following the game I am quite aware what each player is trying to do and where they are directing the ball without engaging in the detailed simulation.

[315] It seems to me that the evidence for parallel processing of action affordances at initial stages is robust and tightly connected to our normal process of visual attention. See for example Tucker & Ellis (1998). Further one might speculate that a lot of clinical problems are due to a lack of normal function in this parallel process. With visual neglect, for instance there is not a normal incorporation of affordances on the neglected side. With anarchic hand syndrome, on the other hand, suggests that affordances of one hand somehow have gotten out of the central control that would normally prevent some of the parallel affordances from overt action.

[316] See for example: Schubotz (2004).

[317] See Gobbini et al. (2007).

processes that do not easily translate to full simulations of actions in one's motor repertoire nonetheless seem to be supported by processes with a more abstract motor aspect. The implicit acceptance of this narrow view of motor cognition therefore appears unwarranted and is a great impediment to the advancement of our understanding of sensorimotor integrations like those observed in mirror neurons.

However, Jacob and Jeannerod do not want to let anybody 'tinker' with the idea of motor simulation and suggest that much social cognition must depend on extra-motor areas like the STS, TPJ, amygdala and medial frontal areas. I think that they are right in pinpointing all these areas as important to many social cognitive functions. However, it is not clear that any proponents of the traditional mirroring hypothesis would be against the importance of these areas for certain social cognitive functions. There are many important issues to discuss in relation to the various contributions of all of the above mentioned brain areas, but the important point here is that contributions from these various areas do not *exclude* a functional role of mirror neurons and other motor and sensory-motor processes in relation to cognition in general and social cognition in particular. Thus, Jacob and Jeannerod's arguments against the importance of mirror neurons in social cognition ride on an in my view ill-conceived logic of motor cognition. We have seen that Jacob and Jeannerod understand mirroring processes as covert action simulation and that they argue that the function of such action simulation processes cannot have much to do with social cognition, but rather support motor learning and action anticipation. Turning now to Csibra's criticism, we shall see that he in many ways shares the views expressed by Jacob and Jeannerod, but he focuses his energy on an analysis of the mirroring process and the understanding and attribution of goals to others.

5.4. Csibra - action emulation and teleological social perception

Gergeley Csibra is a Hungarian developmental psychologist, who is doing fascinating work on the cognitive development of infants. Among many other things, he has studied how infants from very early on naturally understand perceived actions as teleological, i.e. as goal-directed. He has also done very important work on the role of eye gaze and gaze following in regard to how we understand and perceive both objects and also social and communicative intentions. Interestingly, in recent years he has also become a prominent critic of various elements of traditional mirroring theories. I think that his criticism is thought provoking, but it is born out of a notoriously computational and serial view of cognition and motor cognition in particular. Thus, his critique relies on the assumptions about motor cognition as serial input-output simulation processes, which I precisely see mirror neuron findings and broader sensorimotor findings as challenging.

We saw how one of the core discoveries of the Parma group's pre-mirror neuron research was that many F5 premotor neurons seemed to be organized around goals rather than specific motor commands. This finding of the importance of goals in these areas led me to suggest that such neurons and areas maybe should be seen as inherently sensorimotor rather than primarily motor with additional sensory connections. My proposal is that premotor and parietal circuits play a key role in monitoring the broader 'affordance space,' and therefore it is exactly an attempt to hypothesize a function that makes sense of the sensorimotor goal integrations found in these areas. We saw in regard to Jacob and Jeannerod that they see these areas as primarily motor and see motor cognition in general as covert simulations of fully re-enacted actions from one's motor repertoire. I argued that this picture does not do justice to the empirical findings of simultaneous sensorimotor processes and more abstract event predictions in premotor areas. I want to hypothesize that these premotor and parietal circuits might represent 'goals' and teleological and causal relations at a more abstract level than the constraints of our overt and fully specified action repertoire allow for. Csibra, as we shall see, stays within the same narrow picture of motor cognition as the full simulation suggested by Jeannerod, and similarly logically infers without empirical backing that the teleological abilities of infants (without the appropriate motor repertoire) *must* necessarily depend on extra-motor and possibly purely perceptual processes. I, on the other hand, want to argue that given a broader notion of motor cognition the teleological focus and abilities of infants might indeed rely on sensorimotor integrations. More specifically, the empirical findings in premotor-parietal areas point to such a sensorimotor basis for goal representation.

In Csibra's 2007 article 'Action mirroring and action understanding: an alternative account,' he summarizes his view as follows:

> Observed actions elicit motor activations in observers that in the case they were executed, would generate similar actions to the observed ones. I challenge the most popular interpretation of these phenomena, according to which such action mirroring is generated by direct matching and serves the function of action understanding in

> terms of their goals. I propose that action mirroring is generated by action reconstruction via top-down emulation from action interpretation produced outside the motor system. Such action mirroring does not follow but anticipates ongoing actions and enables, beyond predictive tracking, action coordination with others. (Csibra 2007, p. 435)

Thus, the main points of contention regard 1) the idea of mirroring as direct matching of the observed action and 2) the suggestion that this mirroring process should serve the purpose of action goal understanding. Alternatively, he suggests that 1) motor activations during action perception occur due to 'top-down action emulation' and that 2) this process serves purposes of action anticipation and coordination rather than action understanding. I shall in the following develop his arguments on each of these points. As mentioned, I find that Csibra's sharp analyses often lead to excessively categorical conclusions due to certain common assumptions about what motor cognition is in general and how intentional actions work in particular. In the quote above Csibra summarizes mirroring activations during action perception as "motor activations in observers that in the case they were executed, would generate similar actions to the observed ones." So, given this wording, he can be taken to suggest that there are neuronal populations responsible for action production processes up until to the actual overt execution and that these exact neuronal populations are active in the same way during action execution and action perception. These activations are presented as the indisputable phenomena the process and function of which now need to be reinterpreted. As already discussed, this 'action minus execution' view of motor cognition and mirror neurons in particular is highly questionable. I have repeatedly suggested that the empirical evidence from mirror neurons or other behavioral studies of motor resonance precisely challenge this idea of identity between action observation and execution up until execution. I suggested in the previous chapters that the heterogeneous evidence does not even lend support to the idea that mirroring is one uniform and automatic and context independent mechanism. In the previous sections on Jeannerod I pointed to how the assumption of action mirroring as being like overt action minus execution is intertwined with a general theoretical view of cognition, in which the motor system is seen as an independently organized output system controlled by separate non-motor representations. Mirror and simulation processes might serve various 'off-line' purposes in regard to perception and central cognitive processes, but the 'system' or mechanism is strictly motor, and all motor process are like actions with or without execution.

In extension of this view, Csibra's analyses stress a particular top-down directionality of motor cognition, which is very important for his critique of mirroring as a direct matching process. Thus, he suggests that motor cognition is a one-way machinery for 'goal-to-action command simulation + overt execution,' and like Jacob and Jeannerod focuses exclusively on full covert simulation. One of the central aims of this dissertation is to challenge this view of the motor system as 1) one-way top-down and 2) as producing only full simulations, always one at a time. This input-output simulation framework of motor cognition seems to stand in the way of many of the good intuitions of both the proponents and critics of mirroring theories and their analyses of the possible social function of motor cognition. This narrow view of the motor system can be seen as inherent already in some of the initial mirroring hypothesis interpretations. In many ways I think that the criticisms of Csibra, Jacob and Jeannerod, by taking this view of the motor system so 'seriously,' really make explicit some of the tacit tensions that this view creates in the classical mirroring theories. But whereas Csibra and other critics seem want to fit the mirroring story into their view of the motor and social cognition, I on the contrary want to use the broader sensorimotor context of mirror neuron research to reinterpret these traditional frameworks of motor and social cognition.

Now, I will take a closer look at Csibra's arguments, starting with his reinterpretation of the nature of the neurological process. He argues that motor activity during action perception is not as a 'direct matching' process, but rather a reconstructive or emulative process.

5.4.1. Interpreting sensorimotor integration and 'direct matching'

Csibra points out that the mirroring metaphor gives the impression that the motor system somehow directly mirrors or copies the observed action. He labels this idea the 'direct matching' hypothesis. It is not obvious that any proponent of the mirroring theory has ever held the view that observed actions are directly matched in the sense that they are copied at the finest kinematic level. As we saw in Chapter 2, already from the first discovery of mirror neurons it was obvious that the modulation of many of these visuomotor neurons depended on the perceived action goals and that kinematic details were thereby often abstracted away from. However, even if mirror matching is supposed to happen at a more abstract level, Csibra rightly points out that any version of a direct matching theory would have to account for 1) how the mirror neurons are tuned to a given sensorimotor integration and 2) what exactly it is that is matched if it is not the exact movements. And, it is clear when reading the mirror neuron literature that there does not seem to be much consistency and agreement in regard to these challenges. I have proposed that one of the key reasons for the many inconsistencies might be that there are some classification mistakes going on. Maybe one such mistake is forcing the phenomena by trying to find one *universal mirror mechanism* for all motor activity induced by action perception. Like the traditional mirror theory proponents, Csibra does not question the category or the idea of one mirroring mechanism covering all instances of motor resonance. He wants to reinterpret precisely the process of this universal mechanism. He argues that maybe motor mirroring during action perception should be seen as action reconstruction rather than direct matching. The main point is that he wants to deny the ’directness’ of the mirroring process, the idea that an action is supposed to be mirrored before it is recognized or properly understood. Csibra argues that the action on the contrary must be recognized prior to any non-arbitrary sensorimotor translation of the action perception into a motor format, and thus that this translation happens at a relatively high level of cognitive processing that cannot be seen as ‘direct:’

> If observed actions receive mid- or high-level interpretation within the visual system before becoming transformed into motor code, the result of this analysis will provide the input to the motor system to be reproduced. What mirroring can then achieve is the reconstruction of the motor command needed to perform that action. (Csibra 2007, p.438)

The point is that only by analyzing the observed action to a certain perceptual level can a non-arbitrary choice be made as to the right level of description at which to mirror the action motorically. Csibra argues that the action must be understood at this very level prior to the mirror process, and that the function of the mirroring process then is to reconstruct motor commands given the specific level of action understanding description used. Thus, we can see how Csibra’s analysis of the mirroring process also leads to a reinterpretation of the function of mirror neurons as relating not to action understanding but rather to action anticipation.

Csibra might be putting his finger on something interesting when he argues that mirroring must happen at a non-arbitrary level of analysis. This issue relates to the general question of what it is that gets 'mirrored,' but it also brings up the general question of whether there is an element of, if not choice*,* then 'picking out' in mirroring and sensorimotor integration as such. The traditional mirror mechanism interpretation proposes that the mirroring is automatic, ubiquitous and direct. Csibra wants to criticize the idea that mirroring should be automatically evoked by 'low-level' perception, and argues that the perception must have reached a "mid- or high-level interpretation within the visual system before becoming transformed into motor code." His thought is that the action must be interpreted enough to provide the grounds for a *choice* of the appropriate level of motor mirroring. He understands mirroring along the same lines as overt action choice and as facing the same issues as overt imitation. However, my question is whether one should think of basic sensorimotor integration processes - of which mirror neurons are one instance – as chosen or even intentionally chosen. Or if, as I have argued, one should rather see these sensorimotor integrations as the basis on which our action choices become possible. For Bergson it is a key organizing theme for the understanding cognitive processes that information only is gathered for the purpose of a better action choice. We saw that his view is that the motor aspects of cognition as always involving a limitation in regard to the heterogeneous chaos of perceptual stimulation, and therefore that it is this limiting specification that makes the perception relevant and meaningful to our point of view as agents.[318] In regard to Csibra's claim about choosing a 'level of description,' the question is if Bergson's thoughts might not point to a less deliberate and more automatically evoked alternative to the idea of 'choice.' Sensorimotor integrations might simply be established not by random statistical association but might instead be organized around action goals and pragmatic parameters. Thus, automatic sensorimotor activations might actually be seen as what informs perception about what counts as action relevant levels of description.

These ideas of sensorimotor integration as pragmatically structuring perception also fit very well with Erik Rietveld's analysis of unreflective intentional action choices as being performed in a field of affordances. This is, in other words, the view that such unreflective choices depend on action invitation perception and emotional evaluation. I

[318] Bergson (1986) p. 202-204

argue that this view in combination with an understanding of sensorimotor integration might hint to a road forward not only for an understanding of intentional action choice, but also for some of the intuitions of Parma group thinkers like Gallese and Rizzolatti in regard to social cognition. A sensorimotor grounded social affordance monitoring process would give motor cognition an interpretive role in social perception, but without any need to insist on either an insulated agent neutral mirroring mechanism or on full covert simulation of all observed actions.[319]

But these Bergsonian ideas by no means figure into Csibra's own view of the process of choosing the right level at which to mirror action perceptions. His thought is that the perception is interpreted and the choice is made prior to any motor system involvement. To fully understand Csibra's view, it is again important to see how it ties to a more general view of the motor system and the nature of action production. He describes how actions are organized hierarchically with action goals over sub-goals of various action means, and how these sub-goals again can be actualized by various means etc. The goals constrain each other hierarchically but without fully determining the lower level sub-goal, i.e. different action means but not just any action could do the job. Following this idea that action decisions and commands pertain to a specific level of hierarchical goal representation, Csibra draws a parallel to overt imitation. He suggests that the distinction often used in social learning literature, between respectively emulation of action goals and imitation of both goals and means, obscures the fact that imitation always at some level is emulation. It is simply a question of the level of detail at which the other's action goal/sub-goal is emulated:

> My point here is not that imitation as such does not really exist, but rather that any instance of imitation is actually achieved by emulation. When someone imitates, she chooses a certain level of description of the observed action, and reproduces that level by reconstructing it in her own motor system. From this perspective, whether something is imitation or emulation is not a well-formulated question. Instead, one should ask how the imitator decides what is the relevant level in the action hierarchy on which an action is to be reproduced. (Csibra 2007, p.440-441)

Thus, according to Csibra imitation is a product of emulation at various levels of precision. Or, in other words, the precision of the emulative imitation depends on the level of action description that the imitator chooses as relevant for the current situation. I

---

[319] In addition to the fact that – contrary to what Csibra vehemently insists - there would not be any need for any deliberation prior to any motor involvement.

think this might very well be a correct analysis of overt imitation. Csibra here also conveys some important insights not only about imitation but also about overt action planning in general. He writes that 'the imitator decides' the relevant level of description. I will suggest that this decision process would not make sense independently of the broader social, pragmatic and motivational context – and that our understanding of this context is exactly rooted in 'always already' engaged sensorimotor representations of the affordance space. Csibra himself never explains how the interpretation and choice on the part of the imitator/observer is made. His point is that it must be made prior to the initiation of the imitation action. Then he quickly moves on to suggest that covert action mirroring just like overt imitation depends on an interpretation and choice of the relevant level of description. It is exactly this move from overt imitation as *global behavior* to covert sub-personal sensorimotor 'translation' that I challenge, as I think that many of the sensorimotor integrations precisely provide the foundation on which the overt/emulated action choice is made. I order to help put my challenge in perspective I will quote at length:

> Just as the *mechanism* of imitation is always emulation at a lower level, the *mechanism* of motor mirroring is always reconstruction. There is no mysterious mirroring process that directly transforms action observation into motor code. Rather, the observed action is analyzed at some level of precision and the result of this analysis is mapped onto the observer's motor system. One can call this mapping process 'direct matching' (Rizzolatti et al., 2001) and such mappings may be established by 'direct' associations, but what is mapped during mirroring is not an uninterpreted signal but a description of the observed action at some level of the action hierarchy. The fine details of the resulting motor activation in the observer do not directly originate from the observation but are reconstructed from this description. (Csibra, 2007,p. 441)

Csibra thus argues that the observed action cannot be reproduced un-interpreted. I think that he might be right that the full intentional actions are imitated based on an interpretation and a choice of an appropriate goal level determining the level of emulation detail. The question again is whether all *motor resonance* fits this overt imitation - or covert 'intentional action minus execution' - picture. In this quote, Csibra opens the possibility that mirroring might be 'direct' in the sense that neurological 'associations' have been established. His point seems here to be that for this association or mapping to be formed *in the first place* the action must have been visually understood and interpreted prior to the motor association. Thus, now the 'choice' seems to be deferred from each mirroring action to the ontogenetic establishment of the association.[320] In other words, now the claim seems to be simply that the first mapping must be based on a prior interpretation. Of course, my alternative order of events – following the Bergsonian proposal above - would be that it is the experience shaped sensorimotor integrations that create the various meaningful levels of description. In any event the idea of a stored sensorimotor mapping – irrespective of whether it came about via Csibra's mid/high level of interpretation – opens the door for many *simultaneous* perceptually induced sensorimotor activations. In other words, just because we control overt actions in a certain 'top-down' way and sensorimotor integrations might be shaped by action interpretations, that does not exclude the possibility that we have multiple parallel sensorimotor activations and bottom-up processes happening concurrently with top-down control. The question is why Csibra would assume that all the already established sensorimotor integrations would be idle or quiet when not under top-down control. But he does not seem to see such simultaneous top-down and bottom-up activity as an option. He moves on to say that "the crucial question about action mirroring and its relation to action understanding concerns the level of action interpretation where the mapping from visual to motor code takes place."(Csibra 2007, p.441) Note the use of the singular - the level, the mapping. Again, from the standpoint of action control in overt action, I think Csibra is right that one specific level of goal representation serves to organize actions hierarchically. But what is at stake here is not controlled overt actions but all motor resonance during action perception.

As already mentioned Csibra understands mirroring on the model of overt imitation. And thus when he talks about how the mirror mapping takes place he is tapping into the big debate in the imitation literature referred to as the 'correspondence problem.' That is the question of how we come to associate visual input with a corresponding action that we in basic cases such as facial imitation have motor but no visual experience of.[321] Csibra here argues that the mapping or translation of mirroring can only take place at a certain level if

[320] I shall not try to solve the mystery of how sensorimotor integrations and mirror neurons 'get wired' as these issues are hotly disputed and complex. Cecilia Heyes and also Keysers and Perrett (2004) have put forth powerful arguments in support of the proposal that the sensorimotor integrations in mirror neurons maybe simply should be understood as a product of experience based on Hebb's law of associations ('what fires together wires together').

[321] See Brass & Heyes, C. (2005) for a critical discussion of mirror neurons in relation to the correspondence problem.

the action is understood as this level first, and following this logic he concludes that the action must always be understood but at the appropriate level before the mirroring process is initiated.

Given his framework of mutually exclusive bottom-up ‘action-to-goal’ or top-down ‘goal-to action’ simulation, Csibra explains his disagreement with traditional mirror theories as pertaining to not just the question of what level the action mirror mapping is thought to take place, but maybe even more importantly as being about the ‘direction’ of the simulation process. Traditional ‘direct matching’ theories focus on low-level action description mirroring that propagates “bottom-up” to a higher level action and intention understanding. In contrast, Csibra’s own view is that we start by mirroring the action at a high level of interpretation and then via a “top-down” reconstruction produce the more fine-grained motor simulation of the observed action. He summarizes the differences as follows:

> The difference between the two models lies in two factors: the action interpretation level at which visuomotor translation is performed, and the propagation direction of activation within the action control system following mirroring. (Csibra 2007, p.441-442)

Given his self-imposed ultimatum between low-level and high-level/bottom-up and top-down, Csibra chooses a top-down emulation mirroring process. I think his view of imitation as emulation is very plausible, and I also agree that actions are organized hierarchically in a way where higher goals constrain sub-goal choices. But again the unasked question is whether one needs to assume that sensorimotor integration only happens at one level of action description, and whether the strict dichotomy between bottom-up and top-down processes is empirically and theoretically tenable. I suggest that Csibra might be misled by an overly strong analogy between overt action imitation and the perceptually induced activity of mirror neurons, between top-down organization of action control and the possible ‘bottom-up’ or low-level sensorimotor integrations that concrete action choices are based on.[322] The acceptance of the sharp dichotomy is illustrated by an interesting discussion Csibra offers of the ‘tension’ in the mirror neuron literature between bottom-up and top-down processes and between ‘goal to action’ and ‘action to goal’ inferences:

---

[322] Csibra’s reliance on the overt action choice and control to understand covert sensorimotor integrations is parallel to what I earlier labeled the ‘one body in the mind’ assumption.

