



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Head Office: Università degli Studi di Padova

Department of Philosophy, Sociology, Education and Applied Psychology (FISPPA)

Ph.D. COURSE IN PHILOSOPHY

CURRICULUM: THEORETICAL AND PRACTICAL PHILOSOPHY

SERIES: XXXII

THE PROBLEMS OF CONSCIOUSNESS AND ITS RELATIONSHIP WITH THE WORLD.

AN INVESTIGATION BETWEEN PHILOSOPHY AND NEUROSCIENCE

Coordinator: Ch.ma Prof.ssa Francesca Menegoni

Supervisor: Ch.mo Prof. Antonio Da Re

Ph.D. student: Federico Zilio

Table of Contents

Abstract.....	7
Introduction.....	9
1. The Problems of Consciousness.....	17
1.1 Consciousness is said “in many ways”	17
1.1.1 Consciousness by Definition, Method, Perspective and Theory.....	19
1.2 The Hardness of Consciousness	24
1.2.1 The Nature of Qualia.....	33
1.3 The Meta-problem of Consciousness	38
1.3.1 What does “Meta” mean?.....	39
1.3.2 The Specificity of Consciousness	42
1.3.3 Consciousness as Explanans	48
1.3.4 The Illusionist position on Consciousness	52
2. The Logical and Phenomenological Structure of Consciousness	55
2.1 What Starting Definition for Consciousness?	55
2.1.1 Kripke’s Direct Reference Theory	58
2.1.2 A Kripkean-like definition of Consciousness	60
2.1.3 A necessary a posteriori Statement for Consciousness?	66
2.1.4 The Consciousness-Brain Relation	68
2.2 What am I aware of when I am aware? On Pre-reflective Self-awareness	74
2.3 Regarding the possibility of Unconscious Experiences	79
2.3.1 The Refrigerator Light Problem and the Long-distance Truck Driver	80
2.3.1.1. Bodily Awareness as Pre-reflective Self-consciousness.....	86
2.3.2 Unconscious Perception and the Blindsight case.....	92
3. Neuroscience of Consciousness and Neuro-centrism.....	99
3.1 The Neuroscience of Consciousness	100
3.1.1 Approaches and Operationalization of Consciousness	101
3.1.2 The Neural Correlates of Consciousness	107
3.1.2.1 Criticisms and Refinements of NCCs	110
3.1.3 Drastic Scientific Approaches towards Consciousness.....	113
3.1.3.1 Reduction: Daniel Dennett’s Heterophenomenology	114
3.1.3.2 Transmutation: Churchlands’ Eliminative Materialism.....	117

3.1.3.3	Dismissal: Elisabeth Irvine's Scientific Eliminativism	119
3.2	The Roots of Neuro-centrism.....	121
3.2.1	The Brainhood condition	123
3.2.2	Pre-modern concepts of Brain and Consciousness	125
3.2.3	Modernity as the Cradle of the contemporary concept of Consciousness.....	128
3.2.4	From Modern Philosophy to Contemporary Cognitive Sciences ...	135
3.2.5	The Brainhood condition and the Body-Machine model.....	137
4.	The Appearance-Reality Fallacy as the Root of contemporary Studies of Consciousness.....	141
4.1	The Cartesianism in contemporary Cognitive Sciences.....	142
4.1.1	The persistent Primary-Secondary Qualities Dichotomy	145
4.2	Consciousness beyond Surface and Deep Properties.....	147
4.2.1	The intrinsic Perspectivalness of Consciousness.....	151
4.3	The Appearance-Reality Fallacy	154
4.3.1	The Insufficiency of the Physicalist account of Reality	159
4.3.2	Lifeworld as the Horizon of all the Worlds	161
4.3.3	From Ontological Pluralism to an Interdisciplinary and Relational Approach to Consciousness.....	165
4.4	Problems with Neuro- and Mento-centrism	169
4.4.1	Neuro-centrism	170
4.4.1.1	The Myth of Localization	171
4.4.1.2	The Mereological Fallacy	174
4.4.1.3	Fallacy of Misplaced Concreteness	175
4.4.2	Mento-centrism.....	176
4.4.2.1	Hypostatization of Psychic Objects	177
4.4.2.2	Confusing Actual and Objectual Consciousness	181
5.	Pluralist and Interdisciplinary Approaches to Consciousness.....	185
5.1	Looking for an Alternative Approach	186
5.1.1	Regional Ontologies	186
5.1.2	The Relation between Phenomenology, Philosophy and Neuroscience	191
5.1.2.1	Neurophenomenology.....	193
5.1.2.2	Non-reductive Neurophilosophy.....	196
5.1.2.3	Neuropragmatism and 4E Cognition	199
5.1.2.4	Philosophy of Neuroscience and Neuroscience of Philosophy	205

5.2	Conceptual and Phenomenological analysis: Non-mentalistic and Non-physicalist account of Consciousness	206
5.2.1	Intentionality with No Contents: Anti-idealism and Phenomenological Realism.....	206
5.2.2	Phenomenality I: Sartre and the Intentionalist approach	209
5.2.2.1	Perception and Imagination: the Disjunctivist approach	213
5.2.3	Phenomenality II: Self-awareness and the problem of Subjectivity	216
5.2.3.1	Pre-reflective Non-positional Consciousness	220
5.2.4	Intentionality goes Adverbial.....	224
6.	Experience as Embodied-Being-in-the-World and the Environment-Body-Brain Relation. A Neurophilosophical account	229
6.1	Consciousness as Embodied-Being-in-the-World.....	230
6.1.1	Why Sartre?.....	231
6.1.2	The Body Surpassed towards the World.....	233
6.1.2.1	Embodiment, Embedment, Extendedness, Enactment.....	237
6.1.3	The Primacy of the World.....	243
6.2	The Worldhood Paradigm	246
6.2.1	Being-in-the-World as Spatio-temporality.....	248
6.2.1.1	Temporality and Spatiality	249
6.2.1.2	Brainhood versus Worldhood on the interpretation of experience	252
6.3	Neurophilosophy of Consciousness	257
6.3.1	World and Body in neuroscientific Theories of Consciousness	258
6.3.1.1	Integrated Information Theory	258
6.3.1.2	Global Neuronal Workspace Theory	259
6.3.1.3	Temporo-spatial Theory of Consciousness.....	260
6.3.2	Northoff's World-Brain Relation and the link between Neuroecology and Phenomenology	263
6.4	<i>Experimentum crucis</i> : consciousness without body and environment?	271
6.4.1	Locked-in Syndrome	272
6.4.1.1	4E cognition challenged by LIS.....	276
6.4.1.2	LIS as “locked inside” but not outside of the world	278
6.4.1.3	Complete Locked-in Syndrome	279
6.4.1.4	Environment-Body-Brain relation as Predisposition of Consciousness.....	281
6.4.1.5	CLIS as potentially being-out-of-the-world.....	285

6.4.2 Dreaming: Disconnected Consciousness?	289
6.4.2.1 Hallucinations and Dream Worlds.....	291
Conclusion	295
References	305

Abstract

The main subject of this dissertation is the relationship of consciousness with the world. Consciousness is extremely familiar and mysterious at the same time, and it is considered one of the most complex phenomena to describe and explain. Recently, cognitive science and neuroscience have assumed a significant role in the search for the fundamental structures of consciousness, investigating the neuronal basis that allows the emergence of our experience. For this reason, the present work is not only aimed at philosophical and phenomenological analysis but also at the latest results of neuroscience on the relationship between consciousness and brain.

The first chapter outlines the main topics of contemporary consciousness studies (first-, second-, third-person perspective, the nature of qualia, phenomenal contents, the hard problem and the meta-problem of consciousness, etc.). After presenting a provisional model for the phenomenological (Sartre) and logical (Kripke) structure of consciousness, represented by intentionality, phenomenality and self-awareness (Chapter 2), I will critically discuss reductive approaches that are influential both in neuroscience and in philosophy of mind. Some of them support the possibility to conceptually and ontologically reduce experience to brain activity (neuro-centrism; Chapter 3), while others perform the hypostatization of numerous mental states within the mind/brain (mento-centrism; Chapter 4).

In this regard, it is possible to identify a common paradigm, that is, the distinction between reality and appearance that occurred especially during modern philosophy (in particular Descartes) and after the modern scientific revolution (in particular Galilei). That has led to the reduction of experience to a mind- or brain-related quality/property, as we can see nowadays (Chapter 4). Given the perspectival, multifaceted structure of experience and following the analysis of several non-reductive approaches, I will argue for an ontologically and epistemologically pluralist account of consciousness, that it can be developed only through an interdisciplinary approach, by philosophy, phenomenology and neuroscience together (Chapter 5).

This will allow me to address the issue of consciousness from the so-called non-reductive neurophilosophical approach, through which I will analyse the way we experience the world from both a phenomenological and a neuroscientific perspective. Supported by Sartre's phenomenological analyses of consciousness in relation to world and body, I argue that experience is not given without the embodied-being-in-the-world condition, a condition that will then be connected to the neuroscientific concept of environment-body-brain alignment (adapted from Northoff's world-brain relation), that is, a necessary predisposition for the maintenance of consciousness (Chapter 6). In conclusion, the concept of consciousness-in-the-world will be tested through neurophilosophical experiments regarding the Locked-in syndrome and the dreaming state, suggesting the need to reconsider consciousness not merely in terms of brain- or mind-relation, but intrinsically as world-related.

Introduction

Here was just the thing [Sartre] had been longing to achieve for years – to describe objects just as he saw and touched them, and extract philosophy from the process.

De Beauvoir, 1960/1965, p. 135

I am sitting in front of my desk as I type these letters on my laptop keyboard, my fingers spin fast as I watch the words appearing on the screen but at the same time I reflect on myself reading the very same words – it is a strange sensation. The lava lamp (so vintage) dimly illuminates the pages of a book on my left, from time to time my gaze is captured by the slow movement of the coloured bubbles inside. I unknowingly move my hand towards the cup of tea to my right, but my eyes are so focused on this paragraph that a finger sticks on the greases of the succulent plant that decorates the desk (it looks very bad, I do not have a green thumb). I immediately withdraw the hand in pain and take off the headphones that produced strange ambient sounds and whisperers, just as the annoying buzz of a mosquito shows up behind me. Well, it would be easier to write if I were not constantly wrapped by and immersed in a world around me while I do it. But this is the condition that determines our experience, that is, to be in the world. A long list of issues and themes emerges from this apparently banal paragraph: intentionality, bodily awareness, self-consciousness, unconscious perception, pain-consciousness, phenomenality, and so forth. Another equally long list of questions emerges and understanding and discussing these problems is the task of the present research. What is consciousness? How does it “work”? What are the conditions of possibility of consciousness? Am I always conscious of the environment? Where is my body when I am immersed in some specific task? What does it feel like sitting on this desk writing on the computer? What happens to my body and my brain while I do it? What relationship does my body have with the world? How does the world look full of meanings to me? Paraphrasing the words of Aron to Sartre

in 1932¹: if you are a phenomenologist, you can talk about this paragraph and make philosophy out of it!

As the title says, the present research aims to investigate the problems of consciousness, with a focus on how we – as conscious subjects – are related to the world. A brief analysis of the title can already suggest a lot about what will be exposed in this research. I used “problems” because consciousness is not a unitary and solid philosophical problem, rather it presents several issues from different perspectives, but all linked together. It is indeed the most pervasive phenomenon we can share and at the same time one of the most evanescent to explain, so elusive as to be difficult to describe by rigorous scientific approaches. The main argument addressed here will be the potential relations that consciousness can instantiate with the objects of the environment and with (through) our body; in a nutshell, with the world. As I figuratively showed in the previous paragraph, our experience never seems to be made of brief flashes of consciousness in specific space-time points, rather it seems to be a fluid and total immersion in a world made of objects and backgrounds, meanings and references, actions and perceptions. Thus, it seems impossible to examine the phenomenon of consciousness without questioning how we engage the world we live in.

The relationship between consciousness and world is one of the most important fundamentals of classical and contemporary phenomenology, and it will be, at the same time, the starting point and the main topic of this research. However, this will not be developed only from a philosophical or phenomenological point of view but will also be supported by a neuroscientific perspective. Since the last decades, in fact, consciousness cannot be defined anymore as the exclusive domain of phenomenology or philosophy of mind, given that several scientific articles on the neural correlates of consciousness have been published, with the aim of explaining the relationship

¹ «We spent an evening together at the Bec de Gaz in the Rue Montparnasse. We ordered the speciality of the house, apricot cocktails; Aron said, pointing to his glass, ‘You see, my dear fellow, if you are a phenomenologist, you can talk about this cocktail and make philosophy out of it!’ Sartre turned pale with emotion at this. Here was just the thing he had been longing to achieve for years – to describe objects just as he saw and touched them, and extract philosophy from the process. Aron convinced him that phenomenology exactly fitted in with his special preoccupations: by-passing the antithesis of idealism and realism, affirming simultaneously both the supremacy of consciousness and the reality of the visible world as it appears to our senses» (De Beauvoir, 1960/1965, p. 135).

between brain activity and our own mental life, especially our conscious experience. For this reason, if the purpose here is to investigate the current state of consciousness studies, analysing critically their characteristics with reference to the consciousness-world relationship, and in order to offer an original perspective about it, then it is not possible to ignore recent neuroscientific developments. It would be like presuming to talk about life completely ignoring biological studies on the subject. That explains why I put in the title “An investigation between philosophy and neuroscience”, where by “philosophy” I mean a whole series of tools of historical-conceptual analysis and phenomenological description. In this sense, the present work aims to address the issue of consciousness through a mainly philosophical approach, but which is able at the same time to establish links on an interdisciplinary level with the results of neuroscientific research. Thus, the epistemological and methodological goal is to present an interdisciplinary approach towards experience, which can consequently offer an ontologically pluralistic account of consciousness.

The first chapter, “The problems of consciousness”, highlights the complexity of the issue of consciousness. The concept of consciousness is here examined in terms of its possible meanings, from different methods and perspective, exposing the main philosophical positions over it. This immediately raises the epistemological question about adopting a perspective on consciousness which cannot only be an objectifying one, i.e. in the third person, but needs to include also the first and second person ones. There is also a large section dedicated to perhaps the most famous philosophical issue regarding consciousness, that is, Chalmers’ hard problem that argues for the irreducibility of experience to the physicalist and materialist paradigms of contemporary science. This will lead me to discuss several collateral issues that addresses consciousness from both ontological and epistemological perspectives, for example, the explanatory gap, the what-it-is-likeness of experience and the nature of qualia. The examination of these points will help me to outline what I define as the “specificity” of consciousness, that is, the ontological incommensurability that we usually report about consciousness with respect to all the other phenomena in the world, and that Chalmers has recently defined the meta-problem of consciousness. The last part of the chapter is dedicated to this new kind of metaphysical problem, which

also offers the opportunity to expose the two main philosophical positions concerning consciousness, realism and illusionism.

After a general introduction to the problems of consciousness, the second chapter, “The logical and phenomenological structure of consciousness” starts with an attempt to identify a provisional blueprint of consciousness. In other words, in the midst of so much critical literature about the problems of consciousness, I will try to start the analysis of the phenomenon of consciousness from the very same definition that we generally give to the term “consciousness”. In this regard, I will use Kripke’s direct reference theory, in order to analyse the term “consciousness” in a non-descriptive way, as if “consciousness” were a rigid designator in the same way as the terms “gold” or “tiger” with their respective natural kinds. The Kripkean discussion will also focus on the validity of the definition of “consciousness” as “neural activity x”, which is also affirmed by many supporters of the mind-brain identity. Secondly, after offering the provisional definition of consciousness as “being aware of something”, I will try to determine the phenomenological structure of our experience, focusing on the issue of pre-reflective consciousness, developing a fruitful dialogue with various contributions from phenomenology (Sartre, Merleau-Ponty, Gallagher, Zahavi), and delineating some preliminary features, like phenomenality, intentionality and self-awareness. These features will then be tested in the last part of the chapter, through some peculiar thought and real experiments on the possibility of unconscious experiences (refrigerator light problem, long-distance truck driver, blindsight).

After providing a provisional definition and phenomenological structure to the concept of consciousness, there still are several questions about consciousness, from both the philosophical and scientific levels. In the third chapter, “Neuroscience of consciousness and neuro-centrism”, I will look into the contemporary scientific context, and more specifically into the neuroscientific studies on consciousness. If the purpose of neuroscience is to provide an explanation of consciousness, we must discuss about what a science – and, in particular, a neuroscience – of consciousness would be. A scientific account, in a very general way, should offer a theory that could explain, describe and predict the phenomenon of consciousness in all its features. In other words, the theory must define what consciousness is (description), how it comes to existence (explanation), which its functions are (functionality). In this regard, I

expose and discuss various approaches and operationalizations in neuroscience (brain-based, mind-based, brain-reductive, mind-reductive), in particular the typical method of description of consciousness in terms of “contents” and “levels”. Then I critically examine the core theme of the neurosciences of consciousness, that is, the search for the so-called neural correlates of consciousness (NCC), which is often considered the holy grail that will reveal us all the secrets of experience. More generally, most of contemporary scientific researches on consciousness maintains that consciousness can be now explained through all the new methods of neuroimaging available, and that in this way a complete description of how we experience can be reduced to the mechanisms of the brain. This kind of attitude towards consciousness – neuro-centrism – is based on the idea that our personality, our mental life, our consciousness itself are all products or equivalent to our brain activity (the brainhood paradigm), which finds its roots in the modern scientific revolution and philosophy. About that, I will delineate an historical-epistemological analysis of the brainhood paradigm, trying to establish a comparison between, on the one hand, Galileo and Descartes and, on the other hand, the contemporary reductionist and eliminativist approaches to consciousness.

In the fourth chapter, titled “The appearance-reality fallacy as the root of contemporary studies of consciousness”, I proceed with the analysis of the cultural, epistemological and ontological basis on which contemporary researches on consciousness are grounded, including not only empirical studies but also analytical studies from the philosophy of mind. The thesis here is that Cartesianism is diffusely present in contemporary cognitive sciences and also that the modern dichotomy between primary and secondary qualities, instantiated first by Galilei, continues to exercise a considerable influence even today. This represents the background within which what I call the Appearance-Reality Fallacy takes shape, which is the epistemological and ontological error to separate the world and also consciousness in two realms, the apparent/superficial one and the real/deep one, as if the laptop in front of me right now was just the appearance of a truer reality beyond it, like a thing-in-itself. This fallacy is present both in neuroscience and philosophy of mind and I will present a series of problems concerning neuro-centrism (myth of localization, mereological fallacy, fallacy of misplaced concreteness) as well as mento-centrism (Sartre’s critique of mentalism, i.e. hypostatization of psychic objects), in order to

avoid the immediate objectification of experience both in neuronal and mental sense. In this regard, on the basis of the analysis prepared in the first chapter, I will distinguish between actual and objectual consciousness, identifying the subject of the subsequent phenomenological analysis in the intrinsic and unavoidable perspectivalness of the actual consciousness, while instead most of contemporary studies on consciousness are focused on how consciousness is objectified by means of neuronal correlation or philosophical concepts (representations, qualia, self, etc.). This will lead me to propose ways of escape from reductionist naturalisms or irreconcilable dualisms, towards an alternative approach based on ontological pluralism and interdisciplinary epistemology.

While in chapter four I simply suggest a new interdisciplinary and pluralist approach on the basis of the problems and criticisms of the mentalist and neuro-centric paradigms, in chapter five, “Pluralist and Interdisciplinary Approaches to Consciousness” I propose an alternative perspective, mainly grounded on Sartre’s phenomenological ontology. First, I present a method of analysis that considers various dimensions of our experience, physical, phenomenal, psychological, etc., which consequently needs the contribution of many disciplines; in particular I will argue for a fruitful interconnection between phenomenology, philosophy and neuroscience, exposing different types of non-reductive approaches (neurophenomenology, non-reductive neurophilosophy, neuropragmatism, 4E cognition). Secondly, regarding the concept of consciousness as such, I focus on the lived experience as the starting point and at the same time the presupposition of every other experience (introspection, empirical analysis, conceptualization, etc.). From here, the analysis will focus mainly on the necessary condition of consciousness as being-in-the-world, in a concrete sense (we exist as immersed in an objectual and pre-given context) as well as the activity of sense-disclosing (consciousness is at the same time an activity of meaning-apprehending and meaning-making in the world). In this sense, the provisional features I have introduced in the second chapter, will here be inserted within the context of consciousness as being-in-the-world, which works as a pre-condition for intentionality, phenomenality and self-awareness. The attempt will be to propose a non-mentalistic and non-reductionist account of experience, starting with a philosophical-phenomenological analysis that will then develop it in chapter six

through a phenomenological-neuroscientific approach. In the last part of the chapter, I will also build a bridge between Sartre's phenomenology and some contemporary philosophical position about consciousness, like intentionalism, representationalism, disjunctivism and adverbialism, to ensure greater conceptual clarity.

This long itinerary finds its final synthesis in the last chapter six, "Experience as Embodied-Being-in-the-World and the Environment-Body-Brain Alignment. A Neurophilosophical account", where I expose in detail an account of consciousness understood as embodied-being-in-the-world (phenomenological perspective) and environment-body-brain alignment (neuroscientific perspective). The aim indeed is to give bones, flesh and tendons to this interdisciplinary theory, focusing on the relation between the body and the world, from both phenomenology and neuroscience. First, I will rely on Sartre's early phenomenological studies (from *The Imagination* to *Being and Nothingness*), in particular to re-evaluate his complex analysis of corporeity, often unfairly overshadowed by the studies on Husserl and Merleau-Ponty. In this sense, I will analyse bodily consciousness as intrinsically situated and connected with the world, through a body that is often "surpassed" and "passed by in silence", to use Sartre's words. In order to build a bridge with contemporary studies of consciousness and the body, I will compare Sartre's phenomenological analysis with the key concepts of the 4E Cognition (embodied, embedded, enacted, extended). This will also lead me to consider the possibility that, in order to understand how we exist our body, we need to understand first how we "exist" the world, in search of a new paradigm for the study of consciousness that is not founded neither on the brain nor on the body alone, but on our relationship with the world. This relationship may be called "embodied-being-in-the-world" and I will expose and discuss it, highlighting its characteristics in relation to how we experience the world in a spatio-temporal way.

Finally, in order to be faithful to my interdisciplinary project, I will analyse consciousness from a neurophilosophical point of view, discussing how the consciousness-world connection is conceived by some of the most relevant neuroscientific theories of consciousness (Integrated Information Theory, Global Neuronal Workspace Theory, Temporo-spatial Theory of Consciousness), consequently relating the embodied-being-in-the-world condition to the concept of "world-brain relation" proposed by Georg Northoff, neuropsychiatrist and

neurophilosopher. From here, I will discuss Northoff's theory and give an account for what I call the environment-body-brain alignment as the empirical description of the ontological concept of world-brain relation, and also as a necessary precondition for the maintenance of consciousness. This will involve a final test for the theory here exposed, both at a phenomenological and a neuroscientific level, a sort of *experimentum crucis* that can help us to understand if I can really provide a description of our experience as embodied-being-in-the-world and environment-body-brain alignment. In this regard, therefore, I will propose two cases in which consciousness seems to exist in absence of relationship with the body and the environment, namely the Locked-in Syndrome, a rare and severe clinical condition, and the dreaming state intended as a case of "disconnected consciousness". These paradigmatic cases can offer us an empirical, phenomenological and conceptual test to the theory presented in this research, a neurophilosophical and neurophenomenological account of experience that is, above all, based on the concept of being in connection with the surrounding world, as I described it at the beginning of this introduction. This concept of consciousness-in-the-world will suggest the need to reconsider consciousness not merely in terms of brain- or mind-relation, but intrinsically as world-related.

1. The Problems of Consciousness

The first chapter highlights the complexity of the issue of consciousness. The concept of consciousness is here examined in terms of its possible meanings, from different methods and perspective, exposing the main philosophical positions over it. This immediately raises the epistemological question about adopting a perspective on consciousness which cannot only be an objectifying one, i.e. in the third person, but needs to include also the first and second person ones. I will dedicate a large section to Chalmers' hard problem that argues for the irreducibility of experience to the physicalist and materialist paradigms of contemporary science. This will lead me to discuss several collateral issues that addresses consciousness from both ontological and epistemological perspectives, for example, the explanatory gap, the what-it-is-likeness of experience and the nature of qualia. The examination of these points will help me to outline what I define as the "specificity" of consciousness, that is, the ontological incommensurability that we usually report about consciousness with respect to all the other phenomena in the world, and that Chalmers has recently defined the meta-problem of consciousness. The last part of the chapter is dedicated to this new kind of metaphysical problem, which also offers the opportunity to expose the two main philosophical positions concerning consciousness, realism and illusionism, where I will argue for the impossibility to deny in a coherent way the existence of acts of consciousness.