> All these findings reflect a tension between two conflicting claims about action mirroring implied by the direct-matching hypothesis: the claim that action mirroring reflects low-level resonance mechanisms, and the claim that it reflects high-level action understanding. The tension arises from the fact that the more it seems that mirroring is nothing else but faithful duplication of observed actions, the less evidence it provides for action understanding; and the more mirroring represents high-level interpretation of the observed actions, the less evidence it provides that this interpretation is generated by low-level motor duplication. (Csibra 2007,p.446-447)

Csibra's logic is that if simulation is a one-way street with an input and an output, then one must make a choice between a 'goal to action' and 'action to goal' inference process. But the question remains if the motor system represents such a one-way input-output machinery. Csibra insists:

> One cannot have one's cake and eat it too: the discharge of a set of MNs cannot represent the activation of the observer's motor system at low and high levels at the same time. (Csibra 2007,p.447)

I am exactly suggesting that in a world of dynamic processes and reciprocal and looping integration one might just be allowed to have some of one's cake and eat it too! Csibra could of course be right in assuming that mirroring motor activations during action perception is like overt imitation. But via my discussion of Jeannerod's influential version of this theory of motor cognition as action simulation, I have vehemently tried to argue that this sort of process only represents a small part of the story and that it is too narrow a concept for a general understanding of motor cognition and sensorimotor integration. I have pointed out how many findings of perceptually induced motor activity cannot be made sense of as 'covert action simulation,' i.e. simply as actions that are inhibited before they reach the execution phase, as actions identical to overt actions up to that point[323]. By way of gesturing towards a view of intentional action where relevant action choice is seen as dependent on already being in a field of affordances, I propose the hypothesis that overt actions and full simulations depend on there already being perceptually induced sensorimotor activations. I thus argue that there are essential differences between action imitation and action affordances and action goal monitoring. Furthermore, I want to point out that the empirical findings of the Parma Lab and others of complex sensorimotor integration during action perception exactly suggest that it is problematic to pigeonhole motor cognition as one process of inverse modeling from pre-

[323] See here my discussion of Jeannerod's notion of motor cognition as simulation (Section 5.2.1) and also my discussion of the broader empirical findings of sensorimotor intergration (Section 2.2).

existing 'goal to action command' schemas. The point for now is that Csibra's quick logically inferred conclusions depend on a specific view of motor cognition where the motor system is an instrument for computing action commands given goals that are already specified and chosen outside of the motor system.

5.4.2. Emulation rather than low-level mirroring?

Csibra wants to deny 'direct matching,' in the sense that sensory and motor information is integrated prior to an at least 'mid-level' perceptual analysis of the observed action. Sometimes it seems like he is denying low-level sensorimotor integration per se. But, as I have suggested, that would make it rather tricky to see how his action control hierarchy would become specified in a concrete situation. After all, how would an action goal of eating an apple and the sub-goal choices of how to do it be made without perceptual guidance and sensorimotor affordances at a lower-level? Csibra's argument against low-level sensorimotor integration centers not on perceptual guidance of actions, but narrowly on action mirroring. I will therefore push the general sensorimotor integration issue aside for now, but will simply draw attention to the seemingly arbitrary distinction between mirroring and other sensorimotor integration.[324]

In regard to mirroring, Csibra's claims are 1) that action mirroring is based on a purely perceptual analysis and understanding of the action goal at a relevant level of description prior to a transformation of this goal into a motor format, and 2) that this 'mirroring' transformation does not happen at a 'low' or 'direct' or un-interpreted level of processing. As mentioned my hunch is that the plausibility of these ideas rides on a very narrow understanding of motor cognition. I have here challenged the idea that all mirroring processes can be characterized as 'goal-to-action' emulation, but I have also earlier questioned whether a purely perceptual goal understanding is plausible. These matters can be further elucidated by looking a little closer at Csibra's own arguments.

In regard to the exclusion of low-level un-interpreted mirroring Csibra discusses the 'end goal' modulation found in some mirror neuron studies. Fogassi's single cell study on

---

[324] Csibra might also in the case of other sensorimotor processes insist that the perceptual analysis is rather advanced even before the sensorimotor integrations like those of for example the dorsal visual stream and it's functional role in 'perception for action.' But if the 'multi-level' sensorimotor connections exits anyway as various possible levels of interpretation and emulative mirroring are possible, then one might suspect that lower-level connection would not be totally idle during action perception at a higher level. And such parallel sensorimotor activity would question the strict bottom-up top-down dichotomy.

monkeys and Iacoboni's fMRI study on humans discussed in Section 4.3.1 are often heralded as proof that mirror neurons play a role in intention understanding. I do think that Csibra is right in pointing out that these studies as such really only show that mirror neurons reflect intention/goal understanding rather than provide direct evidence that the functional role of mirror neurons is to produce such understanding. However, Csibra wants to make the further point that the studies provide evidence that mirror neurons are top-down rather than bottom-up modulated. He writes: "This is a clear demonstration that MNs take into account the further goal, and not just the perceived action, when responding to observed actions.“ (Csibra 2007, p.445) In light of my discussion of these truly intriguing findings in the previous chapter the end goal does seem to be important for the modulation of many of these mirror neurons. However, Csibra does not mention that it was not all but rather only around 50% of the neurons that were thus modulated by end goal inferences. Could it be that Csibra's own story is too simplistic here? I think so. In the earlier chapters I have tried to point to the heterogeneity of the mirror neuron findings and questioned the attempt to squeeze the broad spectrum of found action sensitive visuomotor neuron properties into a simplistic mirroring theory. Csibra's top-down emulation theory is different but as simplistic.

This being said, I think it is important to raise the question of whether low-level mirror matching really exists. It is certainly not obvious from the heterogeneous findings what exactly - if anything - is mirrored in various cases and by various types of mirror neurons. Perhaps Csibra is right to say that insofar as actions are actually 'mirrored,' in the sense of covertly sequentially simulated, this mirroring happens at the level of goal recognition and not at un-interpreted levels of movement command. In other words, we can say that just like in the case of imitation it is not the movements that are mirrored but their goals and sub-goals. A possible difficulty for this story is that we saw already in Chapter 2 how certain types of mirror neurons seem to be attuned to very specific temporal and hierarchical segments of the observed action. As I read him, Csibra's claim here is that these neurons still do not 'mirror' an un-interpreted observed movement. The activity can in different scenarios be either a part and a consequence of an emulative reconstruction of the observed action starting at a higher level, or mirror a very specific and already interpreted sub-goal, for example 'impact with object.'[325] The point is that in both cases there is a top-down process that determines the activity of the mirror neurons.

Another argument that Csibra raises against low-level mirroring is that we can actually mirror actions that are biologically non-executable. He refers to a study by Ferrari et al (2005) showing that some monkey premotor mirror neurons can be trained with purely perceptual exposure to respond to goal-directed actions with tools that are not in the monkey's own motor repertoire.[326] Csibra concludes:

> In other words, MNs responded to the sight of a non-executable action with a different action that the monkey could have used to achieve the same goal. This is exactly what the emulation model of action mirroring predicts for observed actions whose goals are interpreted outside the motor system and then fed into the observer's action control system for reconstruction. (Csibra 2007, p.446)

Csibra argues that we recognize and understand the action goal at the appropriate level perceptually prior to sensorimotor integration and only after this understanding is the goal mirrored and emulated by the motor system. However, as was that case in the Iacoboni study referred to earlier, not all but rather a smaller subset of mirror neurons learned to respond to the tool using goal-directed action. In other words, one problem with this argument is that it attempts to make the broad conclusion that we always understand the action goal before any sensorimotor integration or 'transformation from perceptual to motor format,' to use Csibra's terminology. Another, and probably bigger problem with Csibra's argument is that motor activity is not always easily understood in terms of his 'goal-to-action' emulation model. I have already repeatedly referred to Ricarda Schubotz's findings that we seem to use the very premotor areas that have been implicated in mirror neuron studies to predict non-biological external events[327]. These findings are interesting as they problematize the idea that motor cognition is limited to our full re-enactment of action within our motor repertoire, and what I have earlier labeled the 'on body in the head' view. The idea is that if premotor areas and therefore presumably motor cognitive processes support prediction of non-biological events, then these processes must work at a level of abstraction significantly removed from the

[325] This latter case would be a rather rare case where the observer was very focused on the exact details of the observed action. The problem of course is that this analysis does not allow for parallel low-level 'pre-choice' sensorimotor activity that could support an 'affordance space' representation. I have argued that without such parallel processing, action choice and top-down determination would be tough to make sense of in the first place.

[326] See: Ferrari et al. (2005).

[327] Schubotz, R. (2007). See also Schubotz & von Cramon (2004) for similar findings.

concrete body and its kinematics. As we saw one of the significant early sensorimotor findings of the Parma group was that many premotor neurons seems to be modulated by the action goals in relative abstraction from the concrete movements. Mirror neurons mostly appear to operate at a similar level of abstraction. But this finding by Schubotz et al. takes things much further as it appears that sensorimotor areas supports that anticipation of perceptual outcomes that does not even involve an acting body or an intentional action. Thus, Schubotz's findings do not provide evidence for low-level mirroring but question the idea that motor activations are always emulative in the sense of specifying concrete motor commands on the basis of explicit goals. And if there are no such limitations on motor cognition then the 'tool mirroring' findings of Ferrari and his coworkers could be seen as representing the goal of the tool using action at a level of abstraction above its precise motor command implementation. Csibra's argument is that since the monkeys cannot use tools themselves their motor activity is due to a reconstruction of how they with their specific motor repertoire would perform the action in question. However, the question is how tightly anchored sensorimotor integrations need to be in our specific motor repertoire, and whether all motor activity represents specific motor commands - or whether, alternatively, we via sensorimotor integrations can come to understand and predict action goals and events that we cannot exactly perform. I think that the evidence for motor activation in regard to non-biological actions and external events supports the latter.

5.4.3. Purely perceptual goal attribution

This brings me to the issue of how to understand the 'purely perceptual' goal comprehension. It is exactly because Csibra sees motor cognition as going from 'goal-to-action' and as tightly anchored in motor commands and action repertoire he argues that the evidence suggests a non-motor goal comprehension and attribution mechanism.

I share the belief that motor commands are organized around goals. As a matter of fact, I suspect that the ideomotor theory has got it right and that motor commands can only be initiated and specified if in terms of some level of goal - and thereby an anticipatory idea of what the action command is supposed to do. Even a simple act like bending a finger seems to involve somatosensory or visually based anticipation of the end-goal. The sensorimotor relation between goal and action commands and perceptual 'end state' representations is the reason why I have posed the question, already in Chapter 2, of whether a purely motor goal representation makes sense, and further what such a non-perceptual goal representation could look like. Now Csibra demands of me to ask the opposite question, namely what a pure perceptual goal understanding or representation would look like. But instead of offering a positive account of purely perceptual goal understanding, Csibra uses examples to argue that motor simulation is not necessary for action understanding. Given his narrow concept of motor cognition, and similarly to Jacob and Jeannerod, he takes the fact that we are able to understand many actions that we cannot ourselves produce as all the evidence he needs to draw his conclusion. Thus, on the basis of examples of action understanding beyond motor repertoire, he not only concludes that a detailed covert action simulation is not essential for action understanding, but he also goes a few steps further and concludes that motor cognition in general is not necessary and that we can understand action goals by purely perceptual means. This argument obviously depends on the assumption that all motor cognition is covert action simulation. It is given this assumption that he excludes the relevance of motor processes and infers the ‘purely perceptual’ hypothesis. However, I have in this chapter never stopped contesting that assumption, and can therefore not so easily swallow the conclusion.

But beyond the negative argument Csibra says the following about what it means to understand an action or an action goal: “Understanding the goal of an observed action involves figuring out the content of the intention that generated the action.” (Csibra 2007, p.447) This wording clearly also gives us a glimpse of the underlying action theory – which Csibra seemingly takes for granted-- namely, that action initiation is a causal process that is ’generated’ by an intention with a certain goal content, which can be represented prior to and independently of the action, which it produces. Via the idea of actions being chosen and specified in a field of affordances, I hypothesize that the intentional or motor representation of the goal on some level already is active before one really decides to perform this action at this time, i.e. before the action choice is definitively made and the overt action execution initiated. Csibra’s language and theory is here quite different from mine. He says that the task of ‘extracting the intention’ from an action represents an ‘inverse problem;’ it cannot simply be deduced because the action always underdetermines the intention, and the question is how we figure out the most probable intention given the action. Csibra's idea of underdetermination here also seems similar to Jacob and Jeannerod's insistence that prior intentions cannot be determined by observable behavior. Csibra refers to three candidate mechanisms for extracting the intention and goal,[328] but focuses on the point that mirroring is not always necessary for action understanding. He points to empirical evidence from looking time studies with infants showing that they understand action intentions well beyond their action means.[329] Firstly, Csibra refers to a Heider and Simmel style[330] study showing that animated geometrical figures with no resemblance to the effectors of the infants' bodies are interpreted as performing goal-directed actions. Twelve month old infants were first habituated to seeing a circle 'jumping' over an obstacle 'to move to' another circle. Then the obstacle was removed, and looking times indicated that infants given this change in the action context expected the circle to adapt to the new environment and run along the ground, rather than repeat the very jumping action they had been habituated to watching. Csibra concludes that, since these infants neither share bodily features nor can for that matter jump, the action expectation must be based on something else than motor simulation.[331] Thus, Csibra uses these geometric cartoon studies to argue that goal attribution must be purely perceptual. But, interestingly enough, some recent evidence seems to implicate premotor 'mirror' areas in the processing of such non-biological intentional stimuli.[332] In other words, given his preconceived notion of motor cognition, Csibra uses the studies to make the logical argument that action understanding must be purely perceptual. However, the twist is that these very studies might provide evidence that his assumptions are wrong and his concept of motor cognition too narrow.

The second experimental paradigm involves cartoons of effective but motorically impossible actions.[333] Infants were here habituated to a hand reaching behind an obstacle

[328] Csibra mentions three possible methods of intention estimation: association, mirroring and teleological reasoning. I shall discuss his analysis of these three action understanding methods in the following, with reference to a joint article with György Gergeley where each method is further conceptualized and where the authors suggest that the methods might to some extent be complementary.

[329]In the context of my discussion of Onishi's alternative false belief test in Chapter 4 Section 4.2.6.A, I introduced the 'looking time' and 'expectation violation' paradigms used in many cognitive studies of pre-linguistic infants, and these shall be brought up again in the next section.

[330] This is another of the usual empirical suspects thought to support non-motor theories of social cognition. I discussed in the previous section how these experiments might actually rather support a more complex notion of motor and social cognition. See Heider and Simmel (1944).

[331] See original report of the experiment Gergely et al (1995).

[332] See Gobbini et al. (2007).

[333] Southgate, Johnson and Csibra (unpublished) Infants attribute goals to biologically impossible actions.

to grasp an object and then an additional obstacle is introduced that makes the first direct reaching action impossible. Given the choice between a two-step process of removing the new obstacle and then reaching or engaging in an impossible one-step zigzag snake-like action towards the goal, the evidence showed that 6-8 month olds expect the more effective but impossible one step action. Note that the biomechanically 'impossible' action has one goal and one action as opposed to the alternative of a goal and a sub-goal with each their goal-directed action. Again, Csibra concludes that motor simulation is not necessary for understanding the goal of the action. But, again, I have to point out that the experiment does not exclude the possibility that the action expectations rely on motor cognition as such but simply that it relies on a detailed motor re-enactment. Interestingly, Csibra at another point in the articles points to the findings by Costantini and colleagues,[334] which exactly show activity in fronto-parietal circuits when watching impossible actions to argue against a direct matching process:

> Action mirroring can even occur in response to biologically impossible actions. Costantini et al. (2005) measured observers' brain activation by functional magnetic resonance imaging while they were presented with finger movements that were within or outside the normal range of such actions. The impossible action depicted a hand with the little finger moving laterally for 90°. The results showed that the human ventral premotor cortex (part of the human 'mirror neuron system') was activated equally by the possible and impossible actions. Costantini et al. (2005) concluded, "the premotor system does not take into account the biomechanical constraints the observed movements would involve if they were actually executed. (p. 1765)" (Csibra, 2007, p. 446)

I very much agree with the conclusion of Constantini and colleagues here, namely that premotor areas might be able to represent action goals at a level of abstraction beyond the biomechanical constraints. And this would of course also fit with the results reported above that babies prefer a biomechanically impossible but simpler 'one-step' action over a possible 'two-step' action. However, Csibra wants to insist that motor cognition is about action commands available in the motor repertoire. He therefore reasons as follows:

> Is it possible that such 'mirroring' of an impossible action is produced by 'direct matching'? As there is no matching action in the observer's repertoire, this is unlikely. However, it is conceivable that the visual system can provide an appropriate description of the end-state of such an action ('the little finger is perpendicular to the others'), which then the motor system attempts to approximate, albeit unsuccessfully,

[334] See Costantini et al. (2005).

> using the available motor programs. In other words, mirroring the observed action can be attempted by driving the motor system top-down from a mid-level interpretation of it. (Csibra, 2007, p. 446)

It is of course an empirical question whether this ‘indirect approximation’ suggestion is right, but it is clearly theoretically rather than empirically motivated as it stands, and thus emblematic of Csibra’s narrow assumptions about motor cognition.

Lastly, Csibra refers to an experiment by Onishi et al. (2007)[335] involving 15 month olds watching sequential pretend actions. Here, the experimenter pretends to pour water into one of two cups that have just been turned over so as to show they were empty when the pretending began. Then she pretends to drink from either of the two cups, and the infants’ expectations clearly seem to be violated when the experimenter is drinking out of the ’empty’ cup. This is a really interesting study showing that these small kids really “recognized the causal and teleological relatedness of these actions” or, to use my own terminology, that they are monitoring the changing socially instantiated affordances of sequential pretend actions. Again, Csibra argues that since 15 month olds do not know how to pour or drink from a cup these pretend actions cannot have been understood via motor simulation. I presume I at this point do not have to repeat my objection to his conclusion.

Overall, I think that Csibra is right in pointing out that these examples seem to show that we can understand actions in the absence of detailed motor mirroring or reenactment. But, at the same time, I repeatedly insist that it takes a few further steps to argue that these are cases of ‘purely perceptual’ action understanding without any motor involvement. Certainly, there are plenty of affordance structures present that could involve some level of sensorimotor integration and analysis. Further, as mentioned, there seems to be neuroimaging evidence that premotor areas are modulated by both geometrical cartoons like the one referred to in the first study, and also by certain forms of sequential reasoning that do not involve action or action perception at all.[336] The question is how one accounts for the many seemingly more abstract cognitive functions of premotor areas if one insists on a narrow ‘covert action simulation’ interpretation of motor cognition? One reply could be that these motor areas are known to have various

[335] Onishi, Baillargeon, and Leslie (2007).

[336] Schubotz and von Cramon (2004). Gobbini et al. (2007). Somewhat different focus, but see also Mar et al. (2007), on the difference between agency attribution to biological and non-biological agents.

non-motor functions and these findings could point to a few more. Premotor areas are for example implicated in working memory tasks and language sequencing. Thus, one might simply want to start with a narrow definition of motor processes as motor repertoire simulations and then suggest that these motor functions arbitrarily cohabitate within parts of the brain with non-motor functions. My alternative approach is to argue that the narrow notion of motor cognition is problematic in and of itself, and that we must allow for a more liberal notion of sensorimotor activity to understand the perception of affordances, the representations of goals and the process of intentional action choice and specification. In short, the narrow notion of motor cognition does not even seem capable of explaining the core motor functions let alone more abstract cognitive functions. I suggest that armed with a more liberal notion of motor or sensorimotor processes, one might start to understand what on a functional level is shared between for example working memory, linguistic performance and action orchestration. My proposal here is that fronto-parietal areas might support a more general ability to create and navigate between real and hypothetical affordance spaces and functions. In brief, I suggest that these various findings might help us go beyond the narrow 'full simulation' notion of motor cognition and guide a new theory of a motor role in social cognition and other higher cognitive functions.

5.4.4. An action anticipation and prediction function of mirror neurons

Given Csibra's narrow interpretation of motor cognition and mirroring processes as a goal-to-motor command simulation, he argues that the function of mirroring is not action understanding but rather action anticipation and prediction. Given his idea of motor cognition this hypothesis makes a lot of sense. But additionally there are a series of empirical findings that support the idea that motor areas such as the premotor mirror neuron areas are modulated in a predictive rather than simultaneously mirror matching way.[337]

For *particular instances* of detailed simulative mirroring I think it is reasonable to say this process is supported by fronto-parietal sensorimotor processes. Still, as I have

[337] See Kilner et al. (2003, 2004, 2007a, 2007b), Southgate et al. (2009), Flanagan & Johansson (2003) and Falck-Ytter, et al. (2006).

argued, such specific perceptually induced action simulations depend on intentional focus and are the exception rather than the rule of normal action perception. Thus again, my main point of contention with Csibra is that he thinks of all motor processes uniquely in terms of overt or covert 'goal-to-motor command' processes. It seems to me that the empirical findings do not support this claim, and further that many of the findings of motor anticipation during action perception could be explained by schematic goal and affordance representations that lead to motor resonance in the absence of competing actions plans. In order to substantiate this suggestion I will first tease apart two links between goals and action anticipations.