1.1 Consciousness is said "in many ways"

The concept of consciousness appears ambiguous and complex, devoid of precise definitions and full of variable meanings (Ned Block even defined it as a "mongrel concept"; Block, 1995). Nevertheless, at the same time, we seem to immediately apprehend the relevance of the term through a kind of natural and pragmatic acquaintance of it. In effect, before being a controversial umbrella concept discussed

in philosophy of mind and more recently in neuroscience, consciousness refers *prima facie* to the simple «states of sentience or awareness that typically begin when we wake up in the morning from a dreamless sleep and continue throughout the day until we fall asleep again» (Searle, 1999, pp. 40–41)¹. The word “consciousness” represents both the most familiar sensation of being awake and aware of the surrounding world and one of the most evanescent and opaque conceptual issue of contemporary philosophy and science. In the words of David Chalmers: «[t]here is nothing that we know more intimately than conscious experience, but there is nothing that is harder to explain» (Chalmers, 1995, p. 200).

Consciousness has not always been a matter of scientific interest. For a long time, the study of consciousness has been an exclusive privilege of philosophy. Indeed, it seems paradoxical, or at least odd, to think about consciousness as a “matter” of scientific research, and it seems apparently impossible to catch such a phenomenon with a scientific method, for many reasons I am going to illustrate below. Actually, up until a little more than a century ago it seemed impossible to empirically see an electron or, better still, to discover the elementary particles of the universe, however nowadays we know that it was only a contingent impossibility, progressively dissolved by the scientific progress. As a matter of fact, what has been for long an exclusive matter for theoretical and conceptual analysis of philosophy, has become over time an object of experimentation and empirical consideration. Just consider the revolution from philosophical theories about the nature of the universe to the scientific determination of the law of physics.

For that reason, it seems consistent to say that also consciousness might be a potential subject of scientific study, thanks to new methodologies, disciplines and technologies. The huge increment of neuroscientific researches about consciousness can be a proof of this belief²: in the last decades, neuroscientists, besides focusing on perceptual and cognitive abilities – and impairments – of the brain, they also began to deal with the concept of consciousness as such, trying new quantitative measurements within the empirical and experimental domain of cognitive sciences. Consequently,

¹ Also, Strawson: «We know that consciousness exists, we know this for certain, and we know what it is. There is a fundamental sense in which we know *exactly* what it is» (Strawson, 2017).

² Medline trend for “Consciousness” (year: number of citations): 1997: 703; 2007: 1319; 2017: 2280.

they started to talk about “visual consciousness”, “perceptual consciousness”, “level and content of consciousness”, “arousal and cortical awareness”, “attention”, “self”, “altered forms of waking consciousness” (trance, absorption, hypnosis, dissociation, meditative states, drug states, and out-of-body experiences), “REM/dreaming” versus “slow-wave/deep sleep states”, etc. (Faw, 2009; Laureys & Tononi, 2009b).

1.1.1 Consciousness by Definition, Method, Perspective and Theory

As said, consciousness is a very general and vague term and there is not a univocal definition for it³. It seems more like an umbrella concept with a great range of significances (Van Gulick, 2018) and it could be interpreted in many different ways: sentience, i.e. the ability of sensing and responding to the environment; wakefulness, i.e. the awake and alert state in opposition to sleep, coma or other abnormal states⁴; interaction with external and inner world; self-consciousness, i.e. the epistemic state of being aware that we are aware; the “what it is like to be” and the phenomenal content, i.e. those specific subjective characters of being in a particular conscious state; intentionality or transitive consciousness, i.e. the fact that consciousness is always and necessarily consciousness of something; awareness of surrounding objects and reportable contents. These are only some of the meanings of this concept. Ram Vimal describes and lists forty distinguishable and non-exhaustive meanings attributed to “consciousness” discovered from the literature and some cultural tradition (Vimal, 2009). These meanings can broadly be divided into those that mainly refer to

³ It is also important to note that when we talk about finding a definition for consciousness, we should first understand what we mean by “definition”. Generally speaking, a definition is the statement of the meaning of a term, how it is used by people. Of course, there are little variations in the common use of a term like consciousness, e.g. when I say “Tom is conscious” it could mean that Tom is awake and reactive to the surrounding environment, that Tom has regained consciousness after having passed out, or even that Tom is aware of something in particular. Then, to be more specific, there is also a way of understanding “definition” as the determination of the essential qualities/conditions/boundaries/extent of the entity in question (from the Latin *de-finire*: to limit) (Faw, 2009). In this case, the definition of consciousness means the determination of the necessary condition we need in order to state that there is consciousness or that someone is conscious. This evaluation, as I will show in the next chapters through some empirical case, needs more than our common sense and usage of the term “consciousness”.

⁴ Sentience and wakefulness may be grouped under the more general category of creature consciousness, as a assumed property of creatures (Rosenthal, 1993).

“function” and those that are primarily about “experience” (I shall expose later how important an intrinsic is this intuitive distinction).

The first group mostly contains meanings related to a materialistic approach to consciousness (according to Chalmers’ categorization, materialisms from type A to C; Chalmers, 2003, 2011), which aims to describe the function (how) of consciousness in nature. Just to illustrate some in addition to those already mentioned, consciousness as neurobiological processes, distributed and interrelated mental capacities (e.g. working memory, cognition, attention, inner monitoring, etc.), thought processing, transforming processing from sensory stimuli to information, stream of intentional information, acting functions involving sensorimotor interactions, dynamic organism-environment interactions, etc. Now, it is not important which one of these meanings is the most empirically plausible or logically coherent, what matters is that they all are attempts to identify a function, that is, how consciousness “works”, and which mechanisms are involved in the way it operates. Within the second group fall all those meaning related to non-materialistic but physicalist ontology (from type D to F), idealism and modern constructivism. For example, perceptual and emotional experiences, raw feelings, bodily sensations, sense of mineness, subjective experience, qualia, phenomenal characters of consciousness, pre-cognitive/pre-reflective experience, unity and continuity of stream of awareness, cosmic source from which matter emerges, feature of the universe. Differently from the previous group, these meanings, with some exception, focus mainly on a description of the essence of an experience, i.e. which characters properly define consciousness⁵. Additionally, the concept of consciousness may be used not only for humans, but also for animals (Allen & Trestman, 2017; D. B. Edelman & Seth, 2009), artificial intelligence (Aleksander, 2017), and even in the debate about the fundamental features of the universe (e.g. panpsychism, Goff, 2017; Brüntrup & Jaskolla, 2016).

Just as much interpretations of the concept of consciousness, there are also as many disciplines that approach the problem of consciousness with their own methodology, e.g. phenomenology, introspection, qualitative differences and dissociation methods, neurophysiological, psychological and behavioural methods,

⁵ As I shall expose later, this distinction between “functional” and “experience” meanings are related with the distinction between easy and hard problems of consciousness.

etc. (Irvine, 2013a; Seth et al., 2008). Regarding this methodological richness, the meaning and the ontological status of consciousness seem extremely related to the kind of approach we use for it. For example, findings from recent experimental approaches on consciousness are often in contrast with our common sense. Just consider the discordances between the apparently colourful and rich visual perception of the surrounding environment we daily have and the results from cognitive sciences about our colour blindness in the visual periphery (Dennett, 2005, pp. 82-84; Levy, 2007, pp. 40-42; Odegaard, Chang, Lau, & Cheung, 2018; for a critical examination see Haun, Tononi, Koch, & Tsuchiya, 2017; Cohen, Dennett, & Kanwisher, 2016). There is a lack of convergences even among different measurements and analysis within the same domain, and, therefore, it is not clear if the results from all these different methods refer to a single common phenomenon or a variety of features partially related to the consciousness macro-phenomenon, like cognition or memory, etc. In this sense, it might not be so easy to develop a genuine, empirical method for a scientific study of consciousness.

Moreover, we need to consider that any method/approach towards a phenomenon is epistemically dependent on the perspective we use. And interestingly, consciousness can be approached from different points of view, as well as other phenomena. We here consider the first-, second- and third-person perspectives (or epistemic stances). The peculiarity is that consciousness is – I might say – intrinsically “pluri-perspectival”, while for other natural phenomena a single epistemic stance could be enough. The first-person perspective (FPP) concerns our immediate and subjective experience of a phenomenon from a single point of view. The experience in FPP can be presented by simple reports by voice (“I see a red apple”), but not always, as the external reporting is not a necessary condition for conscious experience. FPP mainly refers to phenomenal and qualitative characters of experience, that is, how things immediately appears to us, and is related to the so-called “phenomenal consciousness” (Block, 1995; details later). There can be also a cognitive evaluation in FPP upon our own conscious states, through self-reflection, i.e. I can re-live the phenomenal experience and see myself as another so that I can objectify my own states

and conceptualize my stream of consciousness⁶. The third-person perspective (TPP) refers to objective and measurable descriptions of phenomena, through a common language and conceptual scheme. It is typical of empirical sciences, but it is also necessary for any kind of shared knowledge, since in order to make objective judgments we need a sharable and detached perspective towards an object. The second-person perspective (SPP), as the name may suggest, is halfway between FPP and TPP, and it involves intersubjectivity, because it is a relation between an epistemic subject and another sentient being's mental states (Pauen, 2012). While FPP is related to pre-reflective experience or operates through self-reflective activity from the very same point of view of the subject, the SPP occurs when we answer to some specific questions about our conscious state made by another person. This is typical of the phenomenological interview⁷, in which one subject analyses another subject, and the interviewer tries to take up an emphatic position so that the subject can also have a second-person perspective on him/herself through the other (I am aware of the other, but also of myself as addressed by the other) (Høffding & Martiny, 2016).

Regarding consciousness, all the three perspective are sources of important information, indeed we might want to analyse it in the way we consciously experience the world (FPP), through phenomenological interview (SPP) or neuropsychological measurements (TPP). As exposed here above, any single perspective has its own features but there are no unbridgeable frontiers between one specific epistemic stance and another. The three perspectives are related to specific methods, which can be put together during the conceptual and empirical analysis of consciousness. Just consider the translation of the lived experience into psychological scales like the Verbal Rating Scale (VRS), the Numerical Rating Scale (NRS) or the Visual Analog Scale (VAS), that is, the passage from FPP (the subjective experience) to third-person methodology (TPM) (the scales)⁸. On the other side, consider the experiments with the real time

⁶ There also are some examples of introspection with phenomenological training; e.g. the Experience Descriptive Sampling: DES subjects carry a random beeper in natural environments; when the beep sounds, they capture their inner experience, writing down notes about it (Heavey et al., 2010; R. T. Hurlburt & Akhter, 2006).

⁷ See, for example, the Explicitation Interview, a technique that aims to create conditions conducive to access to pre-reflective knowledge, of which the subject has no knowledge in the mode of reflective consciousness (Hogan, Hinrichs, & Hornecker, 2016; Maurel, 2009).

⁸ This is what, later and in the next chapters, I will call "operationalization", i.e. the mediation of an entity through methods and instruments (Azzouni, 2013).

fMRI-Neurofeedback (rtfMRI), during which the subject learns to modulate his/her sensations through an immediate feedback on his/her brain activations (for example, through the mental regulation of an image of a flame put in front of the eyes)⁹. This is an instance of the translation of a third-person method (the functional magnetic resonance) into FPP perspective (the self-regulating activity of the subject). As I will discuss later in details, one of the greatest challenges for future researches on consciousness will be to develop an integrated science of consciousness (Olivares et al., 2015), beyond all assumed divisions in methods, for a complete “triangulation” of the phenomenon¹⁰.

Given the richness and diversity – one might say vagueness – of definitions of consciousness and methods to approach it, there consequently are different theories that attempt to offer an exhaustive explanation of experience (for an introduction, Seager, 2016). I would here suggest three main sets, in which the theories are collected on the basis of their ontological commitment towards consciousness: non-physicalist, physicalist and specific theories. The first one represents the most important philosophical positions about the nature of consciousness: dualisms (substance d., property d., fundamental property d., emergent property d.), monism (neutral m., anomalous m.) and panpsychism. The second set represents the physicalist positions, often – but not always – preferred in the scientific domain: eliminativism, identity theory, supervenience, functionalism, non-reductive physicalism. Lastly, the third set represents all those theories that do not perfectly fit within a specific metaphysical

⁹ «RtfMRI neurofeedback is a noninvasive technology that allows us to give an individual feedback on activation of single or multiple brain areas shown to be involved in specific functions. An added advantage to using rtfMRI is that we can provide feedback on either activation, connectivity, or both, allowing for a direct yet comprehensive approach to altering brain function. The goal of rtfMRI neurofeedback is to train patients to cognitively manage their own pain, but with the added aid of directly training neural systems underlying their successful modulation. This approach puts the patient in control, promoting a sense of self-efficacy and strengthening the patient’s mind-body relationship, thus giving rtfMRI the potential for producing long-lasting benefits» (Chapin, Bagarinao, & Mackey, 2012). For a systematic review see Thibault, MacPherson, Lifshitz, Roth, & Raz, 2018.

¹⁰ «Triangulation derives from the use of stars for navigation and can be defined in geometrical terms. It allows the navigator to determine the location of a point (specifically where the navigator is) by measuring angles. That is, you can specify your location on Earth by using a fixed baseline defined by two stars (by measuring the angles from your location to those stars). [...] The concept of triangulation has been used in cognitive science to indicate a combination of methods designed to intersect on a target subject matter. Owen Flanagan (1992), for example, suggested that to study consciousness one needed to combine psychology, neuroscience, and phenomenology - a combination reflected precisely in the neurophenomenological program as proposed by Francisco Varela (1996)» (Janz, Reinerman-Jones, Bockelman, Gallagher, & Trempler, 2016, p. 60). I shall come back to this point in the last chapters.

category. Van Gulick groups them in seven classes: higher-order theories – e.g. higher-order thought (HOT) and higher-order perception (HOP) – reflexive theories, representationalisms, narrative interpretative theories, cognitive theories, integrated information theories, some neural theories, quantum theories (Van Gulick, 2018).

Taken together, the huge variety of definitions, methodologies, perspectives and theories gives a multi-faceted, intricate and confounding picture of the contemporary studies of consciousness, and, to this day, it is not possible to talk about a solid and structured “science of consciousness”, from both conceptual and empirical domain.

1.2 The Hardness of Consciousness

As I will expose in detail later, despite the enormous progresses in contemporary science, there are still many doubts about whether consciousness concretely is suitable for empirical research. One of the main authors of the recent debate on consciousness is certainly David Chalmers. The problem of the nature of consciousness is a constant presence in his research, at least since Chalmers, 1987¹¹. In 1995, he officially raises the issue about the intrinsic dissatisfaction of any functional explanation of consciousness, with the groundbreaking paper *Facing Up to the Problem of Consciousness* (Chalmers, 1995, see also Chalmers 1996). The main point was that consciousness raises a hard problem that, by definition, is different from other easy problems of cognitive sciences: «The easy problems of consciousness are those that seem directly susceptible to the standard methods of cognitive science, whereby a phenomenon is explained in terms of computational or neural mechanisms. The hard problems are those that seem to resist those methods» (Chalmers, 1995, p. 200). This distinction between easy and hard problems is still matter of debate in different domains (philosophical, neuroscientific, cognitive, linguistic,

¹¹ See also Chalmers, 1990: “A completed theory of mind must provide (C1) An account of why we are conscious. (C2) An account of why we claim to be conscious; why we think we are conscious. Further (C3), accounts (C1) and (C2) must cohere with each other.”

phenomenological), and there is still no agreement about a possible solution or dissolution of the issue.

According to this distinction, most of problems raised by the phenomenon of consciousness are easy problems: the ability to discriminate environmental stimuli, integrate information, access our own mental states and report them, control our behaviour and make decisions, etc. Undoubtedly, they are “easy” only in a metaphorical sense, Chalmers was not arguing these issues can be solved with a simple solution, rather they are easy because they can be solved through functional or mechanist approaches. All the easy problems are important and necessary for a complete theory and science of consciousness, but the most important is the hard problem, which can be summarised in the fundamental question “[w]hy should physical processing give rise to a rich inner life at all?” (Chalmers, 1995, p. 201). All the solutions given through the functionalist or mechanist approach miss the point about the hard problem: the explanation of the performance of a particular function, e.g. the mechanisms that permit the conscious access to our own mental state, is not going to help the solution of the hard problem of consciousness in any way. There will be still there the unsolved question: why is the performance of this mechanism accompanied by experience? How is this internal access experienced? These questions go «beyond problems about the performance of functions» (Chalmers, 1995, p. 203).

The hardness of consciousness seems something specifically related to consciousness; indeed, it would be quite odd to ask the same kind of questions for other natural phenomena. Chalmers takes the example of DNA: «If someone says ‘I can see that you have explained how DNA stores and transmits hereditary information from one generation to the next, but you have not explained how it is a *gene*,’ then they are making a conceptual mistake» (Chalmers, 1995, p. 203). The same goes with other complex explanations, like the physics of black matter or the concept of life. As I will expose later, many problems about natural phenomena and/or concepts are extremely “complex”, instead consciousness refers to some specific “hard” character of existence. There are also other problems of consciousness related to the hard one. Clark’s “meta-hard problem of consciousness”, i.e. the problem whether there really is HP, «something special about phenomenal consciousness [...] or is it just a matter of explaining a pattern of responsiveness and report» (Clark, 2001, p. 185), Aaronson’s

“pretty hard problem”, i.e. how scientific theories can match our intuitions about which systems are conscious (Aaronson, 2014), the harder problem, i.e. an epistemological problem closely tied to the problem of other minds¹² (Block, 2002) and the meta-problem of consciousness (below for details)¹³.

Certainly, this problem of consciousness in relation to objective reality and physical science existed, in similar versions but with different facets, well before¹⁴ Chalmers’ definition of the hard problem. For example, Nagel’s “what it is like to be” in a specific conscious state: «fundamentally an organism has conscious mental states if and only if there is something that it is like to be that organism – something it is like for the organism» (Nagel, 1974, p. 436). With this definition, Thomas Nagel meant the subjective character of experience that is not analysable in terms of any explanatory system of functional states. Indeed, any objective, physical theory, at least in principle, must abandon all the single points of view on the world, but the “what it is like to be” is essentially connected with a single point of view and existence¹⁵, therefore it seems reasonable to say that these explanations inevitably leave something out. And this is the reason why any functional explanation cannot answer to the question “why is the performance of this mechanism accompanied by experience?”. Also, more specifically Nagel’s “what it is like” can be divided between “qualitative” and “subjective” characters or dimensions: the subjective character is the fact that there is something it is like “for me” to have the experience of a red tomato (and hence not “for the tomato”

¹² «Are the Hard and Harder Problems really different problems? The Hard Problem is: why is the scientific basis of a phenomenal property the scientific basis of that property rather than another or rather than a nonphenomenal property? The question behind the Harder Problem could be put so as to emphasize the similarity: why should physically different creatures overlap phenomenally in one way rather than another or not at all? This way of putting it makes it plausible that the Harder Problem includes or presupposes the Hard Problem. In any case, the Harder Problem includes an issue that is more narrowly epistemic than the Hard Problem the Hard Problem could arise for someone who has no conception of another person, whereas the Harder Problem is closely tied to the problem of other minds. Finally, the Harder Problem involves an epistemic discomfort not involved in the Hard Problem. My claim is that the “Harder Problem” differs from the “Hard Problem” in these ways independently of whether we choose to see them as distinct problems or as part of a single problem» (Block, 2002, pp. 423–424).

¹³ There also is an antagonist problem, the so-called “real problem” of consciousness: «[H]ow to account for the various properties of consciousness in terms of biological mechanisms; without pretending it doesn’t exist (easy problem) and without worrying too much about explaining its existence in the first place (hard problem)» (Seth, 2016).

¹⁴ For example, Paolo Pecere exposes all the similarities between Chalmers’ hard problem and du Bois-Reymond’s thesis (1872) on the impossibility of a scientific explanation of consciousness (Pecere, 2018).

¹⁵ «It does not mean “what (in our experience) it resembles,” but rather “how it is for the subject himself.”» (Nagel, 1974, p. 440n).

per se); the qualitative character, instead, refers to the “what” of that experience, that is, the reddish of the tomato¹⁶ (Kriegel, 2009; Levine, 2001). To sum up Nagel’s argument:

- 1) «Fundamentally, an organism has conscious mental states if and only if there is something that it is like to be that organism – something it is like for the organism» (p. 166).
- 2) «If physicalism is to be defended, the phenomenological features of experience must themselves be given a physical account» (p. 167).
- 3) «But when we examine their subjective character it seems that such a result is impossible. The reason is that every subjective phenomenon is essentially connected with a single point of view, and it seems inevitable that an objective physical theory will abandon that point of view» (p. 167).

Chalmers’ hard problem is related to Nagel’s “what it is like” (see Chalmers, 1995, p. 214), nevertheless the two issues are not identical. Nagel approaches consciousness from an epistemic perspective: what interests Nagel particularly is the subjective character of experience beyond the reach of the objective science through the alleged “point of view from nowhere” (T. Nagel, 1986). Every single person has his/her personal perspective and at the same time, as human being, he/she can consider the world in a detached way. The point, according to Nagel, is how to reconcile and integrate these two epistemic dimensions. A reductive science that aims to objectify everything will never be able to comprehend also the single personal standpoints. Conversely, the personal perspective is not exclusively private and impenetrable, because we reciprocally share similar biological structure and we all are embedded in the same semantic, linguistic and social context¹⁷. Even with the example of the bat

¹⁶ «To notice the difference between these two dimensions is not to say that they are separable in reality; what is intended, rather, is a conceptual distinction between simultaneously present and interdependent facets of phenomenal consciousness» (Guillot, 2017).

¹⁷ «The subjective character of the experience of a person deaf and blind from birth is not accessible to me, for example, nor presumably is mine to him. This does not prevent us each from believing that the other’s experience has such a subjective character. [...] I am not adverting here to the alleged privacy of experience to its possessor. The point of view in question is not one accessible only to a single individual. Rather it is a type. It is often possible to take up a point of view other than one’s own, so the comprehension of such facts is not limited to one’s own case. There is a sense in which phenomenological facts are perfectly objective: one person can know or say of another what the quality of the other’s experience is. They are subjective, however, in the sense that even this objective ascription of experience is possible only for someone sufficiently similar to the object of ascription to be able to adopt his point of view-to understand the ascription in the first person as well as in the third, so to speak» (Nagel, 1974, p. 440, 441-2).

that gives the title to the famous paper in 1974, Nagel is not arguing that we cannot know what it is like to be a bat, at least in principle (Nagel, 1974, p. 442n), indeed we already have extensive scientific knowledge about how echolocation works in bats and other animals (M. B. Fenton et al., 2016; Surlykke et al., 2014). Rather he wants to highlight the fact that in order to form even a simple conception of what it is like to be a bat, we must take up the bat's point of view, that is, the bat's single point of view. And this is far from the domain of «objective facts *par excellence* – that kind that can be observed and understood from many points of view and by individuals with differing perceptual systems» (Nagel, 1974, p. 442). This is an epistemic argument, as Nagel is not strictly accounting for the existence of something super- or extra-natural, but for the undeniable peculiarity of the subjective point of view. And this is related with the issue of consciousness, because experience seems specifically tied to our personal standpoint, differently from other natural phenomena that progressively needs a more and more fine-grained objective description¹⁸. This is what I will describe later in detail as the “specificity” of consciousness.

For the reasons stated here above, recent attempts to solve the “what it is like to be” issue through a neuroscientific theory are definitively missing the point. Naotsugu Tsuchiya argues that the application of the Integrative information theory (IIT, for details see chapter three) on Nagel's problem will be in the future a possible, or even the only, solution to it (Tsuchiya, 2017). Through comparisons between different integrated information patterns in human and bat brains, we could predict in the next future whether bats' echolocation should feel more like our vision, audition or neither of them. This is completely in accordance with Nagel's claims¹⁹, still it does not deal with the fact that any potential objective comparison will never reach what it

¹⁸ «And in understanding a phenomenon like lightning, it is legitimate to go as far away as one can from a strictly human viewpoint. In the case of experience, on the other hand, the connection with a particular point of view seems much closer. It is difficult to understand what could be meant by the objective character of an experience, apart from the particular point of view from which its subject apprehends it. After all, what would be left of what it was like to be a bat if one removed the viewpoint of the bat?» (Nagel, 1974, p. 443).