In regard to the idea of action anticipation, we have already seen that Gallese and his colleagues under the 'chains of action' version of the mirror theory link action anticipation to goal and intention anticipation or attribution. I agree that there is a sense in which action intentions themselves involve anticipated action outcomes – in that we somehow have to anticipate the outcome/goal of action to even understand or represent it. This anticipation thus has to do simply with the anticipatory future-directedness of goal directed actions in that the goal, so to speak, is placed in the future.

Csibra suggests that action goals/intentions lead to this general anticipation of obtaining the goal. But additionally, via the goal-to-action simulation process, he argues that there is a much more specific anticipation of the detailed kinematic actions that come between the initial position and the end goal. I agree that both these sorts of anticipation can be linked to goals. Our disagreement is over whether the *normal function* of mirror neurons and other sensorimotor neurons has to do *only* with the latter simulative anticipation and not the former goal representation, as Csibra suggests. As argued above, I think his view that the goal is attributed via a purely perceptual process is based on a faulty logic and little empirical evidence.[338]

I agree with Csibra that it seems implausible that we would have to go through the detailed sequential simulation process to infer the action goal when the goal of the action

[338] There is empirical research supporting a purely perceptual theory of action understanding, or at least supporting the view that much action goal understanding happens in non-motor areas such as, for example, the STS (See for example: Jellema & Perrett (2005 & 2007)). However, as I understand the evidence, it is not quite clear which neurological processes are the more basic in goal understanding. I therefore conclude that sensorimotor processes which are also heavily empirically implied in goal understanding at least should be given the benefit of a doubt, and not be ruled out on the basis of a logical reasoning on questionable premises. (See for example: Hamilton et al (2006)

can be understood via dynamic relations between the parallel action representations and affordances of the context. For example, if I see your face light up as you see the coffee is brewing, and that you move towards the cupboard – do I need to go through an internal simulation of your movements to predict that you will search for a cup and then pour coffee in it? I propose that the link between action intentions and anticipated action outcomes should be understood not via a sequential process but rather in the ideomotor sense that action representations via prior experiences are organized and stored around multiple hierarchical goals and anticipated outcomes. Thus, I see that you see and smell the coffee, that you want coffee and that you see the cupboard etc. and via my perception of your actions and affordances, I can represent your goals and sub-goals and to a large extent understand not only your actions but also the choices that you are making. All the findings of the Parma Lab suggest that these representations of the goals and affordances are supported by fronto-parietal circuits. Each teleological representation seems to imply a certain element of action anticipation and prediction, but only yield the sort of detailed action sequence anticipation that Csibra talks about if a specific action is picked-out and simulated. To summarize, I think that fronto-parietal circuits are involved in both general goal anticipation and detailed action sequence simulations. But just as I do not think we understand the goals because of action sequence simulations, I do not think that we necessarily simulate the detailed action sequences just because we understand the goal. I thus agree with Csibra that action sequence simulation does not yield goal understanding. Where we fundamentally disagree is about whether action sequence simulation is all the fronto-parietal circuits do. As opposed to him, I think that mirror neurons and other sensorimotor neurons of the fronto-parietal circuits are very important precisely to the process of conjecturing the goals of others as they act in a perceived affordance space. Interestingly, Csibra's most fascinating research is on how human infants are naturally inclined to understand their perceptions of movement and change in terms of teleology and causal efficacious properties. I think that his findings are very important – but disagree that such teleological and causal understanding should be produced by a special non-motor 'teleological reasoning mechanism.'

5.4.5. Csibra on teleological action understanding

Csibra argues that mirroring is not essential to action understanding and goal attribution, and that these basic social cognitive processes rather depend on 'purely perceptual' analyses and processes. In a joint article with György Gergeley entitled ''Obsessed with goals': functions and mechanisms of teleological interpretation of actions in humans,'[339] the authors expound on what they think the 'mechanism' of action understanding is. Csibra and Gergeley center on what they call the 'teleological stance' of children, that is on how even young infants naturally recognize agents as such and interpret the observed actions in 'teleological' terms, i.e. as functional, intentional and goal-directed. Csibra and Gergeley want to ask what function this teleological stance has, and to discuss what mechanisms could support it. They see the function of goal attribution as linked to action anticipation and prediction. This is interesting, as we just saw that Csibra thinks that motor mirroring serves as a mechanism for action anticipation. They write that the goal seems to play a predictive role in the two senses just discussed. First, in the simple sense that to act intentionally we need to be able to reach goals in the future by somehow representing these goals prior to the action that brings about that intended future state. Hence, the goal attribution predicts the future of this goal being reached. Secondly, Csibra and Gergeley suggest that goal attribution also informs more specific action anticipations, via our own knowledge and ability of not only recognizing the goal but also reaching it, i.e. our ability to anticipate the intervening actions. It is this process they see as a process of motor simulation. And the resulting motor anticipation is what they see as the 'output' of motor simulation in general but also mirroring emulation in particular. Csibra and Gergeley, again using the schema of 'goal-to-action' versus 'action-to-goal' inferences, classify action means anticipation and social learning of new means as goal-to-action' inferences, and the anticipation of goals and learning new goals from observing others as 'action-to-goal' inferences. They argue that these two kinds of processes are supported by two distinct computational mechanisms:

> In computational terms, either the percept of an action provides the input of the system, activating an appropriate mechanism to identify and output the goal to be attributed to the actor (as in goal prediction and social learning of functions), or it is the goal that serves as the input to the system for which the mechanism is expected to produce a (predicted or learnt) action as the output. (Csibra & Gergeley 2007. p.64-65)

[339] Csibra & Gergely (2007).

According to Csibra motor simulation can only give us 'goal-to-action' inferences, and thus never the inference to a goal by action means given an already inferred goal. Csibra and Gergeley continuously refer to Marr's vision model as providing an analogy to goal-to-action and action-to-goal inference computations. They thus suggest that action-to goal-inferences represent underdetermined or inverse problems just like the task of creating/determining a 3D percept from a 2D retinal impression does. This diagnosis of the inverse problem of going from 2D to 3D vision depends on the questionable and dare I say outdated theory of Marr's that we actually base our perception on relatively non-temporal 2D representations in the absence of occulomotor and other sensorimotor information. Csibra and Gergeley suggest that in the case of action and goal inferences we solve the problem by the help of contextual and prior information and "a very effective goal extraction mechanism:"

> Our knowledge about people, kitchens, cups, and water, and further information about the actual situation will help us to find an answer, but perhaps nothing in the pouring action itself can serve as the basis for choosing among these possibilities. Just like a certain arrangement of edges on the retina can be projected by many arrangements of objects in the world, an action can be performed to achieve many different kinds of goals. The fact that our goal attribution system could answer this question easily in most situations is not an indication of the simplicity of the problem, but evidence for a very efficient goal extraction mechanism. (Csibra & Gergeley 2007)

Thus, they argue that action mirroring is insufficient for goal understanding. The question is why they assume that just because the goal of the action is understood on a broader contextual basis than the present action observation, they must hypothesize a mysterious extra-motor teleological reasoning goal extracting mechanism. I of course want to suggest that the broader fronto-parietal sensorimotor circuits play an essential role in monitoring and anticipating the overall affordance space and its relational and teleological structure – this might indeed prove to be "a very effective goal extraction mechanism."[340]

Csibra and Gergeley themselves discuss three different mechanisms of goal attribution that have been proposed in recent years - what they call 'action-effect associations,' 'simulation procedures,' and 'teleological reasoning.' The first one, 'action-effect

[340] In regard to the understanding of teleological properties and causality in general, I think it is interesting to notice the parallel between my psychological theory of affordances and abstract sensorimotor representations and the increasingly popular interventionist theory of causation proposed by Woodward and others (Woodward, 2003).

associations,' is rooted in James' theory of ideomotor action and action organization. They explain the view as follows:

> According to this view, the representation of actions in the actor's cognitive system includes and is related to the representation of their desired distal effects, and these inherently related action-effect representations are linked to each other through bidirectional associations. Thus, the 'idea' of the goal state, i.e., the desired effect, automatically activates the corresponding action, while the activation of an action elicits the anticipation of the distal effect associated with it. These links are established by simple associations upon observing the effects that one's actions have produced, and these associations start to build up from early on in infancy. (Csibra & Gergeley, 2007, p.68)

They emphasize the idea of 'bi-directional' links between actions and action effects. I think that the implications of a symmetric associative relation between motor actions and their effect might be slightly misleading in that the focus is taken away from the third necessary element of intentional actions, which is the environment acted upon. There could be various reasons for the narrow focus on a bidirectional association between action and effect, but one might be that the authors want to talk about action goal attribution and understanding based on action perception:

> An interesting extension of this approach of action control is the idea that perceiving and understanding other people's actions are also based on similar bidirectional associations between actions and their effects (e.g., Elsner, present issue). Thus, observed actions are proposed to be represented by being linked to the effects they have been seen to bring about (rather than, for example, just in terms of their antecedent conditions). (Csibra & Gergeley, 2007, p.68)

They refer to evidence that infants expect similar outcomes from similar goal-directed actions but that this action-effect association does not happen as a product of watching unintentional actions. They suggest that this sort of associative learning can be useful for goal attribution and action anticipation in so far as the observed actions are routine actions performed without much variation over various contexts. Hence, Csibra and Gergeley think that this ideomotor mechanism does not really solve any of the inverse problems of goal attribution and action anticipation. The focus on the bidirectional association at the expense of context evaluation is in my eyes presenting a particularly simplistic version of the ideomotor theory, which reduces the possibly broader explanatory power of learned sensorimotor integrations. One might ask how to even make sense of this narrow bidirectional version of the ideomotor theory. If the context and the preceding perceptual conditions are not taken into account I do not really understand how we could ever talk about ideomotor action initiation. Thus, it seems that the very idea of ideomotor action depends on association not only between actions and their sensory consequences – but more broadly the sensory change associated with the action. In other words, one might characterize the sensorimotor association grounding ideomotor actions as having to do with teleological and causal powers of actions in relation to temporally extended perceptual contexts.

Csibra and Gergeley then refer to simulation theories of goal attribution. Their view is here as I have discussed above, basically that simulation can only go from 'goal-to-action' and that simulation therefore cannot infer the goal but relies on an external mechanism to come up with a goal hypothesis. The goal hypothesis is then the 'input' of the simulation mechanism. The 'output' is the next most likely motor act, and the simulation procedure can therefore be used to predict and anticipate specific actions and movements in so far as we share the motor repertoire with the observed agent. If there is a mismatch between our simulated/emulated prediction and the observed sensory feedback we might use the mechanism to learn and adjust our motor commands. However, all in all the claim is that mirroring emulation cannot give us action or goal understanding but relies on extra-mirror mechanisms for this conjecture. As I discussed in relation to Gallese and Goldman's mirroring simulation theory, a narrow non-contextual action mirroring processes does indeed seem to hypothesize rather than produce the goal. In the later 'chains of action' version of mirroring the idea seems to be that the specific action chains are activated not just narrowly by observed actions but by action in contexts. Further, the idea is that the anticipation of the next most likely action often reveals the overall goal. In the Foggasi et al. 2005 study for example, the idea is that based on the most activated action chain we understand that the goal of the action is either eating or placing. Csibra argues that in order to choose the right pre-wired chain to activate based on the context, we must somehow already have inferred the goal by other means. This argument is a clue to Csibra's implicit use of a computational model of simulation as initiated at one specific level of action specification and then serially unfolding in a rather encapsulated manner. I on the other hand do not see why one would assume that our sensorimotor experience is not addressed by the perception at multiple levels simultaneously and on-goingly. As mentioned, I am not too fond of the idea of 'pre-wired chains' as I would suggest that action sequences are specified and customized with respect to the concrete context and therefore are not totally 'pre-wired.' However, I do think that Gallese's intention with this model is to posit that various 'chains' and actions are simultaneously addressed by the perceived action; the most likely next action and action goal is thought to be represented by the most modulated group of neurons. Thus, the idea of 'chains' seem to be an ill-chosen metaphor for a much more complex pattern of modulation and facilitation including not only one specific 'most likely next action' but also activations of less likely but somewhat probable next actions. The important difference is therefore that Csibra insists on an exclusive action chain choice prior to the simulation process and presumably to any sensorimotor or mirror activation. Gallese on the other hand seems to suggest that multiple chains are co-activated and one wins out so to speak. This might be what happens when we actually try to 're-enact' or overtly specify an action into a series of movements. However, my proposal is that one should quit the terminology of pre-wired chains, in particular when it comes to simply perceptually induced sensorimotor activations as in the case of mirror neurons and canonical neurons. I think it would be better here to talk about a spectrum of schematic simultaneous sensorimotor activations representing the observed action and the broader affordance space. Upcoming actions are anticipated not based uniquely on a specific observed action or specific context but on the dynamic and relational affordance structure that these yield together. Given that such affordance space understanding already is in place, we might in some cases simulate the realistically or wishfully projected actions of others in greater detail in light of the perceived affordances. My point is that *not all* perception induced sensorimotor activity should be seen as such full action simulation. Rather, the precise specification of such re/pre-enactment depends on an already sensorimotor engaged perception of the affordance structure.[341]

With this idea of sensorimotor affordance structure monitoring in mind, let me present Csibra and Gergeley's last proposal for a 'goal attribution mechanism,' namely what they call 'teleological reasoning.' I think teleology seems appropriate to goal attribution, but the problem is to decipher what this cognitive description of the 'mechanism' would

[341] Based on findings in songbirds, Scott Grafton has written an interesting article relating to this issue of what is stored and what is specified in the concrete context. He suggests that motor sensorimotor skills are malleable, in the sense that there is a rather rigid stored core template that is then filled in and adjusted with use. (Grafton (2008)).

translate into on a neurological level and how more concretely it would be different from what is suggested by the ideomotor and motor simulation theories. Csibra and Gergely talk about including background information about "the physical constraints of the situation and the actor:"

> What is required for this task is a causal analysis of the affordance structure of the observed actions and the artifacts they involve in order to recover which elements of those actions (and artifact use) are necessary and sufficient for producing the desired effect. Teleological reasoning could provide exactly this kind of analysis. (p.72)

Interestingly enough we see here that Csibra and Gergely even use the term 'affordance structure.' Such affordance structures might not be included in simulation or ideomotor theories per se, but the question is whether it is not reasonable to hypothesize that this sort of pragmatic knowledge in some way is linked to sensory-motor integrations. In this article the motor role in teleological reasoning is not excluded, but, as we have seen, in Csibra's other articles he clearly argues that action goals are inferred on a 'purely perceptual' basis. Interestingly, in this article it is suggested that the three proposed mechanisms are not necessarily mutually exclusive theories but they are rather posited as complementary cognitive processes that each serve their purpose. Each proposed mechanism does seem to have its strengths, but rather then adopting them in an additional form, one might want to reinterpret the hypothesized mechanisms and possibly integrate the insights as features of interrelated processes. As we saw in Chapters 2 and 3, a part of the original mirror mechanism idea is that this process yields action knowledge exactly because of action-effect associations. Further, the question is whether ideomotor action initiation and goal-centered organization of actions in general make sense in isolation from a broader contextual understanding. Differently put, can action-effect knowledge be understood in isolation from the contextually based change brought about? And then there is the issue of how affordance structures, goal and action possibilities and their causal interrelations are 'picked up' in one's perception of the environment.

Csibra and Gergeley note that "observing successful goal-directed actions can also inform us about the causally relevant properties of actions, objects and situations." (p.63) I think this is true, but I also wonder how one understands these causal relations and properties. Csibra and Gergely do themselves often use the terminology of affordances, and they talk about tools as 'frozen goals.' Somehow we represent the causal, relational and intentional structures of the environment around us. I share the sense that this sort of relational representation is essential not only for our understanding of others and the world around us, but also for our own ability for suitable action goals and plans. Again, the question is how and why Csibra thinks that we get to understand goals and affordance relations by way of purely perceptual analysis. As mentioned, I think that the 'why' rests on a narrow conception of motor cognition as covert action simulation. And the 'how' is left in the dark, or rather in the esoteric haze of 'pure perception.'

5.5. Chapter summary and conclusion

With this chapter I hope to have brought some of the pervasive cognitivist assumptions in regard to both motor and social cognition to the fore, but also to have shown how they have a hard time co-inhabiting the world of mirror neuron and broadly sensorimotor integration findings. By way of the critique I further advanced and detailed my positive project of creating new post-cognitivist frameworks, which better fit the empirical findings.

I started by presenting Jeannerod's view of motor cognition as action minus execution, and discussed the problematic nature of this narrow idea of motor processes as it excludes the possibility of parallel sensorimotor activations. Motor cognition is seen as simply actions in the head and under the same constraints as overt actions. I suggested that in opposition to this 'one body in the mind' conception it is important to distinguish between various levels of abstraction and specificity within the broader category of 'covert motor activity,' and particularly between schematically stored representations of for example goals and fully specified spatiotemporally simulated covert actions.

I further analyzed how this view of the motor system in general neglects the functional importance of sensorimotor integrations in the organization of our cognitive processes. I highlighted this problem with a discussion of the dualism between motor and perceptual representations that seems inherent not only in Jeannerod's view but also in the mirror neuron debate and terminology in general. I suggest that this dualism is driving, among other things, the misguided debate about social perception as either motor or purely perceptually based, and the idea of motor simulation as pertaining uniquely to overt observable behavior. I then sketched how alternative understandings of sensorimotor cognition might exactly suggest the importance of these processes in goal representation and for intentional action. I pointed out that the dualism between sensorimotor processes and the idea of intention choice as happening prior to motor involvement might be central to the intuition that mental states are hidden 'behind' the observable action. The point is that if sensorimotor processes can be seen as grounding more abstract mental representations such as goals – as the empirical research of the Parma group and others precisely indicates - then it is not so obvious why mental states should not to some extent be portrayed in our overt engagements. The dichotomy between overt behavior and minds as inner and hidden does not seem as neat and clear-cut as often assumed.

As to Jacob and Jeannerod's critique of the mirror theory - or, as they call it, the 'motor theory of social cognition' - it is exactly the idea of motor simulation as a full reenactment of the overt action that drives the argument. This critique is additionally married to the core assumption that social cognition has to do mainly with attribution and prediction of hidden mental states. A strict and unbridgeable dichotomy has thus been created between observable behavior – or the covert simulations thereof – and hidden mental states that are inherently underdetermined by the overt action. And, given these assumptions, the argument is simply a logical exercise reiterating the assumed gap. I use the discussion of Jacob and Jeannerod to interject with my alternative proposals and empirical findings, which I think effectively undermine their assumptions regarding motor and social cognition.

I then turned to Csibra's critique, which is based on similar frameworks. His views bring out important issues of how we are supposed to understand intentional action choice and the role of sensorimotor integration. By way of a discussion of his strong input-output information processing assumptions, we are brought to the issue of whether cognitive functions must be one-directional, serial, encapsulated or functionally exclusive. I use this debate to further argue that normal intentional action specification becomes hard to understand within such frameworks. I thus try to build the argument that we need a theory that integrates multiple automatic simultaneous sensorimotor activity with top-down endogenously controlled action organization. Furthermore, I argue that the idea of a sensorimotor based affordance space monitoring function might – in contrast to Csibra's own intuition - represent an answer to the question of how small infants can understand teleological and causal relations beyond their motor repertoire.

I have now discussed all the ingredients, I will in the last Chapter try to bring all of them together and recapture my alternative post-cognitivist interpretation of mirror neurons.

Further, I hope to show how this interpretation can change many debates within the philosophy of mind, cognitive science and also shape experimental protocols in actual scientific laboratories.

Chapter 6

Implications of a sensorimotor based social affordance model

6.1. Recapitulation of the overall project

I have now over the course of the many preceding pages sketched how mirror neuron research has systematically been misinterpreted in ways that reflect lingering problematic notions of the nature of motor cognition and of social cognition. Mirror neurons have on the one hand been hailed as the empirical answer to what in philosophy is known as the 'other minds problem' and as providing some sort of neurological link between self and other and, therefore, the basis for various social cognitive capacities, from action understanding and empathy to language. These conclusions regarding the mechanism and cognitive functions of mirror neurons have on the other hand been under intense fire as blowing the social importance of these neurons out of proportion. The core argument runs as follows: Since action mirroring is a motor process it cannot give us much in terms of 'true' social cognition, which pertains to mental states that are by definition exactly beyond observable behavior. In other words, due to the assumption that social cognition is 3rd person mindreading and motor cognition is limited to kinetic output coordination, action mirroring is concluded to have to do with basic motor intentions and concrete action prediction at best.