¹⁹ «It may be possible to approach the gap between subjective and objective from another direction. [...] This should be regarded as a challenge to form new concepts and devise a new method an objective phenomenology not dependent on empathy or the imagination. Though presumably it would not capture everything, its goal would be to describe, at least in part, the subjective character of experiences in a form comprehensible to beings incapable of having those experiences. We would have to develop such a phenomenology to describe the sonar experiences of bats» (emphasis added) (Nagel, 1974, p. 449).

is like to have the echolocation for a bat, i.e. its phenomenal world²⁰. Rather, it is more likely that some blind people that click with the tongue as a way of understanding where they are in space are closing to experiencing, from their personal standpoint, what it is like to be a bat (Schörnich et al., 2013; Thaler et al., 2017)²¹. Nevertheless, this is a way of approaching another point of view simply by analogy, the blind person still has his/her own human mental schemes (language, concepts, etc.) and consequently a phenomenal world that is probably different from the bat's one. Interestingly enough, the so-called “new mysterianism” (category coined by Flanagan, 1991, pp. 312–314, within which he includes Nagel, 1974, 1986 and McGinn, 1991) claims that consciousness exists but it will never be understood, not because of some dualistic theory of mind and body (the “old mysterianism”), but because there is no plausible naturalistic account that is able to explain how a single conscious event can be both phenomenological and physical. This impossibility would not be due to some extra-natural principle, but to our cognitive limits towards the relation between the natural substrate of consciousness and consciousness itself (McGinn, 1991). Strawson, in response to this point, says that the “mystery argument” is a “Very Large Mistake” because we think we know enough about what physical stuff is to have good reason to think that physical goings-on in the brain can't be conscious goings-on, instead there is an “astonishing chaos” about the meaning of the word “physical” (Strawson, forthcoming). I agree about that, but I would rephrase the claim saying that not only we presuppose to know enough about what physical stuff is, but also to know enough about the brain for arguing that the brain would be enough for generating and/or explaining consciousness. Even thinking that physical goings-on in the brain can potentially be conscious goings-on is a presupposition.

About the concept of “phenomenal world”, we need to distinguish the features of consciousness not only as opposed to other natural phenomena, which do not present indeed such epistemic and ontological characters, but it is also important to identify different components of consciousness. Interestingly, Ned Block proposed the

²⁰ A Martian scientist with no understanding of visual perception could understand the rainbow, or lightning, or clouds as physical phenomena, though he would never be able to understand the human concepts of rainbow, lightning, or cloud, or the place these things occupy in our phenomenal world.

²¹ In the concluding remarks of his paper, Tsuchiya admits that «it may be practically impossible to understand bats' phenomenology in every detail» (Tsuchiya, 2017, p. 12), but that is exactly one of Nagel's main points.

distinction between phenomenal and access consciousness (Block, 1995). Phenomenal consciousness (P-consciousness) is the actual experience as such, what Nagel describes as what it is like to be in a particular state for a particular subject²², and it must be distinct from any cognitive or functional or intentional notion, at least conceptually²³. But most importantly, P-consciousness is different from access consciousness (A-consciousness), which happens when some mental content is processed and made available to reasoning, speech, report, behaviour and reflection. Moreover, access-consciousness «can be understood in two different ways, (1) dispositionally, as a matter of availability to cognitive processes, or, (2) occurrently, in terms of representations reaching the machinery of thought and reasoning» (Block, forthcoming). During our normal daily life, P- and A-consciousness are constantly and reciprocally interacting with each other, however Block's point is to show that there can be – conceptually and concretely – cases of P-consciousness without A-consciousness²⁴ and vice versa²⁵. Given that A-consciousness refers to the functional

²² «P-consciousness is experience. P-consciousness properties are experiential ones. P-conscious states are experiential, that is, a state is P-conscious if it has experiential properties. The totality of the experiential properties of a state are "what it is like" to have it. Moving from synonyms to examples, we have P-conscious states when we see, hear, smell, taste, and have pains. P-conscious properties include the experiential properties of sensations, feelings, and perceptions, but I would also include thoughts, desires, and emotions» (Block, 1995, p. 230).

²³ «Cognitive = essentially involving thought; intentional properties = properties in virtue of which a representation or state is about something; functional properties = properties definable (for example) in terms of a computer program» (Block, 1995, p. 230).

²⁴ «Consider an animal you are happy to think of as having P-consciousness, for which brain damage has destroyed centers of reasoning and rational control of action, thus preventing A-consciousness. It certainly seems conceptually possible that the neural bases of P-consciousness systems and A-consciousness systems are distinct. [...] Here is another reason to believe in P-consciousness without A-consciousness: suppose you are engaged in intense conversation when suddenly at noon you realize that right outside your window there is- and has been for some time - a deafening pneumatic drill digging up the street. You were aware of the noise all along, but only at noon are you consciously aware of it. That is, you were P-conscious of the noise all along, but at noon you are both P-conscious and A-conscious of it» (Block, 1995, p. 233-234).

²⁵ Regarding A-consciousness without P-consciousness, Block only gives conceptual examples: «[...] imagine a full-fledged phenomenal zombie, say, a robot computationally identical to a person, but one whose silicon brain does not support P-consciousness. I think such cases are conceptually possible, but this is very controversial [...].

Now imagine something that may not exist, what we might call superblindsight. [...] But suppose - contrary to fact apparently - that a blindsight patient could be trained to prompt himself at will, guessing what is in the blind field without being told to guess. The superblindsighter spontaneously says, "Now I know there is a horizontal line in my blind field even though I don't actually see it." Visual information from his blind field simply pops into his thoughts in the way that solutions to problems we've been worrying about pop into our thoughts, or in the way some people just know the time or which way is North without having any perceptual experience of it. The superblindsighter himself contrasts what it is like to know visually about an X in his blind field and an X in his sighted field. There is something it is like to experience the latter, but not the former, he says. It is the difference between just knowing and knowing via a visual experience. Let us take his word for it; here is the point: the content that there is

aspects of the mind, and that P-consciousness instead refers to Nagel's "what it is like", only the latter is related to the hard problem, while the ability of use perceptual and cognitive information for behaviour, reflection and speech is part of the easy problems²⁶. Therefore, those who claim that it is possible to give a functional explanation of consciousness is confounding A-consciousness with P-consciousness.

Looking back to the "what it is like", I explained the epistemic character of Nagel's argument, which is different from Chalmers' hard problem that primarily elicits an ontological question. It is true that one of the ways in which we can formulate the hard problem is "how this particular functional/objective/physical theory of consciousness does not explain experience as such", however the main question is exquisitely ontological: "why and how must some particular functional/objective/physical system be accompanied by experience, i.e. by a supposedly non-physical feature?"²⁷. Both version of the problem can be described better if we consider another philosophical argument that came before Chalmers' one and presents interesting fluctuations between pure epistemic and strong ontological claims, i.e. the knowledge argument. The knowledge argument is a thought experiment that Frank Jackson has formulated in 1982, revisiting it in 1986 and thereafter, in order to reply to a great debate created around this issue (Alter, 2007; Nida-Rümelin, 2015). Here I quote the original version of the argument (Jackson, 1982, p. 130):

«Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room via a black and white television monitor. She specialises in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on. She discovers, for example, just which wave-length combinations from the sky stimulate the retina, and exactly how this produces via the central nervous system the

an X in his visual field is A-conscious but not P-conscious. The superblindsight case is a very limited partial zombie» (Block, 1995, p. 233).

²⁶ «It is of course P-consciousness rather than access consciousness or self-consciousness that has seemed such a scientific mystery. [...] I mentioned the explanatory gap partly by way of pointing at P-consciousness: that is the entity to which the mentioned explanatory gap applies. Perhaps this identification is contingent; at some time in the future, when we have the concepts to conceive of much more about the explanation of P-consciousness, there may be no explanatory gap to use in picking out P-consciousness» (Block, 1995, pp. 230–231).

²⁷ See also (Chalmers, 1996, p. xi) «it still seems utterly mysterious that the causation of behaviour should be accompanied by a subjective inner life».

contraction of the vocal chords and expulsion of air from the lungs that results in the uttering of the sentence 'The sky is blue'. (It can hardly be denied that it is in principle possible to obtain all this physical information from black and white television, otherwise the Open University would of necessity need to use colour television.) What will happen when Mary is released from her black and white room or is given a colour television monitor? Will she learn anything or not? It seems just obvious that she will learn something about the world and our visual experience of it. But then it is inescapable that her previous knowledge was incomplete. But she had all the physical information. Ergo there is more to have than that, and Physicalism is false».

Basically, the argument can be summarized in some passages (Nida-Rümelin, 2015):

- 1) Mary has all the physical information concerning human colour vision before her release.
- 2) But there is some information about human colour vision that she does not have before her release.

Therefore

- 3) Not all information is physical information.

The argument claims that the mere physical information cannot explain all the phenomena in the world, and not for any contingent reason. When we say that Mary has all physical information, it should be intended as all the possible physical knowledge about the physical world, i.e. the past, present and future knowledge that a science grounded on physical causal closure and naturalism could potentially give to us. Surely, Jackson's argument is not the first against the idea of completeness of physical knowledge, however in the last decades the debate about the possibility or the impossibility to account for phenomenal states within the approach of natural sciences was very intense and heated. For this reason, the knowledge argument is so important, even today, where new neuroscientific discoveries about the neural mechanism of our consciousness and cognition are a daily occurrence.

There are several ways in which we can interpret the meaning of the argument. The original intent had a genuine ontological valence, but in its development the argument could change its purpose, depending on the ontological commitment we want to sustain²⁸. Assuming for argument's sake that the claim is consistent and coherent, it has to be discussed the meaning of "non-physical". Hence, the argument could be mainly interpreted in two ways: the epistemological (weaker version) and the ontological (stronger version) (Horgan, 1989)²⁹. The former sustains that there is a kind of knowledge which is not concerning physical knowledge, while the latter that there are some non-physical facts *tout court*. The weaker version is similar to the explanatory gap argument from Joseph Levine, also cited by Chalmers for the hard problem of consciousness (Chalmers, 1995, p. 203, 2007). The explanatory gap argument states the inability to bridging the epistemic gap between all the functions described scientifically with an objective approach, e.g. physics, neuroscience, biology, etc., and the lived experience itself, e.g. the world perceived from our subjective point of view: no objective explanation can ever account for the experience itself in an exhaustive way (Levine, 1983)³⁰.

1.2.1 The Nature of Qualia

So far, we have been looking to all those features that show us how consciousness could be considered a hard phenomenon to account for. As I shall expose later in detail, even though we can imagine finding a scientific solution, in the future, to some contemporary complex problems of physics, biology and other so-called "hard

²⁸ «I think the argument is decisive, though it is frequently misunderstood in ways that treat it as merely epistemic and not ontological. It is sometimes treated as an epistemic argument to the effect that, for example, the sort of third-person, objective knowledge we might possibly have of a bat's neurophysiology would still not include the first-person, subjective experience of what it feels like to be a bat. But for our present purposes, the point of the argument is ontological and not epistemic. It is a point about what real features exist in the world and not, except derivatively, about how we know about those features» (Searle, 1992, p. 117).

²⁹ There are also more detailed versions of the argument, which try to develop all the shades and intermingling of epistemology and ontology, such as the distinction between metaphysical and linguistic physicalist, the ability hypothesis, the acquaintance hypothesis, and the non-reductive physicalist position of the new knowledge/old fact view (for a general introduction, Nida-Rümelin, 2015).

³⁰ «The explanatory gap argument doesn't demonstrate a gap in nature, but a gap in our understanding of nature. Of course a plausible explanation for there being a gap in our understanding of nature is that there is a genuine gap in nature. But so long as we have countervailing reasons for doubting the latter, we have to look elsewhere for an explanation of the former» (Levine, 1999, p. 12).

sciences”, nevertheless it seems to be a different matter with consciousness. However, the hardness of the problem of consciousness does not automatically lead to the impossibility of a solution, but in order to imagine a possible solution, we first need to conceive a new and different approach to consciousness, in particular the subjective character of it. The subjective character appears to be an unavoidable element for any theory of consciousness, of course this character could be reduced, operationalized, measured in some way, but it does not seem appropriate to discard it beforehand from the explanation (I shall present some of these attempts in the third chapter). Indeed, the subjective character of experience is considered the most essential element of consciousness, the features that differentiate consciousness from other phenomena within the scientific domain.

So far, I have exposed different ways in which the importance of the subjective character has been argued and defended (e.g. “what it is like”, phenomenal consciousness, explanatory gap, knowledge argument and, of course, the hard problem) and I have tried to describe their similarities but also their differences, in particular from ontological and epistemological approaches. Nevertheless, there is a philosophical problem under which we can include all these other issues as different facets of this one, that is, the so-called problem of “qualia” (singular form: quale). Qualia are instances that determine the qualitative aspect of any individual experience, i.e. the matter of what it is like for the subject of that experience to have that experience. Generally, qualia refer to other concepts like “raw feels”, “phenomenal properties”, “subjective and intrinsic properties” or more generally the “qualitative character” of experience. When we see a colourful flower, we smell the fragrance of ground coffee, we taste the roundness of a precious wine or we touch the rough bark of a tree, of course there are specific neural activities that receive the stimuli and process the information in order to correctly recognize those colour, smell, taste and roughness, also in a way in which we could not confuse these different stimuli with each other³¹. However, all these physical, functional, mechanical processes have little

³¹ This, for example, does not always happen in people with synaesthesia, i.e. «a remarkable way of perceiving the world. One attribute of a stimulus (e.g., its sound, shape, or meaning) may inevitably lead to the conscious experience of an additional attribute. For example, the word “Phillip” may taste of sour oranges, the letter A may be luminous red, and a C# note on the violin may be a brown fuzzy line extending from left to right in the lower left part of space» (J. Ward, 2013). This perceptual phenomenon presents particular cognitive and neural basis (Hubbard & Ramachandran, 2005; J. Ward, 2013).

to do with the “raw feels”, the basic aspects of those experiences without which there would not be any seeing, smelling, tasting, touching, etc. Even conceptual analysis or qualitative, which is always useful for a better determination of a particular mental content, has some difficulties in breaking into this immediate, pre-cognitive, raw and direct aspect of experience.

Daniel Dennett is one of the fiercest critics of qualia – I shall expose his position in the third chapter – but he offered a detailed description of how qualia are generally understood by contemporary philosophical and scientific communities (Dennett, 1988, 1991, 2005). He defined qualia as:

- a. intrinsic, i.e. belonging essentially to consciousness, we cannot have a proper experience without the qualitative aspect;
- b. ineffable, i.e. simple and homogeneous, hard to split, analyse and describe in some way to one unacquainted with that particular qualitative instance.
- c. private, i.e. qualia are properties of “my” experience, and they can be reached only by the first-person perspective of a particular subject.
- d. directly apprehensible properties, i.e. they are non-propositional and non-cognitive properties that can be immediately acquainted by the subject³².

Other – even opposed – features that are sometimes used for defining qualia are non-physicality, incorrigibility, non-intentionality, relationality, etc. (Tye, 2018). As is evident from this general description, the problem of qualia is extremely related to the problem of consciousness, assuming that our aim is to find an explanation of consciousness from both the objective/quantitative and the subjective/qualitative side. Moreover, it seems that any epistemological or ontological issue discussed earlier can be included in and represented by the problem of qualia. For example, it addresses the questions about what it is like to have a raw feel (Nagel), why scientific explanation

³² «To understand the problem, let’s analyze more closely what “feel” really is. Take the example of the color red. One aspect of the feel of red is the cognitive states that red puts me into. There are mental associations such as, among other things, roses, ketchup, blood, red traffic lights, anger, and translucent red cough drops. Additional mental associations may be determined by knowledge I have about how red is related to other colors, for example, that it is quite similar to pink and quite different from green. There are other cognitive states that are connected with having an experience of red, namely the thoughts and linguistic utterances that seeing redness at this moment might be provoking in me, or changes in my knowledge, my plans, opinions, or desires. But all these cognitive states are extra things in addition to the raw sensation. They are not themselves experiences. They are mental connections or “add-ons” to what happens when I have the experience of red» (O’Regan, 2011, p. 95).

seems to leave out the qualitative aspects of experience (Jackson and Levine), and why some particular brain function must be accompanied by qualitative experience (Chalmers)³³. In other words, if consciousness is the Holy Grail of the science of mind, the problem of qualia is the sacred brew in it. Notably, for someone there is no difference at all between the problem of consciousness and the concept of qualia, for example John Searle: «The problem of consciousness is identical with the problem of qualia, because conscious states are qualitative states right down to the ground. Take away the qualia and there is nothing there. This is why that I seldom use the word “qualia”, except in sneer quotes, because it suggests that there is something else to consciousness besides qualia, and there isn’t. Conscious states by definition are inner, qualitative, subjective states of awareness or sentience» (Searle, 1998, p. 21).

As I will describe in detail in the third chapter, there are many philosophical attempts to account for the nature of qualia, for example through reductionist and non-reductionist versions of intentionalism or representationalism (Langsam, 2018). Indeed, the nature of qualia is still highly debated, with positions that go from a total agreement about their ontology to a complete rejection of their existence. We could categorize all the positions in, at least, four groups. The first one claims that qualia exist and do not supervene on the natural world³⁴, arguing for the insolubility of the

³³ Consider the instance made by O’Regan: «For the sake of argument, suppose it was two different populations of neurons that coded the raw feel of red and the raw feel of green. What could it be about these different groups of neurons that generated the different feels? Their connections to other brain areas could certainly determine the extracomponents of the feel, for example, any immediate bodily reactions or tendencies to act or changes in cognitive state that are produced by feels of red and green. But what precise aspect of these neurons might be generating the differences in the actual raw feel itself? *There can be no way of answering this question satisfactorily.* For imagine we had actually hit upon what we thought was an explanation. Suppose that it was, say, the fact that the red-producing group of neurons generated one particular frequency of oscillations in wide-ranging cortical areas, and that the green-producing neurons generated a different frequency. Then we could ask: Why does this frequency of oscillation produce the raw feel of red, and that frequency the raw feel of green? Couldn’t it be the other way round? We could go in and look at what was special about these frequencies. Perhaps we would find that the red producing frequencies favored metabolism of a certain neurotransmitter, whereas the green producing frequencies favored metabolism of a different neurotransmitter. But then we would be back to having to explain why these particular neurotransmitters generated the particular raw feels that they do. (p.98) We would not have any logical difficulty if we were simply interested in understanding how the neurotransmitters determined the extra components associated with red and green by influencing brain states and potential behaviors and bodily states. But how could we explain the way they determined the raw feel *itself*? [...] Finding brain structures that are necessary and sufficient to generate a particular raw feel would no doubt be very interesting, but it would still not explain what we would really like to know about the feel, namely why it is the way it is» (O’Regan, 2011, pp. 97–98, italics in original).

³⁴ «A set of properties A supervenes upon another set B just in case no two things can differ with respect to A-properties without also differing with respect to their B-properties. In slogan form, “there cannot be an A-difference without a B-difference”» (McLaughlin & Bennett, 2018).

problem of consciousness from a physicalist perspective (Chalmers, 1996; Jackson, 1986). The second group, instead, claims that qualia supervene on the natural and can be related to psycho-physical laws as an emergent property (Searle, 1992). The third group argues for the reducibility of qualia to intentional properties or representational states (Lycan, 2019). The fourth group claims that qualia do not exist at all, as they are illusions (Williams, 2019). Moreover, it would be useful to analyse the historical and philosophical roots of this concept, in order to highlight its ontological and epistemological characteristics and maybe offer an alternative interpretation (see second chapter).

Taken together, it seems hard to deal with the phenomenon of consciousness – the problem of qualia and the phenomenal character of experience in general – through both the empirical way (consciousness seems to exceed traditional paradigms of cognitive science) and the conceptual way (we are far from finding a univocal and shared definition of consciousness). Hence, the “hardness” of consciousness, that is, the fact that consciousness appears not only as a complex phenomenon to examine, but also resistant to any attempt of explanation, and sometimes it may look as if it was beyond the reach of our own cognitive abilities. Nevertheless, it is important to note that Chalmers has not introduced the hard problem as unsolvable by principle:

«Most existing theories of consciousness either deny the phenomenon, explain something else, or elevate the problem to an eternal mystery. I hope to have shown that it is possible to make progress on the problem even while taking it seriously. To make further progress, we will need further investigation, more refined theories, and more careful analysis. The hard problem is a hard problem, but there is no reason to believe that it will remain permanently unsolved» (Chalmers, 1995, pp. 217–218).

Before any attempt to give any possible solution or dissolution to the hard problem of consciousness, I believe that it could be important to analyse the reasons why we have a hard problem of consciousness. Indeed, we cannot simply assume that there seems to be a hard problem as a fact, but we need to focus on the grounds on which we believe that consciousness raises such a difficult issue (and, for example, why not other natural phenomena). This may be helpful for defining better the roots of the problem as such,

and then we may suggest a way for approaching it. Now to accomplish this, I shall go through a more recent paper from Chalmers, *The Meta-Problem of Consciousness*.

1.3 The Meta-problem of Consciousness

In this paper, Chalmers comes back to the crucial question about the ontological nature of consciousness and its epistemological complexity (Chalmers, 2018). Basically, the meta-problem addresses the question of consciousness from a new perspective and it could be interpreted as grounding or consequent in respect of the original hard problem of consciousness. While the hard problem (HP) raises the question about the relation between the physical and the “what it is like to be” of any conscious experience, the meta-problem (MP) refers to the “problem reports”, i.e. why we feel that consciousness is something special or different from other phenomena. Differently from HP, MP momentarily suspends the judgment on the possibility or impossibility of an exhaustive explanation of conscious experience, rather it focuses on the reason why consciousness appears as something that deserves a particular ontological (it seems something non-physical) and epistemological (it seems irreducible to functional explanation) investigation. Chalmers exposes many potential intuitions supporting that consciousness is considered special somehow – explanatory, metaphysical, knowledge, modal, value, distribution, etc. Analysing them one by one and defining the most plausible is not the aim of this work. Nevertheless, for our purposes I shall discuss two main points about the relation between MP and HP:

- 1) What does the prefix “meta” mean in the meta-problem of consciousness? In other words, to what extent MP and HP seem intertwined, either from the illusionist or the realist perspective?
- 2) Which are the conditions of possibility for consciousness to be perceived, differently from other complex phenomena, as something with special features (qualities, properties, knowledge, perspective, etc.)? In other words, why does consciousness seem to exhibit a sort of epistemological, ontological, etc. specificity? And why not other complex phenomena/concepts?

1.3.1 What does “Meta” mean?

The first issue is related to the prefix “meta” used in the meta-problem. MP and HP seem reciprocally related, hence the presence of the suffix. Regarding this relation, Chalmers distinguishes between a strong line (the solution of MP will dissolve the HP) and a weaker line (the solution of MP will give a direction towards the solution of HP). He defines MP as an easy problem that nevertheless is relevant for the solution of the hard one, explicitly saying that the two are somehow intertwined. However, he never explains how this relation is structured, leaving the burden of proof with the realists, i.e. those who claim that consciousness is a real phenomenon. In order to understand the relation with the meta-problem and the hard problem, we first need to analyse the implication of that prefix in MP. Chalmers uses the suffix as “about” (Chalmers, 2018, p. 6), i.e. MP is the problem about the problem of consciousness, however he does not account for a detailed explanation of how MP is “deeply connected” with HP (p.8). “Meta” suggests a plethora of meanings (from the Greek μετά-). For example, “metacognition” (cognition about cognition), “metadata” (data with information about other data), “metaphysics” (beyond, after the natural world, φύσις), “metaethics” (the study on the foundation of moral values and words). Epistemologically speaking, the meaning of “meta” is vague and can be interpreted in two main ways: MP is the grounding (HP depends on MP) or the consequent/collateral problem (MP is a collateral product of HP). The issue becomes more puzzling if we think that HP is a hard problem by definition, while MP is an easy one, since the problem reports basically are expressions of human behaviour, therefore they could potentially be explained in physical and functional terms.