I have extensively pointed to the empirical problems with the idea of a 'mirror mechanism,' which during action perception produces a covert simulation or agent-neutral motor copy of the observed actions independently of the pragmatic context. But my project is not one of simply defaming mirror neurons or simply raising various skeptical challenges to the existing interpretations of their social function. My project is, rather, to de-center and reinterpret these neurons in their broader neurological and pragmatic context, and via this reinterpretation spread the fame and excitement from the narrow focus on mirror neurons to fronto-parietal sensorimotor integrations more generally. I do think that something extraordinary has been discovered in this field of research and something that is important not only for our understanding of social and motor cognition but also for a number of core issues in the philosophy of mind, such as other minds, intention, action choice, teleology, language and the concept of mental representation itself. Thus, I am not at all trying to tone down the hype (as some might have hoped for) but to redirect it towards the study of processes of sensorimotor integration more generally.

More concretely, I propose that the fronto-parietal areas in question are essential for monitoring and anticipating the relational structure of what I call our ‘affordance space,’ i.e. the action invitations and anticipated possibilities persisting and changing in our perceptual environment. In this broader function I hypothesize that mirror neurons play various roles in our understanding of the actions of others as they relate to our social affordance structure. My arguments for this interpretation are based on empirically problematizing the above-mentioned notions of social cognition as being about hidden mental states and of motor cognition as pertaining purely to actual motor outputs. I also attempt to show that the sensorimotor research, of which mirror neuron research is a part, is exactly pointing to the need for different theoretical frameworks for both social and motor cognition.

I focus on the notions of teleological/goal representations and affordances. These notions seem crucial to conceptualizing the processing in fronto-parietal circuits, and bring together perceptual, motor and teleological aspects in one inherently integrated neurological phenomenon that might be thought of as pre-intentional. The focus on affordances and sub-personal level sensorimotor-based teleological representations thus presents a new empirical angle on grand questions of the philosophy of mind, which has typically been posed from a personal-level and often too static and non-dynamic point of view. I argue that the motor system research discussed calls for a story of goal understanding and intentional action choices that incorporates relational and pragmatic mental ‘content’ and dynamic temporal and hierarchical sensorimotor processes. I suggest that the sensorimotor findings warrant a changed view of motor cognition and further that this leads to a different view of social cognition as well. There are core social cognitive processes, which do not involve mindreading in the sense of some mental process that leads us *from* action observation *to* attribution of hidden mental states. Much social cognition has to do with understanding others in relation to the ‘affordance space’ and the social affordances their actions yield in regard to the perceiver. In other words, thinking of action perception in sensorimotor terms suggests that much social cognition happens already *in* perception and in our teleological (and emotional) understanding of the social affordance space[342].

My own conclusions regarding the importance of fronto-parietal processes for social cognition and intentional action choice are therefore based on rather different notions not only of what is going on, but also of what motor and social cognitions do and what these abilities consist in. In short, I want to suggest that the empirical findings of visuomotor neurons – of which mirror neurons are but one little and often idealized class – are indeed very important, but for radically different reasons than those typically suggested.

The method of my project is to diagnose the problems with the existing research and the debate around it, and show why certain assumptions are empirically and theoretically misguided and misguiding, and, via this discussion, to advance my own alternative hypotheses not only about mirror neurons but more generally about sensorimotor integration and its role in motor and social cognition.

Consequently, an essential part of my project is to show that the idea of the motor mirror mechanism itself is ripe with systematic and problematic theoretical assumptions. As already mentioned, a core theoretical problem is that motor cognition is mostly thought of as simply pertaining to an output system producing either internally or externally driven motor commands, and social cognition is thought of mainly as the ability to 'mind-read,' i.e. to attribute and reason about the hidden mental states of others. However, another big problem is the idea of mirror neurons as constituting their own functionally semi-independent mechanism. This relative and often tacit modularity bias, I argue, hinges on the conceptualization of mirror neurons as producing a symmetric response for action observation and execution conditions, and thereby the idea of mirroring as agent neutral and context independent. I have repeatedly exposed the empirically unwarranted nature of this caricature view of mirror neuron findings. And if mirror neurons as a class cannot so neatly and distinctively be categorized, what is the justification for assuming that they

[342] It should be noted that the idea of teleological and affordance perception might also the other way around force us to rethink our notion of perception – particularly the assumption that we only perceive what is concretely present at any given point in time. For example affordance perception reveals possible futures for an organism in the present – and we are thus dealing with a more complex dimensional and relational notion of perception than simply detecting the present external phenomena. However, this is just meant as a reminder as dealing with the general issue of perception is well beyond my present focus.

constitute a functional 'mechanism' apart from the overall sensorimotor circuits in which they are anatomically located? In opposition to the received view and the terminology of the Parma group, I hypothesize that it is exactly in this broader neurological context that their function should be understood. I sketch my alternative affordance space hypothesis via the concrete empirical findings of sensorimotor integration in fronto-parietal areas and the relation between the activity in these areas and what one might call teleological representations of action goals and affordances. I use empirical means to push the theoretical validity of the core idea that monitoring the affordance space might be functionally a key not only to motor command initiation but also to the relatively abstract representations of goals and intentions, intentional action choice itself and to social cognition, i.e. key to the understanding of how others perceive, think and intentionally act in regard to the shared affordance space. Hereunder, I have identified the need for an understanding of sensorimotor activity that is different from and independent of singular concrete and fully spatiotemporally specified actions. I suggest that fronto-parietal neurons like mirror neurons should rather be thought of as providing parallel schematic sensorimotor integrations and, further, that they provide us with an example of sub-personal 'mental representation' that defies the earlier behaviorist and classic cognitivist frameworks.

For me, the research and heated debate over mirror neurons and their function has thus provided an ideal empirical Petri dish not only to propose a new theory for this particular and enormously exciting field of study, but also to embark on a philosophical project of approaching and beginning to ground a new cognitive framework that does not assume a strict dichotomy between the mental as inner and observable behavior as outer and between actions as motor outputs on the one hand and perceptions as inputs one the other. This is obviously a somewhat over-ambitious project and, in many ways, I see this thesis project as initial groundwork for a later more developed version of my affordance model and its broader philosophical implications. In this concluding chapter I will simply pull out and reiterate some of the arguments of the main chapters that I see as crucial and then point to the many possible implications that need future empirical and theoretical attention well beyond the present project.

6.2. A sensorimotor 'affordance apace' re-interpretation of mirror neurons

I purposefully started Chapter 2 with an account of the sensorimotor research that had served as a backdrop for the discovery of mirror neurons. The key feature of this research for my purposes are that it seriously challenged the traditional conception of the motor system as a unified output system simply translating intentions into hierarchical motor commands.

6.2.1. A new route from motor to social cognition

Here I will focus on 3 elements of the research of the early Parma group, which seriously challenge not only the idea of the motor system as a unified output system, but which also serve as key features in my critique of the assumption that mental states should be seen as necessarily hidden or in strict opposition to observable behavior. This is pivotal as the philosophical 'other minds problem' and the idea of social cognition as 'mindreading' generally assume that mental states are hidden and that they somehow have to be inferred or derived from behavior but 1st personal and inner states radically differ from observable 3rd person behavior (see Chapter 4). My view of the motor role in social cognition thus hinges on an argument that first tried to broaden the notion of motor cognition and show how social perception and intentional actions depend on sensorimotor processes. It is by way of this focus on teleological representations and their role in our pragmatic evaluations and action choices that I also challenge the idea that minds and other minds should per definition be understood as inner and not outer, and hidden as opposed to observable.

The 3 basic elements of the Parma group sensorimotor findings that support my argument are the following:

1. **Goal modulation.** Firstly, I have pointed to the fascinating discoveries of goal modulation in these premotor areas relatively independently of the exact kinetic movements performed, which shows some level of abstraction away from concrete motor commands in premotor areas.
2. **Affordances.** Secondly, these types of fronto-parietal sensorimotor circuits are found to often reflect a pragmatic functional integration, and I have given various examples of visuomotor neurons that could be characterized in terms of some sort of teleology or 'affordance' integration of appropriate actions given certain perceptual features.

3. **Parallel sensorimotor integration and function.** Thirdly, there is the finding of massively parallel sensorimotor integration, suggesting that the motor system cannot be seen as either functionally or anatomically neatly defined or separated from parietal areas of sensory integrations and 'pre-intentional' attentional biasing.

I have suggested that these findings together seriously challenge not only the idea of motor cognition as pertaining to a unified output system, but, more radically, they put a dent in the idea of separate and 'purely' motor and perceptual representation of goals. This is of course a very important point as the debate on the role of mirror neurons in social cognition normally is construed as one of whether these can yield a motor understanding of goals or whether goals are understood in purely perceptual or extra-motor areas. (See particularly Chapter 5) I hypothesize that goal representations and thus understanding might very well fundamentally be dependent on these sorts of relatively abstract pragmatic sensorimotor integrations, and thus that neither a purely motor, nor a purely perceptual account would be appropriate for such teleological understanding[343]. I have tried to show how the concept of 'affordances' is useful not only in describing the empirical findings of what happens in many fronto-parietal circuits, but also is interesting in that it conceptually brings sensory and motor aspects into one common process and function. Each affordance modulation is inherently both related to tracking the perceptual environment and potential action goals in relations to this environment at a given level of abstraction. This focus on the basic organizational priority of sensorimotor representations is very different from the typical interpretations of mirror neurons research that describe these neurons as 'translating,' 'mapping' or 'linking' sensory goal representations to motor goal representations. (See Chapters 2 & 3). The traditional Parma group interpretation is based on the idea of a 'motor vocabulary' of action types and it emphasizes the relative abstraction from motor kinetics but ignores the potentially fundamental role of sensory processes in the organization of this action vocabulary as abstract in the first place. In short, this interpretation proposes a motor theory of social

[343] As have been pointed out to me abstract and pragmatic rarely goes together – but this combination is exactly what fronto-parietal sensorimotor neurons seem to support. The abstraction comes from the finding that that the modulation depends on goals and affordance relation at more general levels than kinetic movements and specific stimuli. The pragmatic nature refers to the normativity and action determining element of the fronto-parietal circuits. In other words, the finding that these neurons seem to represent teleological relations as they relate to the perceivers own action choice – thus that they monitor the concrete affordacnes of the present environment.

cognition where I propose a sensorimotor theory. This is important as many critiques of the role of mirror neurons in social cognition are based on the idea that motor cognition alone does not give us much in terms of intentions. (See Chapter 5)

I further propose that *parallel* sensorimotor goal and affordance representations in some ways challenge the notion of intentions as notoriously hidden mental states thought to be causally prior to the involvement of the motor system, as chosen and represented independently of sensorimotor processes and as 'sandwiched' between input and output systems. I have suggested via Erik Rietveld's analysis that such parallel goal representations inform an intentional or pre-intentional 'field of affordances,' and that it is due to the overall relational affordance space understanding that we can appropriately chose intentional actions in the first place. This idea relies on an additional feature that I argue is empirically backed by later mirror neurons studies. This feature is namely that goals and affordances are not simply represented and catalogued as 'types,' but understood in *relation to* other teleological features of the environment and ongoing actions. My alternative proposal of a 'social affordance model' understanding of fronto-parietal circuits directly builds on these ideas and findings. Thus, I argue that we have an intentional involvement with the world already in our perceptional pre-action relation to it – and it is this involvement along with background considerations that inform and shape our actual action choice.[344]

In regard to social cognition the consequence is that even though intentions might often not be transparent in the individual actions they produce, that does not entail that the background considerations or intentional involvement in principle are hidden or opposed to what we can observe and communicate in overt relational and behavioral engagements. In other words, a sensorimotor grounding of intentional action choice points to some level of transparency of the intentions of others in their broader action engagements. However, it is also an important consequence of my theory that each intentional action and personal level experience to some extent transcends the actual present observable behavior. The point is that even though action intentions in general are relational and related to past and possible observable engagements, actual singular choices are normally

[344] Elisabeth Pacherie has over the years made some very insightful and useful analyses of the dynamics and hierarchical structure of intentional action, that does not fall into the traditional trap of sandwiching intention between action and perception (Pacherie 2008, Pacherie & Haggard, 2010).

based on both experiences and considerations that are not fully determined or given in the present *resulting* behavior even when considered in its spatial pragmatic context. My theory should here be seen in contrast to the typical focus on singular intentions and their causal relation to a singular action outcome – and the resulting conundrum of how we can infer the former from the latter. In essence, my point is that we do not simply post hoc infer singular intentions from singular actions. Rather, we see how the action is chosen from alternative affordances, which tell us things about the priorities, skills and experiences of others that go far beyond the 'why' or 'prior intention' of their present observable action.

My proposal is not a self-sufficient or complete argument against traditional notions of mental states and representations. Rather, what we see in Chapters 4 and 5 is that my theory, through the analysis of sensorimotor integrations from earlier chapters, comes to question specifically the idea of mental states as necessarily hidden and 'beyond' or 'prior to' observable action, and at the same time accepts the need for a mental category of internal representation that goes beyond conjoined sequences of stimuli and responses. The approach I propose is research into abstract and parallel sensorimotor representations and the relational and hierarchical knowledge these give us in regard to both our own actions and those we see others choose.

6.2.2. The asymmetric and heterogeneous 'mirror' neuron findings

In regard to the history of the discovery and interpretation of mirror neurons developed in Chapters 2 and 3. I have pointed to the many empirical facts, which challenge the often tacitly held 'caricature model' of mirror neurons as informed primarily by the metaphor of 'mirroring' and theoretical assumptions. The key elements of what I have called the 'caricature model' of mirroring are:

1. The idea of mirror symmetry and agent-neutrality
2. Categorization via action-types as 'words' in 'motor vocabulary'
3. The ubiquity and non-contextual nature of the response
4. Semi-modular functional mirror mechanism

Accordingly, the overall idea of the 'mirror mechanism' becomes that it produces a rather modular or pragmatically insulated agent-neutral action type representation, which matches the observed action or action goal. It seems to be a misleading idealization of the findings to say that all action sensitive visuomotor neurons are modulated in a 'mirror' fashion, and therefore also problematic to suggest that they together form an observation-execution matching mechanism. It is not simply that different neurons are tuned to different goals and sub-goals and temporal segments of different actions; this diversity is indeed accounted for in the standard interpretations of mirror neurons. The problem is, rather, the very guiding metaphor of this field of research. Namely the issue that the name 'mirror' appears to pick out neurons that can be seen as 'mirroring' and that the metaphor overshadows the more complex properties of action sensitive neurons and the similarities of these neurons with others sensorimotor neurons in the studied areas. To illustrate why this is a problem I point on the one hand to the empirical data already available that shows agent specificity and context and task dependence of 'mirror' neuron responses. I.e. the findings that some so-called mirror neurons respond differently to one's own and others' actions, and that many mirror neurons do indeed respond differently given the practical and motivational context. (See Chapter 3) On the other hand, and for me at least equally importantly, I point to all the missing data. I.e. the experiments that were simply never done and the categories that were never accounted for due to the tacit force of the mirror metaphor, the ideas of observation-execution 'congruency' and of 'action type' representation. Thus, I have hypothesized an extensive array of action sensitive visuomotor neurons – well beyond strictly and broadly congruent mirror neurons. (See Figures 3.1-3.3 in Chapter 3) Most crucially, I propose that given what is already known there is no reason not to predict the existence also of social affordance neurons – i.e. action sensitive visuomotor neurons that respond not symmetrically but rather integrate a perceived action with an appropriate action response.

6.2.3. Canonical neurons and social affordance relations

If one sees the heterogeneity of the mirror neurons findings in their neurological context of the previous knowledge of the fronto-parietal areas in general and F5 in particular then it seems pivotal to ask why their function should be seen as independent of so-called 'canonical' neurons. Canonical neurons are classified as modulated respectively by perceiving objects and executing goal-directed actions towards such objects, i.e. as monitoring object affordances. I suggest that both canonical and the broad range of action sensitive visuomotor neurons – hereunder classical mirror neurons – should be understood as part of the same functional 'mechanism.' I.e. that they should all be seen as informing our representations of the structure of the physical and social affordance space of the perceiver. But it is important to emphasize that my proposal is not simply that there two types of a parallel monitoring going on of respectively action types and object affordances. On the contrary, I think that there is an important and dynamic integration of information between all the various types of visuomotor neurons in the circuit such that we do not simply get an itemized representation of actions and objects but precisely understand their mutual actual and potential teleological relations. In other words, it is not just that canonical neurons are important for the pragmatic action functions of these areas in regard to one's own action choice, but also that the social cognitive function of these circuits should be understood as based not simply on mirror neurons but on the dynamic relations between such action sensitive neurons and canonical neurons. I find the occlusion study by Umilta and her Parma colleagues very informative in regard to this point.[345] The authors themselves are focused on the finding that mirror neurons seem to track the inferred actual hand–object interaction rather than simply the visually perceived motor movement. The additional point that I stress is that in order to track this unseen hand–object interaction, these 'mirror' neurons must be dynamically regulated by canonical neurons, which in turn monitor the presence or absence of the unseen object. Thus, the important information is carried by way of a *relational mapping* of the social and physical affordance structure.

6.2.4. The cognitive function of mirror neurons & fronto-parietal circuits

So, what exactly is it that I claim mirror neurons and fronto-parietal circuits can and cannot do for social cognition? Based primarily on single cell findings from macaque monkeys, but also consistent with behavioral and imaging studies in humans, I hypothesize that mirror neurons should be understood as dynamically and functionally interdependent with other categories of sensorimotor neurons in the fronto-parietal circuits. I propose an 'affordance space' model of the overall functioning of these fronto-parietal areas, and suggest that various kinds of 'mirror' or action sensitive neurons play different monitoring and predicting roles in regard to our understanding of this affordance structure and our own and others' teleological place in it. Thus, I concede that there

[345] Umilta et al. (2001). For details and a discussion of this study, see Chapter 3, Section 3.2.2.

might be certain 'strictly congruent' mirror neurons that rather narrowly might reflect action goals or sub-goals relatively independent of context and even independently of the agent such that these neurons respond relatively symmetrically to action observation and execution. However, I underscore that most neurons and the circuits overall do *not* show such symmetry. Further, the lack of symmetry seems essential to the overall function of affordance space monitoring, such that the other's action is understood as different from and relating to ongoing and potential actions of the observer. A good example here is the recent finding that certain mirror neurons respond differently to action perception depending on whether the observed action takes place inside or outside the peripersonal space of the monkey.[346] This is a very significant finding as it shows that the modulation portrays not simply the presence of an 'action kind' but the pragmatic relations between the observed action and the observer and her affordances as an agent in her own right. Furthermore, one should note that the neurons studied in this experiment, in spite of the obvious asymmetry, under classic categorizations would most likely be regarded as 'strictly congruent,' in that they in many contexts respond to the same specific goal directed motor act independently of whether it is executed or observed. In other words, because the focus has been on finding 'shared representations' of 'action types' the functional importance of asymmetries has been systematically ignored. It has normally been hypothesized that mirror neurons, via an agent neutral action representation, can be used to understand others. The story is that during action perception we simulate or covertly reenact the shared representations and then via a 'who-system' external to mirror neurons themselves are able to attribute the simulated action and its intentions and goals to others. (See Chapters 3 and 4) However, if there is no such asymmetry in the response of mirror neurons of fronto-parietal circuits then this simulation account seems to need fundamental revision. I suggest that the social function of mirror neurons is not that they let the observer into the head of the other or by way of neural symmetry at an experiential level that lets us share the mental representation of the intentions behind the actual action. Rather, my interpretation suggests that the broader sensorimotor integrations of these circuits let the observer understand and predict not only the actual goal-directed actions and teleological relations, but given representations of the relational affordance structure we can understand something about the potential actions and action choices of others in

[346] Caggiano et al. (2009). For details and discussion of this study in chapter 3.

our affordance space. Thus, mirror neurons themselves might simply say "grasping," but the broader information might be something to the effect of "she wants the peanut over there and if she gets to the lever with the stick she has grasped then she will get it before me – I'd better hurry up, because she is not looking at me and I can reach it first if I steal her stick…" etc. The point is that the relational affordance structure carries information about potentialities and teleological dynamics that cannot be understood by way of strictly congruent mirror neurons or an observation-execution matching mechanism. And, it is the understanding of the ongoing intentional engagements, deliberations and choices of the other via an observation of actions in a shared affordance structure that in my view carries the promise of a significant contribution to social cognition, not narrow simulations of perceived occurring actions.