Suppose MP is the grounding problem of HP, this would mean that the solution of it would lead to the dis-solution of HP. I say “dis-solution” because, if MP is the solution of HP, then HP cannot be really a hard problem, rather it must be an illusory problem. Once we found the solution to the (easy) meta-problem of consciousness, the hard problem itself will not be simply understood as an easy problem but it will be revealed as inconsistent and aporetic. Hence, the strong line of argument exposed above: the meta-problem will be helpful to highlight the structure of biased beliefs upon which HP lies. Indeed, according to some philosophical positions like the eliminative materialism, the very idea of a hard problem is to be challenged (see

Blackmore & Troscianko, 2018). Someone argues that it is useless to address the hard problem without considering first all the easy problems, which are still important for the definition of consciousness and they probably will change our view on the hard problem itself (O'Hara & Scutt, 1996). Peter Strawson labels the hard problem as the “so-called problem of consciousness”, because we know for certain that consciousness exists and we do not need any other proof of it, instead we are confusing the non-existent problem of consciousness with the hard problem of matter³⁵. Others, like David Papineau, argue that there is no reason why we should not be able to refer to the “what it is like” using concepts which do not actually give us the feeling, as when we do not expect that the physical explanation of how a hurricane “works” will give us also the feeling of being inside a hurricane. Patricia Churchland claims that the hard problem is “ridiculous”, a “hornswoggle problem”, a grand hoax, because we do not have any idea how to identify which problems are hard and which are the easy ones, and also because the hard problem of consciousness will have the same fate as phlogiston, ether or caloric fluid (P. S. Churchland, 1996)³⁶. Lastly, as we will see in the third chapter, Dennett argues that the hard problem is an illusion, indeed when someone asks an explanation for the “actual” phenomenology (the what it is like or qualia) Dennett answers that there is no such a thing at all, we only think to have an actual phenomenology of consciousness (Dennett, 1991, 2005). This last argument is what Chalmers calls the “illusionist position”, from Keith Frankish’s illusionism (Frankish, 2016). Illusionists argue that consciousness is not a real phenomenon but is like an illusion created by the brain for different reasons, e.g. evolutionary, cognitive, pragmatic, etc. (see Chalmers, 2018, pp. 8–9). According to this position, if we have a physical explanation of why we seem to have non-physical properties, those “special”

³⁵ The hard problem is the problem of matter. We know a great many facts about the structure of matter, facts which physics expresses with numbers and equations: $E = mc^2$, the inverse-square law of gravitational attraction, the periodic table and so on. But we have no idea about the intrinsic nature of the stuff that fleshes out this logico-mathematically expressible structure – except, of course, in so far as it is or involves consciousness, which we may do well to think of as a kind of stuff, a quantity. (There is an enormous quantity of conscious experience on this planet, whatever is going on elsewhere in the universe.)» (Strawson, 2017, pp. 97–98).

³⁶ Similarly, Dehaene argues that the hard problem exists simply because it is ill-defined: «Chalmers swapped the labels: it is the “easy” problem that is hard, while the “hard” problem just seems hard because it engages ill-defined intuitions. Once our intuition is educated by cognitive neuroscience and computer simulations, Chalmers’ “hard problem” will evaporate» (Dehaene, 2014, p. 262).

properties will simply be revealed as illusory³⁷. It is just like a magic show: the performer does not do “real magic”, but only a sophisticated trick, and once all the parts of the trick will be revealed (the easy problems), the astonishment (the hard problem) will disappear (Dennett, 2005 chapter 3). As I will argue later, according to this metaphor, consciousness should be a very strange and problematic trick, because even after the eventual solution of it, we would keep seeing it as something magical and surprising.

Suppose now MP is the consequent/collateral problem of HP, this would mean that the original problem remains HP. In this case, there would be a collateral/consequent problem about the epistemological constitution of HP. Chalmers says that the solution of it would shed light on HP. However, if MP is an easy problem – it can be solved through functional or mechanist approaches – then it is not clear how it could help in defining a solution for HP, as Chalmers himself argued that easy problems concern functions and abilities, while hard problems resist those explanations (Chalmers, 1995, p. 202). Indeed, explanations of easy problems cannot help in defining why these very same functions and abilities must be accompanied by experience. Notably, Chalmers claims that the solution of an easy problem like MP «will constrain the form of a solution» for HP (Chalmers, 2018, p. 8). Following this line, one might consequently argue that the solution of several easy problems strictly related to consciousness – e.g. the neural correlates of consciousness – will progressively shape the form of a solution of HP. It is a possible answer, but then the constitution of HP as resistant to functional explanation might fall, and consciousness might be considered as a complex phenomenon, yet entirely reducible within the domain of empirical sciences and functional paradigms.

Taking together, the two interpretations of the “meta” suffix could lead to the illusionist position or to the dismissal of the hardness of the problem of consciousness. There is an alternative interpretation of the weaker line though, regarding how MP can “shed light” on HP (Chalmers, 2018, p. 35): MP and HP are epistemologically related, however the eventual solution of MP will not directly help in finding the solution of

³⁷ «For the illusionist, what needs to be explained is not phenomenality or qualia or ‘the experiences themselves’ but our illusory ideas about experience. This means we can avoid the hard problem altogether and replace it with ‘the illusion problem’» (Frankish, 2016, p. 11).

HP, because the two problems are ontologically incommensurable. For argument's sake, let us assume that our problem reports about consciousness are evolutionary products, e.g. a useful belief that helps complex biological systems to preserve their own life and the life of their kind, a sort of "evolutionary intuition", to put it in Chalmers' terms. Now, I do not see how this should help in finding how and why something like a conscious experience would rise from natural selection and physical processes: once found this evolutionary reason about MP, there is no guarantee that HP will be nearer to a solution about the nature of consciousness.

I think that the best way to deal with this ontological incommensurability is to consider MP as regulative while HP as constitutive regarding consciousness³⁸. The aim of HP is consciousness itself, therefore it constitutes the subject in question as such (regardless of which solution it offers) and it must deal with its ontology. Instead, MP has a regulative function, it refers to the conditions of the problem reports, offering heuristic and contextual solutions about them. These solutions can be counted as epistemic edges for HP, i.e. the assessment of the epistemological domain within which it is more likely to found substantial hints for HP. However, MP offers no direct solutions about the subject of HP, as consciousness as such is out of the epistemic reach of MP. Chalmers also claims that this deep connection between MP and HP is bidirectional (Chalmers, 2018, p. 35). The idea about the meaning of "meta" exposed here above refers only to the relation that goes from MP to HP. However, it is likely that also the eventual solution for the original HP will help for the solution of MP, not only in a regulative way but also in a constitutive one.

1.3.2 The Specificity of Consciousness

The second issue refers to what I call the specificity of consciousness. According to Chalmers, MP exists because we seem to have specific intuitions about our conscious states, e.g. regarding their non-physical nature or their peculiar epistemic qualities. The complexity of consciousness, the vagueness of its definition and the need of a

³⁸ For the constitutive and regulative function of HP and MP I took inspiration from the Kantian distinction between the constitutive and regulative use of ideas, in the *Transcendental Dialectics* (Kant, 1787).

coherent and complete explanation is not only discussed within the philosophical community, but also by cognitive scientists and neuroscientists. According to an online survey conducted in 2018 (80% of participants from academia), people believe that scientific research on consciousness has been making progress. However, most participants perceived also a lack of rigor in scientific studies of consciousness than other neuroscientific subfields (Michel et al., 2018). This may be due to the difficulty in assessing a common definition and in isolating the empirical characteristics of consciousness from other mental and behavioural features (Irvine, 2013a).

It seems reasonable to think that research on consciousness has made progress in the last decades. Several neuroscientific and cognitive theories are getting more sophisticated day by day, e.g. Global neuronal workspace theory, Integrated information theory, Temporo-spatial theory (Dehaene et al., 2011; Tononi, 2017; Northoff & Huang, 2017), focusing on every single feature that characterizes or is related to consciousness, e.g. self, agency, ownership, mind wandering, self-reflection, qualia, etc. However, as seen above, any functional, mechanical, physical, biological, cognitive, computational explanation appears to miss the point about the what it is like to be conscious of something (Jackson, 1986; Levine, 1983; T. Nagel, 1974). Accounting for this point would represent an answer to HP, but the issue here discussed is a different one: why does consciousness seem irreducible to standard explanations (HP) and raise intuitions about its presumed specificity (MP)? Other phenomena are particularly hard to explain through empirical research or conceptual theories, but still they do not represent a hard problem for the standard paradigm.

As Chalmers highlights (Chalmers, 2018, pp. 20–21), also other entities or concepts express ontological ambiguity or appears complex to define. Just consider all the difficulties in giving any empirically-based insight on prescriptive (good, bad, right, wrong, etc.) or abstract (fate, justice, beauty) concepts without falling into the naturalistic fallacy, i.e. the attempt to reduce them in terms of natural properties (G. E. Moore, 1903)³⁹. Alternatively, consider some micro or macro phenomena that are open

³⁹ The naturalistic fallacy was initially used by G.E. Moore against Spencer's Social Darwinism and any utilitarianist moral theory (the roots of this argument are to be found in Hume's is-ought problem, Hume, 1738). Basically, the attempt to determine the concept of "good" through natural properties (e.g. pleasure) is a naturalistic fallacy. More generally, every attempt to derive values from natural description is a naturalistic fallacy. Nevertheless, it seems reasonable to imagine an epistemological

to empirical investigation, but whose research is far from offering an ultimate answer and explanation (the number of dimensions that constitute the Universe, the nature of black matter, the smallest entity of physics, etc.). Despite being tough to deal with them, we can still imagine that one day we will find a functional and physical explanation for each of them. Accordingly, there seems to be some difference between “complex” problems and the “hard” problem of consciousness. This issue undoubtedly is related to Clark’s “meta-hard problem of consciousness”⁴⁰, however, assuming that there really is a hard problem, the point here is to understand the reason why there is a hard problem of consciousness and not a hard problem of space, mass, evil, justice, etc. Even though these are some of the most complex issues of contemporary science or persistent problems in the history of philosophy, they do not elicit any immediate intuition about their ontological nature and specificity. Therefore, they can be considered as easy problems, that is, using a more general interpretation of Chalmers’ concept (Chalmers, 1995), susceptible to the standard methods of natural sciences or conceptual philosophical analysis.

As I explained before, there are different ways of approaching a phenomenon, depending on the perspective or epistemic stance we use. It is important to note that not all the three types of perspective (FPP, SPP, TPP) are necessary in order to achieve an explanation or a description of a natural phenomenon. For example, in order to describe how gravity works and give a scientific explanation of it, we do not need the FPP or the SPP, i.e. we do not take into account my own single point of view on the thing that hit the ground during the experiment, nor an evaluation of my experience of gravity through a phenomenological interview. On the contrary, my subjective perspective can be an obstacle for an objective description of gravity. My common sense may lead me to think that a bowling ball falls faster than a feather because of its weight, while we all know that in the “Newton tube experiment” they would fall at the same speed. After all, as I will expose in detail later, the scientific method drawn by Galilei is based on a healthy scepticism towards our common sense and subjectivity,

relationship between the MP and the naturalistic fallacy. Indeed, one might interpret the latter as a sort of hard problem (no materialistic theory could ever give an ultimate definition of “good”, “bad”, etc.) and therefore formulate the meta-naturalistic fallacy, i.e. why do we think that there is a naturalistic fallacy? However, this would be an “*a posteriori*” argument, subsequent and less immediate compared to the intuitive level of the problem reports highlighted by MP.

⁴⁰ The problem whether there really is HP, “something special about phenomenal consciousness [...] or is it just a matter of explaining a pattern of responsiveness and report” (Clark, 2001, p. 185).

and on a rigorous process of experimentation in order to achieve a measurable quantification of some specific features of nature, that is, to take up a detached third-person perspective on natural phenomena.