Going somewhat beyond the current project I would want to hypothesize that it is exactly the contribution of potentiality and 'counter-factual futures' in perception that makes such sensorimotor processes important for what we typically call 'higher cognitive processes.' And, when I hypothesize that certain higher cognitive processes can be grounded in sensorimotor representations, it is exactly this sort of parallel counterfactual but experience based pragmatic integration that I have in mind. But what do I mean by 'counter-factual futures' and how does this relate to affordance structures and visuomotor neurons? I will try to unpack this claim by putting it differently. The sensorimotor findings seem to suggest that simple action and object perceptions are already cognitively interpreted via parallel affordances and teleological predictions. Affordances and predictions are about possible future events and the pragmatic allure of such anticipated and present possibilities. I claim that it is by way of the parallel and pre-intentional sensorimotor mapping of potential goal-directed actions and their yet counter-factual outcomes that we hierarchically can coordinate and choose intentional actions. And the further speculative step is that if this level of hypothetical, teleological and normative integration is happening in normal perception and concrete action choice, then why not hypothesize that hypothetical and counterfactual thinking to some extent relies on similar measures? Or that language and linguistic meaning for that matter can be understood as dealing with abstract affordances and that communication happens via the continuous creation of a shared imagined/pretense affordance space? Again the aim is not to argue for these larger implications, but simply to gesture at these possible theories and thereby the potential broad ramifications of changing our cognitive frameworks from a classical sandwich model to a sensorimotor based understanding of certain mental representations.

6.2.5. The limitations of the fronto-parietal circuits

Notably, these gestures towards a larger framework are not meant to say that fronto-parietal circuits are all there is to say about social cognition or even goal-directed action perception and intention understanding. I have repeatedly admitted to my problematic bracketing of emotional and motivational factors, which are importantly informed by sub-cortical structures and medial reward systems. It is clear that such processes play an important role in organizing our goal understanding and also the normativity inherent in the affordance concept. I have also been mentioning the issue of prior information and endogenously controlled behavior that also very much depend on pre-frontal regions as we saw in the discussion of imitation and utilization behavior. Furthermore, in connection with the false belief tasks I have discussed the importance of the ability to 'decouple' and abstract from one's own concrete affordance structure to reason from the perspective of others – the right temporal-parietal junction seems to be particularly crucial for this attention shifting and abstraction ability. Additionally, I am intrigued by the goal and gaze/attentional direction related activity found in the superior temporal sulcus (STS). Certain areas in the STS have been found to contain neurons that are modulated by perception of biological motion and goal directed actions, but interestingly it seems to be uniquely the perception of the actions judged to belong to others and not the perception of one's own executed actions that modulates STS neurons.[347] If this is the case, then these areas might integrate information that is distinctive to the other – as opposed to relational teleological information, which again fits with the lack of spatial mapping in the STS.[348] All these other contributions to social cognitive abilities are beyond my present scope, and the point here is that I do not deny the significance of other areas, but simply underline that something importantly social and importantly 'mental' is indeed going on when we, on the basis of fronto-parietal sensorimotor circuits, monitor and predict the physical and social affordance space we share with others.

[347] See for example Allison et al (2000) and Puce & Perrett (2003).
[348] A further interesting element here is that such representation of the actions of others in the STS seems to depend not only on visual information but also information or 'reference copies' of one's own initiated actions. See Iacoboni et al. (2001), and also Miall (2003).

And here a further major qualifier is that what I broadly call affordance space monitoring might even turn out to depend on extra fronto-parietal regions. My discussion focuses mainly on a specific fronto-parietal circuit in which mirror neurons were first discovered in monkeys. This circuit spanning the premotor F5 and the inferior parietal lobe has in monkeys been found to respond overwhelmingly to goal directed hand–object and mouth-object interactions. A mapping of, for example, afforded eye movement and a more spatial avoidance and orienting behavior seems to be happening in neighboring fronto-parietal loops. I will not get into copious details here, but the point is that there are many open questions about which sorts of affordances and goals depend on which regions. And, when we turn to the human research, even more is open to interpretation or simply unknown. It has been suggested that humans can represent more abstract goals, intransitive and communicative actions in fronto-parietal circuits.[349] Humans also seem to have the ability to see things as goals that monkeys could not care less about,[350] but it is for now not empirically settled exactly how and where these functional differences are rooted in neurological differences. The fact that Broca's area overlaps with the region hypothesized as the human homologue to F5 has inspired multiple hypotheses about the relationship between language and mirror neurons.[351] In relation to my affordance framework it is interesting that linguistic meaning and grammar might be understood as parsing and constructing a shared communicative affordance space beyond the actual concretely shared affordance space and thereby its own set of social affordances as well. Furthermore, there is the question of the role of the cerebellum, which is known to play important roles in sensorimotor coordination. Maybe the temporal predictions and the dynamic and ongoing modulation and updating of the affordance space monitored depend crucially on the cerebellum. This is naturally an issue of enormous importance as it has been found that the cerebellum is one of the most significantly altered structures in people with autism.[352]

[349] For review see Fabbri-Destro & Rizzolatti (2008)
[350] Michael Tomasello and his colleagues have done interesting work on the cooperative social abilities in humans and non-human primates suggesting that it might be the drive towards and inherent appreciation/valuation of social cooperation that supports the evolution of many higher human cognitive skills. See Tomasello (2008 and 2009).
[351] See for example Rizzolatti & Arbib (1998), Stamenov & Gallese (2002), Gallese & Lakoff (2005).
[352] See Courchesne (1997), Bauman ML, Kemper TL (1986) & (1990)

Many of these observations can be regarded as speculative, but, even though my theory does not depend on them per se, it should be informed by what the future empirical research will teach about the exact coordination, level of abstraction and sorts of affordances monitored in the human fronto-parietal circuits.

6.3. Critique of traditional interpretations of mirror neurons

My interpretation of mirror neurons and their possible role in social cognition should be contrasted with the standard array of mirror neuron interpretations that all to varying degrees stress the idea that mirror neurons are important for social cognition because they give us A) shared action representations or B) some sort of simulation of the mental states or experiences of other agents. The former versions are based on the symmetry of what I call the mirror metaphor or caricature model of mirroring as yielding an agent-neutral context independent action type representation. The latter versions are rooted in the idea that I in Chapter 4 on simulation have called '3rd via 1st person' simulative mindreading. The simulation theory proposes that we understand others by way of a 1st personal simulation of the 3rd person's actions or general situation – of how it would be to be 'in their shoes.' The '3rd via 1st person' expression is supposed to underline that it is still very much a view of social cognition that like theory theory proposals stress not only a view of the other as 3rd personal and somehow problematically out of reach or unknown, but also relies on a fundamental priority of the 1st person perspective. The expression further signals that simulation theories (maybe even more so than theory theories) ignore the '2 person' concrete intersubjective relation in social cognition.[353]
The core idea in both versions of the simulative mirror hypothesis is that the perceiver can know something about the hidden mind of the observed agent via some sort of mirror-matched representation/simulation. I have argued in Chapter 4 that the important difference between the two versions of the mirror theory is the move from a sub-personal shared representation to a temporal and personal level experienced mirroring. I have throughout Chapters 2 and 3 pointed to the vast empirical challenges to the symmetry view of mirroring. For present purposes, the most crucial problem of the latter simulation

[353] My alternative interpretation is that mirror neurons precisely play a role in regard to this relational 2nd person meeting, i.e. they reveal not only something about the other's actions and action affordances but also something about how their actions relate to me and figure in my overall present and anticipated affordance structure.

versions is that the *single cell* mirror neurons findings are now theorized through the prism of fully specified and temporal covert action simulation as yielding a shared qualitative/experiential aspect between agent and observer. I argue that this move raises new problems, primarily due to the lack of attention to the differences between single neuron activity and personal level actions.

6.4. Schematic sensorimotor representation versus full action simulation

My theory fundamentally depends on the distinction between, on the one hand, fully specified actions, be they covertly or overtly executed, and, on the other, much more schematic, learned and stored sensorimotor representations. To illustrate this difference, one could think of watching a tennis match, or even maybe a doubles match, either simply by following the point in order to see who will chose to do what, to see which openings will be created and exploited, etc. Or, as an alternative approach, one could single out one player, and narrowly study the footwork and racket preparation etc. I argue that in the latter case we probably covertly simulate and reenact many of the observed actions, but that in the former case we simply represent the actions and their teleological context at a much more abstract and schematic level. I argue that parallel representations of multiple such schematic forms of action and affordance representations are essential not only to action choice and timing but also to action understanding. The latter sort of action simulation is a different cognitive process with much tighter load constraints, which is confirmed by how we must not only stall our own ongoing actions but also ignore most contextual elements of the overall action space in order to simulate the other. I have proposed that such simulation has great functional benefits, for example motor learning, but that it probably does very little in terms of our regular social understanding. The main point is that the categories of schematic, stored and non-spatiotemporally specified sensorimotor representations often simply fall out of the picture in the literature on mirroring and motor cognition and, as we see in the articles by Jacob and Jeannerod, motor cognition in general is thought of as covert action simulation or action minus execution. This is a crucial problem that cannot be reiterated too many times – as I exactly suggest that mirror neurons should be seen as carrying such stored schematic sensorimotor information. Thus, the problem of talking about 'motor simulations' or 'embodied simulations,' as Gallese has been doing in recent years, is that the distinction between spatiotemporally specified actions and stored representations is denied, forgotten, downplayed or ignored. Due to the load constraints on fully simulated actions, this issue is also connected to the further theoretical issue of ignoring the many parallel sensorimotor activations and the monitoring and anticipation of both observed actions and their teleological contexts and affordances for respectively agent and observer. Gallese often reiterates the importance of such social and physical context information and inter-subjective relations, but as extensively demonstrated in Chapter 4 the thinking behind these insightful comments never really seem to make it into the models and terminology of 'actions chains' or 'embodied simulation' theories.

6.5. Sensorimotor findings beyond behaviorism and cognitivism

I have already emphasized the importance of discarding what I have called the caricature model of mirroring along with the influences of this framework and its metaphors on the whole field of research and the debate around it. A central part of my argument in this thesis project is that this caricature and the traditional interpretations of mirror neurons are entangled in much broader and deeply entrenched conceptual frameworks of cognitive functioning – and that the mirror neuron and sensorimotor research exactly challenges these very frameworks. While presenting the evidence for my reinterpretation, I have already mentioned the need for a new framework for motor cognition, which does not rely on a neat distinction between perceptual and pure motor representations, nor on the idea of a unified motor output system as opposed to a perceptual input system. (Chapters 2 and 5) As to social cognition I have pointed to the need for a framework that does not rely on the idea of a neat distinction between outer observable behaviors and inner hidden mental states and does not regard the primary social relation as one of 3rd person non-engaged and de-contextualized observation. (Chapters 4 and 5)

These resulting critiques of the traditional frameworks can naturally be seen as yet another embodied/enactive/situated/phenomenological/ post-cognitivist attack on the 'classical sandwich' model of cognition. I do indeed see my project as a contribution to this larger attempt at producing a paradigm shift in the cognitive sciences and in the philosophy of mind. However, as is typically the case in theoretical revolts (and as the many 'slashes' above suggest), there is a plurality of counter movements, foci and ideas that barely fits under the uneasy umbrella term of 'post-cognitivist' approaches. I think these are typically called approaches rather than theories exactly because these lines of development are united not by positive theory but by an attempt to break up with certain dichotomies and the non-dynamic and non-temporal nature of the classical cognitive and computational models of the mind. The big question is, of course, exactly which parts of the traditional framework one wants to dismantle and what one plans to replace them with.

6.6. Sensorimotor grounded mental re-presentations

Like many other scholars I have stressed the problem of assuming a dichotomy between perceptual and motor processes, between inner minds and outer behavior, between cognitive machinery as form and mental content as being about things and properties of the world.[354] In extension of the critique of these conceptual divisions, I have also pointed to the empirical problem of assuming sequential and modular functions as opposed to parallel, reciprocal and dynamic processes. But, in contrast with most other 'post-cognitivist' thinkers, I think that it might be a mistake to let go of the notion of mental representation altogether. I agree with many critics of traditional constructivist representationalist theories of perception and their objection that we do not seem to have inner representations as pictures that sort of re-double the external world inside our heads.[355] But, rather than letting go of the notion of representation, I think one might be able to fruitfully reinterpret the notion within a post-cognitivist framework. For me, as we have seen throughout my various discussions in this thesis, the point is to retain the idea of something 'stored,' which can be 're-presented' either via externally provoked or endogenously controlled modulation. The goal is to move towards a new framework that reaps the insights of recognizing the short-comings of traditional cognitive models but also of the recognition of the shortcomings of behaviorist approaches. This proposed

[354] Another pivotal dichotomy that has been attacked by post-cognitivists is the one of cognition versus emotion. My project suffers from the fact that I have skirted the issue of emotion and reward systems, which obviously is essential not only to a full understanding of intentional action choice and social cognition, but also to the further development of my social affordance model. See Damasio (1994) for a classic postcognitivist contribution to our understanding of the role of emotion in cognition.

[355] This anti-representationalist line of reasoning has many contemporary proponents in philosophers like Susan Hurley, Shaun Gallagher, Alva Noë, Dan Zahavi and a neuroscientist like Walter Freeman, but it can be traced far back through the history of philosophy. In my opinion, some of the most poignant early critique comes from Kant, Nietzsche, Bergson and the phenomenological and pragmatist traditions. See also the discussion of Akins below.

reconciliatory idea of mental representation has been sketched and discussed multiple times in the preceding chapters, and I see it as a potentially very powerful - but definitely at this point speculative – idea. Below, I will summarize the main gist of this idea.

The traditional notion of representation focuses on a *personal level* sort of special mental picture that carries a content, which is about *something* in the external/imaginary world, and has a form that reflects this *relational character* of an 'aboutness' link to the outer content. As Kathleen Akins has rightly pointed out, the traditional idea of the 'aboutness' of mental states is deeply rooted in a particular theory of sensory systems and perception, where these processes are thought to construe representational images of external 'things,' which then can serve as the content of other mental processes.[356] Akins criticizes the traditional view of sensory representations as a veridical source of presenting in the brain 'what is where' in the external world and argues that what we know about sensory systems suggests that they rather than veridical and bottom-up are highly 'narcissistic.' Akins' use of the qualifier narcissistic is precisely stressing the relational and pragmatic character of perceptual experiences and processes. What she calls the traditional idea of sensory representation is linked to exactly the sequential input-cognition-output model, the dichotomies between perception and action and between inner and outer, and, finally, between mental process providing the aboutness form and the world as providing the content to be represented. The problem of simply saying that cognition takes place without representation is that it seems to make it hard to exactly pinpoint how such an 'anti-representationalist' theory does not revert to a form of behaviorism that collapses the inner and the outer or obscures the complexity of the internally generated or shaped processes. The idea of grounding cognitive processes in learned and stored sensorimotor integrations is by no means new. Still, I think that my philosophically and empirically informed project presents new and significant contributions to existing theories of sensorimotor grounding.[357]

As I have tried to stress, in order to make sense of the central cognitive roles of such sensorimotor learning one needs to make some crucial distinctions between actual

---

[356] See Akins (1996).

[357] I think it is interesting to note the difference between a theory like mine that starts with a set of neurological sensorimotor processes as opposed to for example Lakoff and Johnson's embodied metaphor theory (Lakoff and Johnson 1999) and Barsalou's theory of modality specific symbolic content (Barsalou 2008).

spatiotemporally specified sensorimotor engagements and more abstractly stored sensorimotor integrations that can be 're-presented' without necessarily being 're-enacted.' However, mental representations are typically thought of as having a reifiable content that is *about something* or a property of the world, which is neatly distinguishable from the mental *form* of presentation as being mentally 'for me.' The focus on presentations and re-presentations is meant to change the focus of form-content, perception-action, inner-outer dichotomies in order to point to an important temporal distinction, but also the dynamic interplay of stored sensorimotor learning and actual perception and action.

Via the analysis of mirror neurons and sensorimotor integration during perception I thus suggest that there is a way of thinking of the *content* of mental re-presentations where it can be not only sub-personal but also relational such that it reveals learned, emotional and pragmatic relations between an organism and its environment. And though one might still distinguish between presentational forms and contents, these clearly are interdependent in relational cases such as, for example, affordance representations. I.e. it might be that what is represented is the actions afforded by present perceptions given prior sensorimotor experience, and that this content is represented via some sort of stored sensorimotor integrations that are evoked by the perceptual present of relevantly similar relations.

Notably, the 'sensorimotor schema' versus 'full action simulation' distinction also becomes a part/whole, sub-personal/personal level distinction. These schemas are components or aspects of the overall experience, executed action or perceived scene. This is a very important feature of my theory. The idea is that sub-personal mental representations can be stored, and then re-presented as *part of* a new personal level 'presentation,' be it perceptually or internally generated. As I see it, this sort of part/whole stored/present understanding is essential to the understanding also of personal level experience and the ways our mental lives are to some extent shared and socially accessible and also private and personal. Given this part/whole idea, it would be interesting to develop a 'multi-dimensional' interpretation of consciousness, which includes the pragmatic and temporal aspects of the experienced and lived present. For the current project the point is that from a social cognitive perspective we might understand elements of others' affordance structures as well as they do – or in the case of our close friends and family possibly better than they do. But that does not entail that we can access their first person experience that precisely is a specific hologram resulting from not just one subset of their affordance structure or past actions but the total 'flavor' of their current attentional focus and mental goings-on.[358]

These are pivotal points both in my more local reinterpretation of mirror neurons and the much broader cognitive implications I here try to alert to. There are a few key aspects that deserves reiteration here. As we have seen I eagerly embrace the project of questioning what one might call the pure 'aboutness' or 'redoubling' view of mental representation as simply an inner presentation of some 'thing-like' content. I focus on the teleological information carried by sensorimotor neurons in fronto-parietal circuits precisely in order to stress the empirical problems or at least limitations of this view. I have ultimately tried to show that the idea of an affordance perception cannot be explained via a traditional view of representation. At least locally an alternative notion is called for to explain the following elements of affordance representations:

1. **Lack of picture or thing-like content coherent in absence of the animal.** The distinction between the form of presentation and content represented cannot be made in the traditional way as the information carried itself reflects a pragmatic relation rather than something that is coherent independently of the perceiving animal.
2. **Challenge to the classical sandwich.** The relational affordance 'content' carried by such neurological circuits seems to depend exactly on this integrating sensory and motor information, thus there is not only an interdependence of form and content, but also exactly of perception and action and thereby a break-down of the neat distinctions of input, central cognition and output, and representations as constructed simply from perceptual inputs.
3. **Challenge to the observable behavior-hidden mental states dichotomy.** This reinterpretation of the 'form/content' of an affordance representation changes not only the distinction between action and perception, but also the typical idea of mental representations as a notoriously hidden and internal version/form with some possibly overtly accessible content.
4. **Mental representations below the personal level.** As affordance

[358] An analysis along these lines would interestingly reframe some of the classic and contemporary debates over the 'what it is like' (Nagel,1974), 'phenomenal' versus 'access consciousness' (Block, 2005), the ideas of the 'transitive content of consciousness' and 'a quality space' of consciousness (Rosenthal, 2005), and the role of attention for consciousness (Prinz, J (forthcoming), Koch & Tsuchiya (2007).), the distinction between consciousness and pre-consciousness (Dehaene et al. 2006) etc.

> representations defy the 'classical sandwich' by representing meaningful sensorimotor integrations in parallel, they also change the notion of mental representations as personal level content available to a coherent introspective presentation for the mental homunculus.

Given such an analysis one might ask why I wouldn't just follow the apparent historical protocol and classify my story as an 'anti-representionalist' theory. The reason why I resist such a move is that I still think it is important to have a term that captures how mental processes can store information and contribute something over and above present movements in response to incoming information. Further, I do not think it is sufficient to point to internal sensorimotor reflexes. In other words, if we want to get beyond the stimulus-response model of behaviorism we cannot simply place the input-output connections in the head. Instead, we must account for the cognitive process of action choice and the analyzable complexity of cognitive processes. This is coherent with what I take to be, for example, Jeannerod's categorization of sensorimotor theories as behaviorist.[359] However, I do not think that action choice can be accounted for by adding a mysterious hidden inner category of cognition sandwiched between inputs and outputs. Rather, as I have argued, a key to understanding action choice is rooted in our capacity of parallel and hierarchical sensorimotor representations. Thus, present perceptions or 'presentations' should be seen as continuously shaped and interpreted through the lens of 're-presentations' of our past experiences both at a personal and a sub-personal level.[360] In so far as our mental lives are shaped by our past and stored experiences our thoughts and intentions will generally not be fully transparent or observable for others in any given present context. Others probably always have elements of their first person perspective that are opaque and inaccessible to any observer. But on the other hand the observer might see certain mental phenomena that we are utterly unaware of and oblivious to ourselves. In the vein of Bergson's work I would suggest that this 'hidden' feature of the mind should not be understood as a feature of mental representation as such. Rather the transcendence and inaccessibility might be seen as a product of memory and the

[359] See Jeannerod (1985) pp. 113-125, where he critiques Bergson and Gibson, but also his recent book on motor cognition (Jeannerod, 2006, p.8-12).

[360] There are interesting parallels to be drawn between my idea here and the Kantian project of reconciling empiricist and rationalist insights and pitfalls, and maybe some neurological elucidation of the famous quote that "Thoughts without content are empty, intuitions without concepts are blind. The understanding can intuit nothing, the senses can think nothing. Only through their unison can knowledge arise." (*Critique of Pure Reason* (1781; 1787), A 51, B 75)

temporally accumulative nature of our mental lives and further what one might call the subjective present synthesis of experience in a first person perspective. The point is here that these considerations paint the 'problem of other minds' in a rather different light. Given the idea that certain mental goings-on are transparent and others transcend someone's present overt engagements, we see that the general query of whether other people 'have minds' becomes a non-issue, while the precise experiences and mental lives of others always to some extent will elude the present observer and even the agent him- or herself.

I cannot develop a full new theoretical framework for mental representations here, but I hope that this discussion of affordances and sketchy proposal show at least the possibility of alternative interpretations and indicate how the neurological findings in fronto-parietal circuits could be seen as sub-personal teleological re-presentations. The key points are on the one hand that the cognitive functions of such representations lie in their parallel and hierarchically dynamical activations, and on the other that such representations should not be equated with our mental goings-on in general but rather simply with the internally stored aspect thereof. Such internal representations are not only historically dependent on emotional, socially and culturally framed sensorimotor engagements but are typically brought together and re-activated via a concrete, situated and overt present engagement such that the mental live of individuals also has an overtly accessible aspect.

6.7. Expelling the lingering dichotomizing assumptions and metaphors

Interestingly the traditional interpretations of mirror neurons, i.e. as providing a motor mirror by which we can understand other perceived actions, are generally seen as belonging under the heading of 'post-cognitivist' approaches. Furthermore, it is obvious that particularly Gallese, but also most of the other core members of the Parma Lab, have many 'post-cognitivist' intuitions and ideas. However, I have continuously argued that their theories, albeit in some ways breaking with the classical sandwich view, in many other ways fail to question their own theoretical assumptions. As we saw in Jacob and Jeannerod's critique, mirror theories of social cognition are seen as 'motor theories' – and the argument becomes about whether or not the motor output system can contribute to perceptual and social functions. This argument is by most seen as a debate between an embodied view of social cognition opposed to a classic cognitivist view. My discussion has hopefully brought out and justified the need for a more fundamental reinterpretation of both motor and social cognition than simply a debate about whether traditional motor functions do some part-time 'moonlight' work in service of traditionally conceived perceptual and higher social cognitive functions.

I argue that exactly the research on sensorimotor integrations, which serves as the larger context of the discovery of mirror neurons, provides an empirical basis for a radical reinterpretation. This is a message has not come across radically enough due to the lingering traditional assumptions. Hence, the pragmatic goal of this thesis is to free the research from these assumptions. I thus focus my critique on the caricature model of mirroring as instantiating a symmetric and encapsulated agent neutral context independent observation-execution matching mechanism, the empirically problematic ideas of motor cognition as covert action simulation, and, lastly, social cognition as 3rd person mindreading, i.e. attribution, prediction and reasoning about the hidden mental states of others.

Given the relevant research discussed in the previous chapters each of these assumptions are empirically misguided and unwarranted. But if this is so obvious, one might suggest that I might be culpable of creating a straw man. Couldn't one in fact object that it is unreasonable that I focus on such caricatured stories that nobody working in the field truly supports? I actually agree that it would be unfair to suggest that this is what most researchers really believe. However the problem, the justification and the need of my project hang on the fact that the *entire debate* about mirror neurons and the role of motor activity in social cognition is, as I have extensively shown, drenched with tacit assumptions exactly taking various elements of these caricatured theories for granted. (See Chapters 2, 3, 4 & 5) Thus, it is very important to note that when I am criticizing Rizzolatti, Gallese, Jeannerod and Csibra my goal is by no means to belittle the value of their overall contribution, but rather to force them and others to truly abandon all those of their well-rehearsed stories and arguments that continuously feed the misunderstanding and, worse yet, which drown out what I see as valuable and enormously important insights of these very same researchers. The alternative story that I give is exactly an attempt to reinterpret the findings with the help of the findings themselves so to speak.

My project is therefore an attempt to dethrone the frameworks of motor and social cognition that I see as standing in the way of seeing the much more complex and radically new findings and ideas. Insights that have generally been overclouded or misinterpreted due to these caricatured and empirically problematic ideas about motor cognition and mirror neurons.

Accordingly, my interpretation of fronto-parietal circuits as supporting social cognition via an affordance space understanding precisely builds on the ideas of the very same people I so vehemently criticize for making unwarranted assumptions. This is to say that my story builds precisely on the broader sensorimotor findings of Rizzolatti and Gallese and other Parma group members, but in addition borrows from and supports Jeannerod's idea of motor schemas and Csibra's developmental findings of the importance and early onset of teleological and affordance understanding.

6.8. Relations to Other Embodied and Phenomenological Reinterpretations

My view in many ways overlaps with phenomenologically inspired views on social cognition advanced by Dan Zahavi and Shaun Gallagher. They have in their respective ways both argued against the theory theory and also the simulation theory of mindreading by pointing to the directness of social perception in many instances. Thus, a core critical point of theirs is that social perception in most normal instances of face-to-face interaction is not mediated by theoretical inferences or by simulative steps.[361] I find their discussion of the issue elucidating and poignant, but I think that for example the notion of direct perception might be somewhat tricky at least when it stands as alone depending on a phenomenological or personal level argument. Many take the idea of direct perception to suggest that social perception is unmediated by other cognitive resources in general, and then object that the question of how social perception is mediated is question of sub-personal processes rather than a phenomenological issue. Interestingly, I think that my account can provide a framework for further filling in and making explicit the ways in which social perception might be 'direct' without claiming that it is unmediated by complex cognitive processes.[362] The points that I have made in regard to this are two-fold. Firstly, I have suggested that we should not see the nature of mental states as in principle opposed to and hidden from observable behavior, but rather that even though the full mental lives of others are always beyond the present behavior there is something

[361] See Zahavi (2008), Gallagher, S. (2007) and Gallagher, S. (2008a).
[362] For an analysis of ways in which perception might be direct or not, see McDermid (2001).

truly mental available in our overt intentional engagements. Thus, this point is about the social as being available as an object of possible perception. Another point that I would like to think is implied by my account has to do with the nature of the process of perception. By way of the analysis of 'canonical neurons' pertaining to object affordances and mirror neurons and other social perception modulated neurons one might say that such neurological structures can support a form of direct perception process that regard not only the pragmatic relevance of objects but also aspects of the mental lives of others. My point is not to argue for a very externalist or ecological account of perception such as that of Gibson or to deny the importance of hierarchical, complex and sub-personal brain processes. Rather, the point is simply that these findings are at the very least compatible with the idea that in spite of sensory processing at much earlier stages than the fronto–parietal circuits, there might not in normal cases be personal level perception at temporally and procedurally prior stages. Or, to shift the burden of proof, I am yet to encounter the empirical evidence that justifies the idea that there is personal level perception first, which then later is pragmatically analyzed, or any other sequential story that could explain the role of sensorimotor modulation during action perception and thereby support the opposite claim that regular face-to-face social perception is 'indirect.'[363]

However, even with the above being said, I must admit I am not so fond of either the notion of 'indirect' or 'direct perception' due to the historical baggage of both of these terms. What is crucial to my story is the idea that the mindedness of others and to some extent their specific mental goings-on are available to be perceived, and that they are perceived via our own experientially educated and sensorimotor-based understanding of the action relations to an actually perceived or mentally constructed affordance space. In other words, our understanding of others can to a large extent be based on our perception of not only their actions but their action *choices* over time given specific contextual affordance structures. Thus, due to our understanding of the social affordance space, we can often *see* the intentional actions choices, motivations, preferences and emotions of others without the need for theoretical inferences or full-blown simulative covert re-enactments of others' actions.

[363] For my discussion and critique of Csibra's attempt to argue that the goals of others are understood independently and prior to motor involvement, see Chapter 5.

I argue that this aspect of social cognition cannot be appropriately construed as either depending on a 'theory of mind' or a 'pretense' or 'embodied simulation' as all these stories leave out the relational and dynamic aspect of our social understanding. Further, I have argued that both the theory theory and the simulation theory, as they stand, are empirically fragile to say the least. I think that many of the 'folk psychological heuristics' that theory theorists point to are often deeply relational in nature and might actually to some extent be reliant on teleological understanding, as also pointed out by Csibra. The empirical problem with theory theorists is that they on the one hand deny that motor cognition plays any significant role in regard to mental states, while they on the other hand postulate a mysterious non-continuous ability and mechanism for the understanding of such hidden states. My analysis suggests that important elements of this mysterious 'Theory of Mind' function might very well be grounded in sensorimotor processes. In regard to the different versions of simulation theory, I have argued that the big empirical problem is the failure to recognize the difference between the highly load-constrained spatiotemporally specified action simulations and more abstract and stored sensorimotor integrations or representations. This distinction is a key to my account. It is the use of multiple parallel teleological 're-presentations' in our perceptual/imaginative 'presentation' that I find most important for our basic social cognitive skills – not our ability to covertly imitate others' exact movements behind a one-way mirror. But this is not to say that teleology is all there is to talk about, or that there are not elements of theorizing and simulation in certain more sophisticated social cognitive tasks. The point is that even these more fancy and detached social thoughts and intentionally controlled simulations would build upon more basic social affordance space understanding.

My story of the role of sensorimotor integration in social cognition also has strong commonalities with many other recent embodied and enactive philosophical contributions well beyond the field of social cognition. I think for example that it is interesting to see my project in relation to that of Alva Noë in regard to enactive perception.[364] Noë argues for the importance of active engagements and sensorimotor knowledge or contingencies for perceptual and other cognitive processes. However, he does this without a reinterpretation of the motor system and of how action relates to

[364] Noë, A. (2005) Action in Perception. Cambridge, MA: The MIT Press. See also the earlier work in collaboration with Kevin O'Regan (O'Regan and Noë (2001)).

perception – one might say that I am attempting to provide the beginnings of such a neurologically inspired reinterpretation. And I would add that on my account there is no reason to say that sensorimotor contingencies are necessarily non-representational or that they are all there is to the phenomenology of perception. Rather, on my view the job of sensorimotor integration has to do only with the potentialities in perception, i.e. predictions, affordances, generalizations etc. And I would also add that there is a phenomenology of the actual aspects that is not captured or explained by sensorimotor activity. But, in an effort to stay within the present project, my point in regard to Noë and other like-minded philosophers is that I think that my social affordance space model can be seen in many ways as both complimentary and corrective in constructive ways.

6.9. Socially afforded conclusions and directions

As a final remark, I hope that this project can lead to practical and theoretical changes within the empirical research and within cognitive science and philosophy of mind and methodologically within philosophy of psychology and neuroscience.

On the empirical side, a clear implication of my project is that we need to stop looking at and theorizing mirror neurons in isolation from other sensorimotor neurons. Furthermore, we need new categorizations of the concrete findings that would allow us to understand various relational elements of teleological and affordance understanding and predictions, one's that are not possible under a simple functional split between 'object-affordance' and 'action-mirror' neurons and within the action category and one-dimensional focus on 'congruency'. As a matter of equal importance, we need new experimental paradigms that do not have as much of an 'imitation bias' as I argue many past and present core experimental setups have. More precisely, we need experiments that do not simply correlate neuron modulation and perceived action types, but which are sensitive to varying pragmatic relational features. In other words, the physical and social affordances of concrete situations and modulations must be experimentally questioned and understood if we are to understand the dynamic modulation that is already observed but not explained in various previous studies.

In regard to theoretical issues in cognitive science and philosophy of mind, I have proposed that we need a new conceptual parsing of mental processes, and that if we succeed in constructing new models of parallel sensorimotor grounded affordance and teleological representations, then this seems to open the door for a new framework for understanding many mental processes, and possibly indicate a post-cognitivist and post-Cartesian horizon to reach out to. I shall in the future, hopefully with the help of others, further analyze the extent and implications of the idea of social affordances and of the affordance space model, which I have developed and argued for in this thesis through an analysis of mirror neurons. Accordingly, I hypothesize that the philosophical implications do not simply have to do with 'mind-reading' and the 'other minds problem' but that the theory has the potential to provide a new angle on issues of action theory, free will, mental representation and the nature of the mental in general.

On a methodological level I see my project as an example of a recent way of engaging the empirical sciences that does not fit neatly within the category of traditional philosophy of neuroscience, i.e. as attempting to categorize and understand the neuroscientific endeavor and the sort of knowledge it yields. Nor can it be seen as a project of neurophilosophy in the traditional Churchland sense of using neurological findings to solve traditional philosophical problems. Rather, it seems to me that there is a third and growing field of truly inherently interdisciplinary work. Instead of merely letting philosophy and neuroscience inform each other's work, this field of empirically engaged philosophy reinterprets and develops philosophical theories *within* the actual empirical scientific practice and might rather than philosophy of science be called philosophy *in* science. The motivating ideas are, on the one hand, that philosophy of mind of the armchair variety seems not only socially deprived but also burdened by fundamental limitations in regard to sub-personal processes, and on the other hand that there simply is no pre-theoretical pure empirical science. I hold that the fields are thus naturally inherently intertwined and it is truly great to see how a form of actual interdisciplinary practice and exchange is reemerging between philosophy and empirical psychology at large. This case study and reinterpretation of mirror neuron research and the developed social affordance theory should be seen as an example of such a multi-disciplinary project with both sources and consequences in multiple fields.

Bibliography

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. Nature Neuroscience 11, 1109-1116.

Agnew, Z., & Wise, R.J.S. (2008). Separate areas for mirror responses and agency within the parietal operculum.The Journal of neuroscience, 28(47), 12268-12273

Akins, K. (1996). Of Sensory Systems and the 'Aboutness' of Mental States, *Journal of Philosophy*, 93: 337–372.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences.* 4, 267–278.

Arbib, M.A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics, Behavioral and Brain Sciences 28**,** pp. 105–168

Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), Natural theories of mind: Evolution, development and simulation of everyday mindreading (pp. 233-251). Oxford: Basil Blackwell.

Baron-Cohen, S. (1995). *Mindblindness: Essay on Autism and the Theory of Mind, Learning, Development and Conceptual Change*, MIT Press.

Barsalou, L. W. (2008) Grounded cognition. *Annual Review of Psychology*. 59, 617–645.

Bauman, M. L., & Kemper, T. L. (1986). Developmental cerebellar abnormalities: A consistent finding in early infantile autism. Neurology 36[Suppl1]:190.

Bauman, M. L., & Kemper, T. L. (1990). Limbic and cerebellar abnormalities are also present in an autistic child of normal intelligence. Neurology 40[Suppl 1]:359.

Bergson, H. (1896). *Matière et Memoire*. (English Translation: *Matter and Memory,* 2004 republication of 1912 MacMillan edition. translators N. Margaret Paul and W. Scott Palmer. Dover Publications).

Bergson, H. (1907). L'Evolution Créatrice, F. Alcan, Paris (English trans. Creative Evolution).

Blakemore, S. J. ,& Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience, 2,* 561–567.

Blakeslee, S. (2006, January 10th) Cells that read minds, *New York Times*.

Block, N. (2005). Two neural correlates of consciousness. Trends in Cognitive Science. 9, 46–52.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica (Amst)*, 106, 3–22.

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, 44, 124–143.

Brass, M., & Heyes, C. (2005). Imitation: Is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, 9, 489–495.

Brass, M., Schmitt, R. M., Spengler, S., & Gergely, G. (2007). Investigation action understanding: Inferential processes versus action simulation. *Current Biology*, 17, 2117–2121.

Brass, M., Schmitt, R., Spengler, S., and Gergely, G. (2007). Investigating action understanding: Inferential processes versus action simulation. *Current Biology*. 17, 2117–2121.

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139-159.

Buccino, G., Bikofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., & Freund, H.-J. (2001). Action observation activates premotor and parietal areas in somatotopic manner: an fMRI study. *European Journal of Neuroscience.* Vol. 13, 400-404.

Buccino, G., Bikofski, F., & Riggio, L. (2004). The mirror neuron system and action recogntion. *Brain and language*, Vol. 89, 370-376.

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition, 112,* 337-342.

Caggiano V., Fogassi, L., Rizzolatti, G., Their, P., & Casile, A. (2009). Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. Science, vol. 324, p.403-406.

Call, J., & Tomasello, M. (1998). Distinguishing intentional from accidental actions in orangutans (Pongo pygmaeus), chimpanzees (Pan troglodytes), and human children (Homo sapiens). Journal of Comparative Psychology, 112, 192-206.

Calvo-Merino, B., Grèzes, J., Glaser, D.E., Passingham, R.E. & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology,* Vol.16, 1905-1910.

Calvo-Merino, B. et al. (2005). Action observation and acquired motor skills: an fMRI study with expert dancers. *Cerebral Cortex*, 15,1243 – 1249.

Carruthers, P. (1996). Simulation and self-knowledge: a defence of the theory-theory. In P. Carruthers & P.K. Smith, Eds. *Theories of theories of mind.* Cambridge: Cambridge University Press.

Carruthers, P. & Smith, P.K. (1996). Theories of theories of mind. Cambridge: Cambridge University Press.

Catmur, C., Gillmeister, H., Bird, G., Liepelt, R., Brass, M. & Heyes, C. (2008). Through the looking glass: counter-mirror activation following incompatible sensorimotor learning. *European Journal of Neuroscience*, 28, 1208–1215.

Catmur, C., Walsh, V. & Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current Biology* 17, 1527-1531.

Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G. & Rizzolatti, G. (2007). Impairment of actions chains in autism and its possible role in intention understanding. *Proceedings in National Acadamy of Science*, 104(45): 17825–17830.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception– behavior link and social interaction. *Journal of Personality and Social Psychology, 76,* 893–910.

Cheng, Y., Meltzoff, A.N. & Decety, J. (2007). Motivation modulates activity of the human mirror neuron system. *Cerebral Cortex*, Vol 17, 1979-1986.

Chomsky, N. (1966). *Cartesian linguistics: A Chapter in the History of Rationalist Thought*. New York, Harper and Row.

Chomsky, N. (1986). *Language and the Mind*. New York, Harcourt, Brace and World.

Chong, T. T., Cunnington, R., Williams, M. A., Kanwisher, N., & Mattingley, J. B. (2008). fMRI adaptation reveals mirror neurons in human inferior parietal cortex. Current Biology, 18, 1576–1580.

Cisek P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical Transactions of the Royal Society London B Biological Sciences*. 362(1485): 1585–1599.

Cisek, P. (2008) The affordance competition hypothesis. In *Embodiment, ego-space, and action* by Roberta L. Klatzky, Brian MacWhinney, Marlene Behrman (ed.). CRC press.

Clark, A. (1997). *Being There: Putting Brain Body and World Together Again.* Cambridge, MA: MIT Press.

Clements, W. A. & Perner, J. (1994). Implicit understanding of false belief. *Cognitive Development*, 9. 377-395.

Costantini, M., Galati, G., Ferretti, A., Caulo, M., Tartaro, A., Romani, G.L. and Aglioti, S. M. (2005) Neural systems underlying observation of humanly impossible movements: an fMRI study. Cerebral Cortex, 15, 1761–1767.

Courchesne, E. (1997) Brainstem, cerebellar and limbic neuroanatomical abnormalities in autism, *Current Opinion in Neurobiology*, Volume 7, Issue 2, April 1997, Pages 269-278.

Craighero, L., Bello, A., Fadiga, L., & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsycholo- gia*, *40*, 492–502.

Csibra, G. (2005). Mirror neurons and action observation. Is simulation involved? *http://www.interdisciplines.org/mirror/papers/4*.

Csibra. G (2007). Action mirroring and action understanding: an alternative account. In: P. Haggard, Y. Rosetti, & M. Kawato (Eds.), Sensorimotor Foundations of Higher Cognition. Attention and Performance XXII. Oxford: Oxford University Press.

Csibra, G. & Gergely, G. (2007) 'Obsessed with goals': functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica* 124, p. 60-78.

Damasio, A. (1994) *Descartes' Error: Emotion, Reason, and the Human Brain*, Putnam Publishing.

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y. & Iacoboni, M. (2006) Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience.* 9:28–30

Davies, M. (1995) "Consciousness and the varieties of aboutness". In C. Macdonald and G. Macdonald (eds.), *Philosophy of Psychology: Debates on Psychological Explanation*. Oxford: Blackwell Publishers, 356–92.

Decety, J., Jeannerod, M., & Prablanc, C. (1989). The timing of mentally represented actions. *Behavioural Brain Research*, 34, 35-42.

Decety, J. and Lamm, C. (2007), The Role of the Right Temporoparietal Junction in Social Interaction: How Low-Level Computational Processes Contribute to Meta-Cognition, *Neuroscientist* Vol. 13, No. 6, 580-593

Dehaene, S. Changeux, J.-P., Naccache, L., Sackur, J. and Sergent C. (2006) Conscious, preconscious, and subliminal processing: a testable taxonomy. Trends in Cognitive Sciences. 10, 204–211.

Dennett, D. (1978). Beliefs about beliefs. Behavioral and Brain Sciences, 1, 568-570.

de Vignemont, F. (2008) Frames of reference in social cognition. *The Quaterly Journal of Experimental Psychology*, 90-100.

de Vignemont, F., & Fourneret, P. (2004). The sense of agency: A philosophical and empirical review of the ''who" system. *Consciousness and Cognition*, 13(1), 1–19.

Descartes, R. (1641/1996) *Meditationes de prima philosophia* (Eng Trans: *Meditations on first philosophy* (with selections from the objections and replies). Translated and edited by John Cottingham. Glasgow: Cambridge University Press).

Desimone, R. (1996) Neural mechanisms for visual memory and their role in attention. Proceedings of the National Academy Sci USA 93:13494–13499.

Dewey, J. (1929). *Experience and nature* (2nd ed.). LaSalle, IL: Open Court.

Dinstein, I. (2008). Human cortex: Reflections of mirror neurons. *Current Biology*, 18, R956–R959.

Dinstein, I., Hasson, U., Rubin, N., & Heeger, D. J. (2007). Brain areas selective for both observed and executed movements. *Journal of Neurophysiology*, 98, 1415–1427.

Dinstein, I., Gardner, J.L., Jazayeri, M., Heeger, D.J., (2008a) Executed and observed movements have different distributed representations in human aIPS. J. Neurosci. 28, 11231–11239.

Dinstein, I., Thomas, C., Behrmann, M., & Heeger, D. J. (2008b). A mirror up to nature. *Current Biology*, 18, R13–R18.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research,* Vol. 91, 176-180.

Dreyfus, H. (1972/92). *What Computers Can't Do: A Critique of Artificial Reason.* New York: Harper and Row. (Third edition: *What Computers Still Can't Do.* 1992. Cambridge, MA: MIT)

Dreyfus, H. L. & Dreyfus, S. E. (1999) The challenge of Merleau-Ponty's phenomenology of embodiment for cognitive science. In G. Weiss & H. F. Haber (eds.) Perspectives on embodiment – the intersections of nature and culture (pp. 103-120). New York: Routledge

Ellis, R., & Tucker, M. (2000). Micro-affordance: The potentiation of components of action by seen objects. British Journal of Psychology, 91, 451-471.

Fabbri-Destro, M. & Rizzolatti (2008). Mirror neurons and mirror systems in monkeys and humans. *Physiology, 23*(3), 171.

Fadiga, L.,Craighero, L. & Olivier, E. (2005). Human motor cortex excitability during the perception of others' action. *Current Opinion in Neurobiology*, vol. 15, 213-218.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology, 73,* 2608–2611.

Fadiga, L., Buccino, G., Craighero, L., Fogassi, L., Gallese, V. & Giovanni, P. (1998). Corticospinal excitability is specifically modulated by motor imagery: A magnetic stimulation study. *Neuropsychologia, 37*(2), 147.

Fadiga, L., Fogassi, L., Gallese, G. & Rizzolatti, G. (2000) Visuomotor neurons: Ambiguity of the discharge or "motor" perception? *International Journal of Psychophysiology* 35:165–77.

Fagiolo S., Hommel, B. & Schobotz, R. I. (2007) Intentional control of attention: action planning primes action-related stimulus dimensions. *Psychological Research*. 71, 22–29.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, 9, 878–879.

Fernandino, L. & Iacoboni, M. (2010) Are cortical motor maps based on body parts or coordinated actions? Implications for embodied semantics, Brain and Language 112, p. 44–53.

Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience, 8,* 1703–1714.

Ferrari, P. F., Rozzi, S., & Fogassi, L. (2005). Mirror Neurons responding to observation of actions made by tools in monkey ventral premotor cortex. *Journal of Cognitive Neuroscience*, Vol. 17, No. 2, 212-226.

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, 424, 769–771.

Fodor, J. A. (1975) *The Language of Thought*, Harvard University Press.

Fodor, J. A.(1983) *The Modularity of Mind: An Essay on Faculty Psychology*, MIT Press.

Fodor, J. A. 1992. A theory of the childs mind. Cognition. 44, 283-296.

Fogassi L. and Ferrari, P.F. (2007) Mirror neurons and the evolution of embodied language, Current Directions in Psychological Science 16, pp. 136–141.

Fogassi L, Gallese V, di Pellegrino G, Fadiga L, Gentilucci M, Luppino, Matelli, M., Pedotti and Rizzolatti (1992). Space coding by premotor cortex. *Experimental Brain Research, 89*(3), 686-690.

Fogassi, L., Gallese, V., Buccino, G., Craighero L., Fadiga, L. & Rizzolatti, G. (2001). Cortical mechanism for the visual guidance of hand grasping movements in the monkey: A reversible inactivation study. *Brain*, 124(3), 571.

Fogassi, L. & Gallese, V. (2002). The neural correlates of action understanding in non-human primates. In M. I. Stamenov & V. Gallese (Eds.), *Mirror Neurons and the Evolution of Brain and Language* (pp. 13-35). Amsterdam: John Benjamins Publ.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F. & Rizzolatti, G. (2005). Parietal lobe: from action organization to action understanding. *Science*, Vol. 38, 662-667.

Fogassi, L., Gallese, V., Fadiga L. and Rizzolatti, G. (1998) Neurons responding to the sight of goal-directed hand/arm actions in the parietal area PF (7b) of the macaque monkey. Society for Neuroscience abstracts. 28th annual meeting, Los Angeles, Calif., November 7-12, 1998. / Volume 24, parts 1 and 2.

Fogassi L. & Luppino. G. (2005) Motor functions of the parietal lobe. Current Opinion in Neurobiology 15:626–631.

Freeman, W. J. & Skarda, C. (1990), Representations: Who Needs Them? in *Brain Organization and Memory Cells, Systems, & Circuits* Edited by: J. L. McGaugh, N. Weinberger, G. Lynch, page: 375–380.

Frith, C. (1996) Neuropsychology of Schizophrenia. British Medical Bulletin 52, (No. 2): 618-626

Frith, U. and Frith, C.D. (2003) Development and neurophysiology of mentalizing. In The Neuroscience of Social Interaction (Frith, C. and Wolpert, D. eds), pp. 45–75, Oxford University Press.

Fujii N, Hihara S, & Iriki A. (2008) Social cognition in premotor and parietal cortex. *Social Neuroscience,* Volume 3, Issue 3 & 4 , pp. 250 - 260.

Gagliardi, J. L., Kirkpatrick-Steger, K. K., Thomas, J., Allen, G.J. & Blumberg, M.S. (1995). Seeing and knowing: Knowledge attribution versus stimulus control in adult humans (Homo sapiens). *Journal of Comparative Psychology*, 109, 107-114.

Gallagher, S. (2001) The practice of mind: theory, simulation or primary interaction? Journal of Consciousness Studies, 8, pp. 83-108

Gallagher, S. (2004). Understanding interpersonal problems in autism: Interaction Theory as an alternative to Theory of Mind. Philosophy, Psychiatry and Psychology, 11, pp. 199-217

Gallagher, S. (2007). Simulation trouble. *Social Neuroscience.* Vol. 2, No.3&4, pp. 353-365.

Gallagher, S. (2008) Direct perception in the intersubjective context. *Consciousness and Cognition* 17: 535-543.

Gallagher, S. & Zahavi, D. (2008) *The Phenomenological Mind: An introduction to philosophy of mind and cognitive science*. Routledge, New York.

Gallese, V. (1999) From Grasping to Language: Mirror Neurons and the Origin of Social Communication in Toward a science of consciousness III: the third Tucson discussions and debates. Ed. by David John Chalmers, Alfred W. Kaszniak.

Gallese, V. (2000). The inner sense of action – agency and motor representations. *Journal of Consciousness Studies, 7*(10), 23-23-40.

Gallese, V. (2001) The 'Shared manifold' hypothesis. Journal of Consciousness Studies, 8, No. 5–7, pp. 33–50.

Gallese, V. (2003). The manifold nature of interpersonal relations: the quest for a common mechanism. *Philosophical Transactions of the Royal Society,* LondonB, 358, 517-528.

Gallese, V. (2006) Intentional attunement: a neurophysiological perspective on social cognition and its disruption in autism. Cogn. Brain Res. 1079, 15–24

Gallese, V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B*, 362, 659-669.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, Vol. 119, 593-609.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In W. Prinz & B. Hommel (Eds.), *Attention & performance XIX: Common mechanisms in perception and action* (pp. 334–355). Oxford, England: Oxford University Press.

Gallese, V. & Goldman, A. (1998). Mirror neurons and the simulation theory of mind reading. *Trends in Cognitive Sciences*, 12, 493-501.

Gallese, V. & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 1*(1).

Gallese, V. & Keysers, C. (2001). Mirror neurons: A sensorimotor representation system. *Behavioral and Brain Sciences, 24*(5), 983.

Gallese, V., Keysers, C. & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, Vol.8, No.9, 396-403.

Gazzola & Keysers. (2008) The observation and execution of actions share motor and somatosensory voxelsin all the tested subjects:single-subject analysis of unsmoothed fMRI data. Cerebral cortex. Doi:10.1093/cercor/bhn181

Gazzola, V., Rizzolatti, G., Wicker, B. & Keysers C. (2007). The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage*, Vol. 35, 1674-1684.

Gentilucci, M. and Corballis, M.C. (2006) From manual gesture to speech: A gradual transition, Neuroscience and Biobehavioral Reviews 30, pp. 949–960.

Gentilucci, M. & Rizzolatti, G. (1990) "Cortical control of arm and hand movements", in *Vision and Action: the control of grasping,* edited by Melvyn Goodale. Published by Intellect Books.

Gentillucci, M., Fogassi, L., Luppino, G., Matelli, M., Camarda, R. & Rizzolatti, G. (1988) Functional organization of inferior area 6 in macaque monkey. I. Somatotopy and the control of proximal movements. *Experimental Brain Research*, 71:475–490

Georgieff, N., & Jeannerod, M. (1998). Beyond consciousness of external reality: A ''who" system for consciousness of action and self-consciousness. *Consciousness and Cognition*, 7(3), 465–477.

Gergely, G., Nádasdy, Z., Csibra, G. and Bíró, S. (1995) Taking the intentional stance at 12 months of age. Cognition, 56, 165–193.

Gerrans, P. (2007). Mental time travel, somatic markers and "myopia for the future". *Synthese*, 159, 3: 459-474.

Gibson, J. J. (1960) The concept of the stimulus in psychology. American Psychologist, 15, 694-703.

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.

Gibson, J. J. (1976) What is it to perceive? In: J. J. Gibson, Purple perils. (unpublished work). http://www.trincoll.edu/depts/ecopsyc/perils/folder6/perceive.html

Glenberg, A. (1999). "Why Mental Models Must Be Embodied." In *Mental Models in Discourse Processing and Reasoning*, Rickheit, G. and Habel, C. (eds). New York: Elsevier.

Goldman, A. (1993) *Philosophical Applications of Cogntive Science,* Westview Press.

Goldman, A.I. (2006) *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading*, New York: Oxford University Press.

Goldman, A. I. (2008) "Mirroring, Mindreading, and Simulation," in J. A. Pineda, ed., *Mirror Neuron Systems: The Role of Mirroring Processes in Social Cognition*, pp. 311-330. New York: Humana Press.

Goldman, A. (2009). Mirroring, simulating and mindreading. *Mind and Language*. Volume 24 Issue 2, Pages 235 – 252.

Gottlieb, J., Balan P.F., Oristaglio, J, Suzuki M (2009): Parietal control of attentional guidance: The significance of sensory, motivational and motor factors, *Neurobiol. Learning and Memory*. 91(2): 121–128.

Gentilucci, M., L. Fogassi, G. Luppino, M. Matelli,R. Camarda, G. Rizzolatti. (1988). Functional organization of inferior area 6 in the macaque monkey: I. Somatotopy and the control of proximal movements. In *Exp Brain Res* 71:475-490.

Georgiou, I., Becchio, C., Glover, S., & Castiello, U. (2007). Different action patterns for cooperative and competitive behaviour. *Cognition, 102*(3), 415-433.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naı¨ve theory of rational action. *Trends in Cognitive Sciences, 7,* 287–292.

Gobbini, Maria Ida, Koralek, Aaron C., Bryan, Ronald E., Montgomery, Kimberly J. & Haxby, James V. (2007). Two Takes on the Social Brain: A Comparison of Theory of Mind Tasks, Journal of Cognitive Neuroscience 2007 19:11, 1803-1814.

Grafton, S. T. (2008) Malleable templates: reshaping our crystallized skills to create new outcomes. *Nature Neuroscience* 11: 248–249.

Grafton, S. T., M.A. Arbib, L. Fadiga, and G. Rizzolatti. (1996). Localization of grasp representations in humans by PET: 2. Observation compared with imagination. In *Experimental Brain Research.* 112:103-111.

Grafton, S. T., Fadiga, L., Arbib, M. A., & Rizzolatti, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *NeuroImage, 6,* 231–236.

Grèzes, J., Armony, J. L., Rowe, J., Passingham, R.E. (2003). Activations related to "Mirror" and "Canonical" neurons in the human brain: a fMRI study. *Neuroimage* 18: 928-937.

Grèzes, J., & Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia, 40,* 212–222.

Grèzes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: An fMRI study. *NeuroImage, 21,* 744–750.

Grill-Spector, K. (2006). Selectivity of adaptation in single units: implications for fMRI experiments. *Neuron* 49, 170–171.

Häberle, A., Schütz-Bosbach, S., Laboissière, R. & Prinz, W. (2008) Ideomotor action in cooperative and competitive settings, Social Neuroscience, 1747-0927, Volume 3, Issue 1, pp 26 – 36.

Haggard, P. & Clark, S. (2003) Intentional action: conscious experience and neural prediction. Conscious. Cogn. 12, 695–707.

Handy, T.C., Tipper, C.M., Schaich, Borg, J., Grafton, S.T. & Gazzaniga, M.S. (2006). Motor experience with graspable objects reduces their implicit analysis in visual-and motor-related cortex. *Brain Research, 1097*(1), 156.

Hamilton, A. F. & Grafton S. T. (2006) Goal representation in human anterior intraparietal sulcus. Journal of Neuroscience 26: 1133–1137.

Hamilton, A., Wolpert, D. M., Frith, U., & Grafton, S.T. (2006). Goal Representation in Human Anterior Intraparietal Sulcus. *NeuroImage* 29:524–535.

Hamilton, A.F., & Grafton, S.T. (2007). “The motor hierarchy: From kinematics to goals and intentions”. In *Attention and Performance xxii*. Y. Rosetti, M. Kawato, and P. Haggard, eds. (Oxford: Oxford University Press).

Hamilton, A., Wolpert, D., & Frith, U. (2004). Your own action influences how you perceive another person’s action. *Current Biology, 14,* 493–498.

Hamilton, A., Joyce, D. W., Flanagan, R., Frith, C. D., & Wolpert, D. M. (2007). Kinematic cues in perceptual weight judgment and their origins in box lifting. *Psychological Research*, vol 71, p. 13-21.

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences, USA, 95,* 15061–15065.

Haselager, P., de Groot, A., & van Rappart, H. (2003). Representationalism vs. anti-representationalism: a debate for the sake of appearance. *Philosophical Psychology* 16(1):5–53.

Haueisen, J., & Knosche, T. R. (2001). Involuntary Motor Activity in Pianists Evoked by Music Perception. *Journal of Cognitive Neuroscience*. 13: 786-792).

Heidegger, M. (1927) *Sein und Zeit*, Gesamtausgabe Volume 2. (Eng translation: *Being and Time*, trans. by John Macquarrie and Edward Robinson (London: SCM Press, 1962); re-translated by Joan Stambaugh (Albany: State University of New York Press, 1996).

Heider, F. and Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology,* 57, 243–259.

Heyes, C. M. (1998) Theory of Mind in Nonhuman Primates. *Behavioral Brain Sciences*. 21, pp. 101–134.

Heyes, C., Bird, G., Johnson, H., & Haggard, P. (2005). Experience modulates automatic imitation. *Cognitive Brain Research*, 22, 233-240.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience* 21(7):1229–1243.

Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24,* 849–937.

Hurley, Susan (1998). *Consciousness in Action.* Harvard University Press.

Hurley**, S.** (2008). The Shared Circuits Model (SCM): How Control, Mirroring, and Simulation Can Enable Imitation, Deliberation, and Mindreading. *Behavioral and Brain Sciences* 31 (1):1-22.

Hyvärinen, J. (1982). Posterior parietal lobe of the primate brain. *Physiological Review*. 62: 1060-I 129.

Iacoboni, M. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology, 3*(3), e79.

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience*. 7, 942–951.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C., et al. (2005). Grasping the Intentions of Others with One's Own Mirror Neuron System. PLoS Biol 3(3): e79. doi:10.1371/journal.pbio.0030079

Iacoboni, M., Koski, L. M., Brass, M., Bekkering, H., Woods, R. P., Dubeau, M. C., Mazziotta, J. C. & Rizzolatti, G. (2001) Reafferent copies of imitated actions in the right superior temporal cortex. *Proceedings of the National Academy of Sciences*. USA 98, 13995–13999.

Iriki A. (2006). The neural origins and implications of imitation, mirror neurons and tool use. *Current Opinion in Neurobiology*. 16:660–67

Jacob, P. (2008). What Do Mirror Neurons Contribute to Human Social Cognition? *Mind and Language* 23 (2):190–223.

Jacob, P. (2009). The tuning-fork model of human social cognition: A critique. Consciousness and Cognition.

Jacob, P. & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences*, Vol. 9, NO.1, 21-25.

James, W. (1890). *The Principles of Psychology*, 2 vols., Dover Publications 1950, vol. 1

Jeannerod, M. (1983) *Le Cerveau-Machine - physiologie de la volonté*, Paris, Fayard. (English trans. The Brain Machine – The development of neurophysiological thought. 1985. Harvard University Press. Translated by David Urion).

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

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage, 14,* S103–S109.

Jeannerod, M. (2004). Visual and action cues contribute to self-other distinction. *Nature Neuroscience* 7, 422-423.

Jeannerod, M. (2006). *Motor Cognition – what actions tells the self*, Oxford University Press.

Jeannerod, M., Arbib, M.A., Rizzolatti, G. & Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends in Neuroscience*, IV, 314-320.

Jellema, T., & Perrett, D. I. (2007). Neural pathways of social cognition. In: *The Oxford Handbook of Evolutionary Psychology* (R. I. M. Dunbar & L. Barrett, Eds), Oxford University Press. pp 163-177.

Jellema, T. & Perrett, D. I. (2005). Neural basis for the perception of goal-directed actions. In: *The Cognitive Neuroscience of Social Behaviour* (A. Easton and N. Emery, Eds.) Psychology Press, Studies in Cognition Series, pp 81-112.

Johnson, S. H. (2000). Thinking ahead: The case for motor imagery in prospective judgements of prehension. *Cognition*, 74, 33-70.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*, 430–454.

Kant, I. (1787). *Kritik der reinen Vernunft*, 2nd edition.Werke.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., et al. (2005). Perception of motion affects language processing. *Cognition, 94*(3), B79-B89.

Keysers, C., Kohler, E., Umilta, M.A., Nanetti, L., Fogassi, L & Gallese, V. (2003). Audiovisual mirror neurons and action recognition, *Experimental Brain Research*, Vol. 153, 628-636.

Keysers, C. & Perrett, D.I. (2004). Demystifying social cognition: a Hebbian perspective. *Trends in Cognitive Sciences*, Vol. 8, 501-507.

Keysers, C., Wicker, B., Gazzola, V., Anton, J.L., Fogassi, L., & Gallese, V. (2004). A touching sight : SII/PV activation during the observation and experience of touch. *Neuron* 42, 335–346.

Kilner, J.M., Friston, K.J., and Frith, C.D. (2007a). Predictive coding: An account of the mirror neuron system. *Cognitive Processing* 8, 159–166.14.

Kilner, J.M., Friston, K.J., and Frith, C.D. (2007b). The mirror-neuron system: A bayesian perspective. *Neuroreport* 18, 619–623.

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology*, 13, 522–52.

Kilner, J. M., Vargas, C., Duval, S., Blakemore, S.-J. & Sirigu, A. (2004). Motor activation prior to observation of predicted movement. *Nature Neuroscience,* Vol. 7, No.12, 422-423.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Sciences 11*, 16–22.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt Brace, p.7.

Kohler, E., Keysers, C., Umilta, M.A., Fogassi, L., Gallese, V. & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, Vol. 297, 846-848.

Koski, L., Wohlschläger, A., Bekkering, H., Woods, R., Dubeau, M.-C., Mazziotta, J. C. & Iacoboni, M. (2002). Modulation of Motor and Premotor Activity during Imitation of Target-directed Actions. *Cerebral Cortex* 12: 847-855.

Kroliczak, G., McAdam, T.D., Quinlan, D.J., & Culham, J.C. (2008). The human dorsal stream adapts to real actions and 3D shape processing: A functional magnetic resonance imaging study. *Journal of Neurophysiology*, 100:2627–2639.

Kurata, K. & Tanji, J. (1986) Premotor cortex neurons in macaques: activity before distal and proximal forelimb movements. *Journal of Neuroscience*, 6: 403-41 1, 1986.

Lakoff, G. & Johnson, M. (1999) *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought,* Basic Books, New York, 1999.

Leslie, A. (1987). Pretense and Representation: the Origins of "Theory of Mind". Psychological Review, vol. 94, no.4, 412-426.

Leslie, A. M. (1992). Pretence, autism, and the theory-of-mind-module. Current Directions in Psychological Science, 1, 18–21.

Leslie, A. & Thaiss, L. (1992). Domain specificity in conceptual development. Cognition 43, 225–251.

Leube, D. T., Knoblich, G., Erb, M., Grodd, W., Bartels, M., & Kircher, T.T. (2003). The neural correlates of perceiving one's own movements. *Neuroimage.* 20:2084-2090.

Lhermitte, F., Pillon, B. & Serdaru, M. (1986). Human Autonomy and the frontal lopes Part I: imitation and utilization behavior: a neurological study of 75 patients. *Annals of Nerology*, Vol. 19, 326-334.

Liberman, A. M., F. S. Cooper, D. P. Shankweiler, & M. Studdert-Kennedy. (1967). Perception of the speech code. *Psychological Review,* 74:431-461.

Liberman, A. M., & Mattingly, I. G. (1985) The motor theory of speech perception revised. *Cognition.* 21:1-36.

Lingnau, A., Gesierich, B., & Caramazza, A. (2009). Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. *Proceedings of the National Academy of Sciences*, U S A 106:9925–9930.

Locke, J. (1690). *An Essay Concerning Human Understanding*. London.

Luppino, G. & Rizzolatti, G. (2000). The organization of the frontal motor cortex. *News in Physiological Sciences* 15: 219–224.

Lycan, W. (2000). Representational theories of consciousness. In (E. Zalta, ed) *The Stanford Encyclopedia of Philosophy*.

Lyons, D. E., Santos L. R. & Keil, F. C (2006). Reflections of other minds: how primate social cognition can inform the function of mirror neurons, *Current Opinion in Neurobiology* 16, pp. 230–234.

Mar, R. A., Kelley, W. M., Heatherton, T. M. & Macrea, C. M. (2007). Detecting agency from the biological motion of veridical vs animated agents. *Social Cognitive and Affective Neuroscience*.2 (3): 199-205.

Marchetti, C. & Della Salla, S. (1998) Disentangling the alien and anarchic hand. *Cognitive Neuropsychiatry*. 3, 191-208.

Marr, D. (1982) *Vision: A computational investigation into the human representation and processing of visual information*. New York: Freeman.

McDermid, D. J. (2001). What is Direct Perceptual Knowledge? A Fivefold Confusion. Grazer Philosophische Studien 62 (1):1-16

Meltzoff, A. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, 31, 838-850.

Meltzoff, A. N., & Decety, J. (2003). What imitation tells us about social cognition: A rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society* (*Series B*)*, 358,* 491–500.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198,* 75–78.

Merleau-Ponty, M. (1942) *La Structure du comportement*, Presses Universitaires de France, Paris. (Eng translation: *The Structure of Behavior* trans. by Alden Fisher, (Boston: Beacon Press, 1963; London: Methuen, 1965).

Merleau-Ponty, M. (1945) *Phénoménologie de la perception*, Gallimard, Paris.(Eng trans: *Phenomenology of Perception* trans. by Colin Smith, (New York: Humanities Press, 1962) and (London: Routledge & Kegan Paul, 1962) translation revised by Forrest Williams, 1981; reprinted, 2002)

Miall, R.C. (2003). Connecting mirror neurons and forward models. *Neuroreport* 14, 2135–2137.

Milner, A. D. & Goodale, M. A. (1996). *The visual brain in action*. Oxford University Press.

Mitchell, J. P. (2008). Activity in Right Temporo-Parietal Junction is Not Selective for Theory-of-Mind, *Cerebral Cortex*. 18(2): 262-271.

Montgomery, K. J., Isenberg, N., & Haxby, J. V. (2007). Communicative hand gestures and object-directed hand movements activated the mirror neuron system. *Social Cognitive and Affective Neuroscience, 2*(2), 114-122.

Mountcastle, V.B., Lynch, J.C., Georgopoulos, A., Sakata, H., Acuña, C. (1975). Posterior parietal association cortex of the monkey: command functions for operations within extrapersonal space. *Journal of Neurophysiology.* 38:871–908.

Mukamel, R., Ekstrom, A.D., Kaplan, J., Iacoboni, M., and Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*. 20, 750–756.

Murata, A., Gallese, V. Kaseda, M. and Sakata, H. (1996). Parietal neurons related to memory-guided hand manipulation. *Journal of Neurophysiology, 75*(5), 2180.

Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997). Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of Neurophysiology, 78*(4), 2226.

Nagel, T. (1974). What it is like to be a bat. *Philosophical Review 83*. October, 435-450.

Nelissen, K., Luppino, G., Vanduffel, W., Rizzolatti, G., & Orban, G.A. (2005). Observing others: multiple action representation in the frontal lobe. *Science,* 310: 332–336, 2005.

Newman-Norlund, R.D, van Schie, H.T.,van Zuijlen, A.M.J, & Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative actions. *Nature Neuroscience*, Vol. 10, No. 7, 817-818.

Nichols, S., & Stich, S. (2000). A cognitive theory of pretense. *Cognition, 74*(2), 115-147.

Nichols, S., and Stich, S. (2003) Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding of Other Minds, Oxford University Press

Nietzsche, F. (1882) *Die fröhliche Wissenschaft* (Eng trans: *The Gay Science* trans. Walter Kaufmann, Vintage, 1974)

Nietzsche, F. (1887) *Zur Genealogie der Moral*, 1887 (Eng. Trans: *On the Genealogy of Morality* in: 'Basic Writings of Nietzsche', trans. Walter Kaufmann, Modern Library, 2000)

Noë, Alva (2004). *Action in perception*. Cambridge, MA: MIT Press.

Oberman, L. M., Pineda, J. A., & Ramachandran, V. S. (2007). The human mirror neuron system: A link between action observation and social skills. *Social Cognitive and Affective Neuroscience, 2*(1), 62-66.

Oberman, L. M., & Ramachandran, V. S. (2007). The simulating social mind: The role of simulation in the social and communicative deficits of autism spectrum disorders. *Psychological Bulletin*, *133*, 310–327.

Onishi, K. & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science* 308:255–58.

Onishi, K. H., Baillargeon, R. & Leslie, A. M. (2007). 15-month-old infants detect violations in pretendscenarios. *Acta Psychologica*, 124, 106–128.

O'Regan, J.K. and Noe, A. (2001) A sensorimotor account of vision and visual consciousness. *Behavioral Brain Sciences.* 24, 939–973

Ortigue, S., King, D., Gazzaniga, M., Miller, M. and Grafton, S. (2009). Right hemisphere dominance for understanding the intentions of others: evidence from a split-brain patient, *BMJ Case Reports,* May 21

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition,* 107, 1:179-217.

Pacherie, E. & Haggard, P. (forthcoming). What are intentions?. In, L. Nadel & W. Sinnott-Armstrong (eds.) *Benjamin Libet and Agency*. Oxford: Oxford University Press.

Penfield, W. & Rasmussen, T. (1952) *The Cerebral Cortex of Man*. MacMillan. New York.

Perner, J., & Leekam, S. R. (2008). The curious incident of the photo that was accused of being false: Issues of domain specificity in development, autism, and brain imaging—Festschrift for Uta Frith. *Quarterly Journal of Experimental Psychology*. Volume 61, Issue 1 January 2008 , pages 76 – 89.

Perrett, D.I., Smith, P.A.J., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D., & Jeeves, M.A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. Proceedings of the Royal Society of London, Series B: Biological Sciences, 223, 293–317

Perrett, D.I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., et al. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology*; 146: 87-113.

Perrett, D. I., Mistlin, A. J., Harries, M. H., & Chitty, A. J. (1990). Understanding the visual appearance and consequence of hand actions. In: Goodale MA, editor. *Vision and action: the control of grasping*. Norwood (NJ): Ablex: 163-80.

Poljac, Schie & Bekkering. 2009. Understanding the flexibility of action–perception coupling, Psychological Research, 73:578–586.

Potter, J. (2000). Post cognitivist psychology, *Theory and Psychology, 10*, 31-37.

Premack, D. G. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1, 515-526.

Prinz, J.J. (2002). *Furnishing the mind- concepts and their perceptual basis.* MIT Press.

Prinz, J.J. (2005). The return of concept empiricism. In H. Cohen & C. Lefebvre (Eds.) *Handbook of categorization in cognitive science.* St Louis: Elsevier.

Prinz, J.J. (forthcoming) *The conscious brain*. New York: Oxford University Press.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology.* 9, 129–154.

Puce, A. & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society London B. Biological Sciences.* 358, 435–445.

Quine, W.V. O. (1960). *Word and Object*, MIT Press.

Ramachandran, V. S. (2000) "MIRROR NEURONS and imitation learning as the driving force behind "the great leap forward" in human evolution" Edge online http://www.edge.org/3rd_culture/ramachandran/ramachandran_p1.html

Raos, V., Umiltá, M.-A. Akira Murata, 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, p.709-729.

Rizzolatti, G. (1987). Functional organization of area 6. In G. Bock, M. O'Connor, and J. Marsh (eds): Motor Areas of the Cerebral Cortex (CIBA. Foundation Symposium 132). Chichester: John Wiley and Sons, pp. 171-186.

Rizzolatti, G., Gentilucci, M., Fogassi, L., Luppino, G., Matelli, M., & Ponzoni-Maggi, S. (1987) Neurons related to goal-directed motor acts in inferior area 6 of the macaque monkey. *Exp. Brain Res.* 67: 220–224.

Rizzolatti, G., Camarda, R. Fogassi, M. Gentilucci, M. Luppino, G., & 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.* 71:491-507.

Rizzolatti, G. & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*(1), 169-192.

Rizzolatti, G. & Fabbri-Destro M. (2008). The mirror system and its role in social cognition. *Current Opinion in Neurobiology,*18 (2):179–184.

Rizzolatti G & Fabbri-Destro M. (2009). Mirror neurons: from discovery to autism. Exp Brain Res. Sep 18.

Rizzolatti, G., & Fadiga, L. (1998) Grasping objects and grasping action meanings: the dual role of monkey rostroventral premotor cortex (area F5). *Novartis Found Symp.;*218:81-95; discussion 95-103.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3, 131-141.

Rizzolatti G., Fadiga L., Fogassi L., & Gallese V. (1999). Resonance behaviors and mirror neurons. Archives Italiennes de biologie, 137, pp.85-100.

Rizzolatti, G., Fogassi, L. & Gallese, V. (1997). Parietal cortex: From sight to action. *Current Opinions in Neurobiology* 7:562–67.

Rizzolatti, G., Fogassi, L. & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience* 2:661–670.

Rizzolatti, G., Fogassi L., & Gallese V. (2002). Motor and cognitive functions of the ventral premotor cortex. *Current Opinion in Neurobiology, 12*(2), 149.

Rizzolatti, G., & Gentilucci, M. (1988). Motor and visual-motor functions of the premotor cortex. In Rakic P. and Singer W. (eds.) *Neurobiology of Neocortex.* Wiley, Chichester, pp. 269-284.

Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron, 31,* 889–901.

Rizzolatti, G., & Wolpert, D. M. (2005). Motor systems. *Current Opinion in Neurobiology, 15*(6), 623.

Rizzolatti, G. and Arbib, M. A. (1998). Language within our grasp. *Trends in Neuroscience 21*, 188–194.

Rock, I. (1983). *The logic of perception*. MIT Press

Rock, I. (ed.) (1997) *Indirect perception*. MIT Press.

Rosenthal, D. M. (2005). *Consciousness and Mind*, Oxford: Clarendon Press.

Rozzi, S., Ferrari, P. F., Bonini, L., Rizzolatti, G. & Fogassi, L. (2008). Functional organization of inferior parietal lobule convexity in the macaque monkey: electrophysiological characterization of motor, sensory and mirror responses and their correlation with cytoarchitectonic areas. *The European Journal of Neuroscience*; 28: 1569–88.

Ryle, G. (1949). The Concept of Mind. New University of Chicago Press edition (2002 ed.)

Saxe, R. (2005). Against simulation: the argument from error. *Trends in Cognitive Sciences* 9, 174–179.

Saxe, R. (2006). Why and how to study Theory of Mind with fMRI. *Brain Research*, 1079, 57–65.

Saxe, R. et al. (2004). Understanding other minds: linking develop- mental psychology and functional neuroimaging. American Review of Psychology 55, 87–124.

Saxe, R. and Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in 'theory of mind', *NeuroImage* 19, 1835-1842.

Saxe, R. and Powell, L. J. (2006). It's the Thought That Counts: Specific Brain Regions for One Component of Theory of Mind, *Psychological Science*, Vol. 17, No. 8, 692-699.

Scarantino, A. (2003). Affordances Explained, Philosophy of Science, Vol. 70, No. 5, Proceedings of the 2002 Biennial Meeting of the Philosophy of Science Association. Part I: Contributed Papers, pp. 949-961.

Scholl, B.J. & Leslie, A.M. (1999). Modularity, development and 'theory of mind'. *Mind & Language*. 14, 131–153.

Schott, N. & Munzert, J. (2002). Mental chronometry in the elderly. Journal of Sport and Exercise Psychology, 24, 109.

Schubotz, R. (2007). Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, Vol. 11, No.5. 211-218.

Schubotz, R. I., & von Cramon, D. Y. (2002). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: An fMRI study. *NeuroImage, 15,* 787–796.

Schubotz, R. I. & von Cramon, D. Y. (2004). Sequences of Abstract Nonbiological Stimuli Share Ventral Premotor Cortex with Action Observation and Imagery, *The Journal of Neuroscience*, 24 (24): 5467-5474.

Schütz-Bosbach, S. & Prinz, W. (2007). Perceptual resonance: Action-induced modulation of perception. *Trends in Cognitive Sciences*, 11(8), 349-355.

Schütz-Bosbach, S., Mancini, B., Aglioti, S. M. & Haggard, P. (2006). Self and Other in the Human Motor System, *Current Biology* 16, 1830-1834.

Searle, J. (1983). Intentionality: An Essay in the Philosophy of Mind, Cambridge University Press, New York.

Sebanz, N., & Frith, C. (2004). Beyond simulation? Neural mechanisms for predicting the actions of others. *Nature Neuroscience*, 7(1), 5–6.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition, 88,* B11-B21.

Shimada, S. & Hiraki, K. (2006). Infant's brains respond differently to live and televised action. *Neuroimage* 32, p.930-939.

Southgate, V. & Hamilton, A. F. (2008). Unbroken mirrors: challenging a theory of autism. *Trends in Cognitive Sciences*, 12:225–229.

Southgate, V., Johnson, M. H., Osborne, T., & Csibra, G. (2009). Predictive motor activation during action observation in human infants. *Biology Letters*, 5, 769–772.

Southgate,V., Johnson, M. H. & Csibra, G. (unpublished). Infants attribute goals to biologically impossible actions.

Stamenov, M. L. & Gallese, V.(Eds.) (2002). *Mirror Neurons and the Evolution of Brain and Language*. Amsterdam: John Benjamins.

Stelmach, G. E., Castiello, U., & Jeannerod, M. (1994). Orienting the finger opposition space during prehension movements. Journal of Motor Behavior, 26, 178-186.

Stich, S. & Nichols, S. (1997). Cognitive penetrability, rationality, and restricted simulation. Mind and Language 12, pp. 297–326

Stock, A. & Stock, C. (2004). A short history of ideo-motor action. *Psychological Research*, Vol. 68, 176-188.

Thelen, E. and Smith, L. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.

Tomasello, M. (2009). *Why We Cooperate*. MIT Press

Tomasello, M. (2008).. *Origins of Human Communication*. MIT Press.

Tucker, M. & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance*. Vol 24(3), 830-846.

Tunik, E., Frey S.H. and Grafton, S.T. (2005). Virtual lesions of the anterior intraparietal area disrupt goal-dependent on-line adjustments of grasp, Nature Neuroscience 8 (4), pp. 505–511.

Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C. & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. *Neuron*, 32, 91-101.

Umiltà, M.A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., Jezzini, A., Gallese, V. & Rizzolatti, G. (2008). When pliers become fingers in the monkey motor system. *Proc Natl Acad Sci* U S A 105:2209–2213.

Van Schie, H., van Waterschoot, B. & Bekkering, H. (*In press)*. Understanding action beyond imitation. *Brain Research*.

van Schie, H. T., van Waterschoot, B. M., & Bekkering, H. (2008). Understanding action beyond imitation: Reversed compatibility effects of action observation in imitation and joint action. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1493–1500.

Varela, F., Thompson, E., Rosch, E. (1991). *The Embodied Mind*, MIT Press, Cambridge, MA.

Wallace, B., Ross, A., Davies, J. & Anderson, T. (2007). The Mind, the Body and the World: Psychology After Cognitivism? Exeter: Imprint Academic

Warneken, F. & Tomasello, M. (2007). Helping and cooperation at 14 months of age. Infancy, 11(3), 271–294.

Wicker, B., Keysers, C., Plailly, J., Royet, J.P., Gallese, V., and Rizzo- latti, G. (2003). Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. Neuron 40, 655–664.

Wilson, M. & Knoblich, G. (2005). The case for motor involvement in Perceiving Conspecifics. *Psychological Bulletin*, Vol. 131, No. 3, 460-473.

Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition, 13, 103-128.

Wittgenstein, L. (1953). *Philosophische Untersuchungen* (Eng terans: *Philosophical Investigations*, translated by G.E.M. Anscombe (1953).

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition* 69:1–34.

Woodward, J. (2003). *Making Things Happen: a Theory of Causal Explanation*, Oxford.

Zahavi, D. (2001). Beyond Empathy. Phenomenological approaches to intersubjectivity, *Journal of Consciousness Studies* 8/5-7, 151-167

Zahavi, D. (2008). Simulation, Projection and empathy. Consciousness and Cognition 17: 514-522.

Zaitchik, D. (1990). When representations conflict with reality: the preschooler's problem with false beliefs and "false" photographs. *Cognition, 35,* 41-68.