This is what generally happens within the domain of science: the TPP is the preferential epistemic stance upon which we can start a scientific analysis. Even in the sociological and psychological sciences, the subjective features are decoded and converted into objective scales for a better measurement of the interested phenomenon. As said before, also consciousness can be analysed through empirical methods, therefore with a third-person perspective (neurophysiological, psychological and behavioural methods, etc.). The problem is that an objective/functional explanation of consciousness seems incomplete without the FPP, while an objective/functional explanation of gravity does not need any supplement from the subjective experience for being considered complete⁴¹. Indeed, the hard problem of consciousness claims for the incompleteness of any possible natural and functional explanation about what it is like to be conscious. From this “what it is like” emerges the intuition about the specificity of consciousness, i.e. problem reports and MP itself. Interestingly enough, “life” is a complex phenomenon and currently there is no complete agreement about the definition of it, hence someone has proposed a corresponding “hard problem of life”: how can information affect the world (Walker & Davies, 2017)? Basically, the main argument is that living systems are different from other physical systems in virtue of their causal efficiency and that the explanation of the intrinsic relation between information and life cannot be reduced in terms of known physics and chemistry, hence the “hard” problem of life. On the contrary, «new principles and potentially even new physical laws are necessary» (p.10). Now, I do not mean to analyse in detail the consistency of the proposal of a hard problem of life in biology, however this is useful for my purpose of explain the specificity of consciousness. Despite certain similarities, the hard problem of life is structurally different from the hard problem of consciousness. Indeed, the problem of how information affects the world is still included within the domain of physics (e.g. see the causal efficiency of living systems), such that the authors of the paper are expecting a revolution of the current physical

⁴¹ In the third chapter, I will analyse and discuss some attempts to reduce, translate and eliminate the subjective stance from the scientific studies of consciousness.

paradigm. Therefore, the problem certainly is complex, but not hard in the same way as Chalmers means, i.e. resistant to physical, chemical, computational or, more generally, functional explanation.

For argument's sake, let us try to reformulate the example of the DNA used by Chalmers, this time with life: "I can see that you have explained how life is related to information and how living systems are different from other physical systems, but you have not explained how it is a *life*". In this case, the question does not seem conceptually mistaken as the DNA's one, rather it seems reasonable to say that this kind of explanation does not solve how it is a life, or better, how it is to be alive. This should mean that the hard problem of life and consciousness are at least structurally the same. However, I would suggest that the explanation of "how it is a life", is not strictly related with the concept of life as such, but with "how it is like to be alive", therefore with the concept of consciousness. Walker and Davies are supposedly right in saying that current paradigms cannot reduce the complexity of the phenomenon of life, nevertheless looking for new paradigms and principles in physics will not help to explain "how it is a life", but only how living systems work. Indeed, the "hardness" of the problem of life does not lie at on the concept of life *per se*, but on the "what it is like to be alive", In other words, while it is possible to give a functional explanation to life, instead no physical or chemical theory will never explain how it is to be alive, i.e. what it is like to feel (therefore consciously) alive. Future physical principles and laws will be probably able to explain the phenomenon of life, but the fact that there is something like being alive is a completely different issue, i.e. the hard problem of consciousness. Then, to what extent the set of "what it is like to be alive" is coextensive with the set of living beings as such – and vice versa – is another more question, e.g. does a paramecium has conscious experience, i.e. is there something it is like to be that paramecium? And from the other side, can a conscious AI be defined also as alive (assuming here for argument's sake the possibility of machine consciousness)? For my purposes, it is enough to say that the problem of life, despite its undeniable complexity, does not seem to possess the conditions for being a hard problem in Chalmers' sense.

Now, let us consider another example of complex phenomenon that seems genuinely hard to explain and describe in a complete way, without contradictions between different perspectives: the concept of time. About time, science or philosophy

has offered explanations that diverge greatly from the way we perceive the related phenomena through common sense. For example, the concept of time in the theory of relativity is very different from our common idea of it: according to the theory of relativity, time is not absolute and linear, moreover it can dilate under particular conditions, completely diverging from the idea of time as a unidirectional arrow or an absolute container or a within which events and ourselves are put. Besides, the concept of time from the theory of relativity is very different from how we phenomenologically perceive time through retention, protention and duration, according to Husserl's phenomenology (Husserl, 1964). Moreover, it seems that contemporary physics is deconstructing our sense of time, claiming that is nothing but an illusion formed by our naïve perception of the world (Jaffe, 2018; Rovelli, 2019). However, the way in which we perceive time through our subjectivity and common sense is hardly replaceable with the scientific notion of time, at least in our daily life. On the other side, from a neuroscientific point of view, it seems impossible to abandon the concept of time (Benini, 2017). In this regard, there are many psychological and neuropsychological findings on the neural mechanisms that permit us to perceive things one after another within a temporal continuity, binding together visual and auditory stimuli (for details, Benini, 2017). Nevertheless, these functional explanations about how the brain encodes different stimuli and collect them together in a temporal order are really similar to those of the easy problems of consciousness, e.g. the ability to discriminate environmental stimuli and integrate information, while the subjective perception of time remains unanswered. That immediately links to Augustine's words, with which he started his analysis about the nature of time: «What then is time? If no one asks me, I know; if I want to explain it to a questioner, I do not know» (Augustine, 2006, p. 242). Accordingly, time seems a good candidate for being a hard problem in the same way as consciousness is, indeed it surely is a crucial subject of study for physics, but the physical explanation of time tells us nothing about what time is “for us”.

At this point, what do these physical and neurobiological explanations leave out? I would say “nothing”, because time actually is a suitable subject for scientific research, as one of the fundamental features of physics or as a neural mechanism crucial for our perceptual abilities. We can easily imagine that all the mysteries about

the physical nature of time and the neural activities related to it will be solved one day through functional explanations. However, this is not what Augustine meant with time, since that concept of time is what the phenomenological tradition, starting from Husserl, calls “time-consciousness” or “temporal consciousness” (*Zeitbewusstseins*; Husserl, 1964; Dainton, 2018). Time as an object of scientific research can be completely addressed by physics and neuroscience, but time as a fundamental structure of consciousness is close – although not identical – to what we mean for the hard problem, approximately “why is the performance of this neural mechanism accompanied by temporal experience?”. Therefore, time does not elicit a hard problem *per se*. As for the example of life, what makes the concept of time similar to a hard problem is that time is intrinsically related to consciousness, understood as self-temporalization⁴².

Taking together, it seems reasonable to think that consciousness specifically – if not exclusively – elicits a hard problem, i.e. it requires something more than a functional/physicalist explanation, while other phenomena not, despite being extremely complex. Note that I still am not claiming that the hard problem is real or the only real problem of consciousness (there are philosophical and scientific position arguing that the hard problem of consciousness is an illusion). I am only showing that consciousness appears as something special compared to other phenomena, with some epistemological or ontological specificity. Hence, the meta-problem of consciousness, i.e. why do we have problem reports about consciousness? In order to analyse better this issue, we need to look back to the concept of consciousness, looking to something that is generally presupposed in our common way of conceiving consciousness and the world.

1.3.3 Consciousness as Explanans

⁴² «Temporality must have the structure of selfness. Indeed it is only because the self in its being is there outside itself that it can be before or after itself, that there can be in general any before and after. Temporality exists only as the intra-structure of a being which has to be its own being; that is, as the intra-structure of a For-itself [consciousness]. Not that the For-itself has an ontological priority over temporality. But Temporality is the being of the For-itself in so far as the For-itself has to be its being ekstastically. Temporality is not, but the For-itself temporalizes itself by existing» (Sartre, 1956, pp. 136). I shall come back to this point in the next chapters.

In 1995, Chalmers defined consciousness as the “key explanandum of the science of the mind” (Chalmers, 1995, p. 206), i.e. what needs to be explained, since it is the central phenomenon of our mental life and the ultimate aim of any theory of mind. I completely agree with this idea that consciousness cannot be discarded by any complete explanation of the mind wannabe, nonetheless I consider that we also need to understand consciousness as the key *explanans*. Strictly speaking, the explanans is the set of statements that explains the phenomenon. In a more ontological context, we might consider the explanans as the condition of possibility of the phenomenon. And in this case, consciousness, besides being the key explanandum of science of mind, could also be understood as the condition of possibility of any explanation – and more generally of any experience. What I am referring to in a metaphorical way is the distinction between “actualist” and “objectualist” conceptions of consciousness, made by Mark Rowlands in *The Nature of Consciousness* (Rowlands, 2001). According to him, one of the peculiar features of consciousness is its hybrid structure, as it can be both object and act of experience (Rowlands, 2001, p. 122). Rowlands bases his claim on Sartre’s distinction between positional (thetic) and non-positional (non-thetic) consciousness⁴³. In particular, non-positional consciousness, in virtue of its non-reflective, non-intentional and non-dyadic features, is very close to be a transcendental condition of the act of consciousness, rather than a form of knowledge⁴⁴ or the awareness of an object (Rowlands, 2001, 2003a). The problem is that most recent

⁴³ Sartre argues that every act of positional consciousness (being conscious of an object) involves non-positional consciousness of that act (every consciousness is pre-reflectively, non-intentionally aware of itself). « If I count the cigarettes which are in that case, I have the impression of disclosing an objective property of this collection of cigarettes: they are a dozen. This property appears to my consciousness as a property existing in the world. It is very possible that I have no positional consciousness of counting them. Then I do not know myself as counting. Proof of this is that children who are capable of making an addition spontaneously can not explain subsequently how they set about it. [...]. Yet at the moment when these cigarettes are revealed to me as a dozen, I have a non-thetic consciousness of my adding activity. If anyone questioned me, indeed, if anyone should ask, "What are you doing there?" I should reply at once, "I am counting." This reply aims not only at the instantaneous consciousness which I can achieve by reflection but at those fleeting consciousnesses which have passed without being reflected-on, those which are forever not-reflected-on in my immediate past. Thus reflection has no kind of primacy over the consciousness reflected-on. It is not reflection which reveals the consciousness reflected-on to itself. Quite the contrary, it is the non-reflective consciousness which renders the reflection possible; there is a pre-reflective cogito which is the condition of the Cartesian cogito. At the same time it is the non-thetic consciousness of counting which is the very condition of my act of adding» (Sartre, 1956, p. liii).

⁴⁴ «The necessity of syntax has compelled us hitherto to speak of the ‘non-positional consciousness of self’. But we can no longer use this expression in which the ‘of self’ still evokes the idea of knowledge. (Henceforth we shall put the ‘of’ inside parentheses to show that it merely satisfies a grammatical requirement.)» (Sartre, 1956, p. liv).

discussions are dealing with the conception of consciousness as an object (Rowlands talks about “ubiquity of objectualism”, Rowlands, 2001, pp. 141–147)⁴⁵, while they are progressively eradicating the epistemic and ontological importance of the act of consciousness as such.

I would point out that this is not a defence of what Chalmers defines as knowledge and/or explanatory intuitions⁴⁶. However, I am using the actualist/objectualist distinction with reference to the conceptual meaning of “consciousness” and without any ontological commitment, so as to avoid to slip into the dualist division between the first- and third person perspectives (Jackson, 1986; Levine, 1983), nor into a revised version of the “what it is like to be” argument or a new definition of some subjective state, like qualia. Moreover, this is not a criticism of the materialistic account of consciousness, since even the qualitative side of experience defended by some realists (intrinsicness, subjectivity, immediate, etc.) is made by particular states and properties that are representations of consciousness, rather than being the act of directing towards a quite distinct object. For example, Rowlands highlights how even Nagel’s “what it is like to be” is grounding on the objectualist conception⁴⁷.

⁴⁵ «Suppose one focuses one’s attention on one’s consciousness, and in doing so becomes aware of various mental items – thoughts, feelings, desires, sensations, and the like – which possess various properties: a thought might be compelling, a feeling irritating, a desire troubling, and so on. When one thinks of consciousness as a collection of these sorts of items, possessing these sorts of properties, one is thinking of consciousness from, as I shall use the expression, the objectualist perspective. Consciousness viewed from this perspective is empirical: a collection of conscious items and properties» (Rowlands, 2003a, p. 207).

⁴⁶ Perhaps the core intuitions for the meta-problem as defined are explanatory intuitions holding that consciousness is hard to explain. These include gap intuitions holding that there is an explanatory gap between physical processes and consciousness, and anti-functionalist intuitions holding that explaining behavioural functions does not suffice to explain consciousness. [...] There are also knowledge intuitions: these include both first-person knowledge intuitions holding that consciousness provides special knowledge from the first-person perspective (like Mary’s knowledge of what it is like to see red on leaving the black and white room), and third-person ignorance intuitions, such as the intuition that it is hard to know the consciousness of other people or other organisms (such as what it is like to be a bat).

⁴⁷ «Nagel’s account of subjectivity rests on the idea that there are certain things – experiential facts – that, by their nature, are accessible only from a certain point of view. Since they are accessible, they are objects of consciousness. And the explanation of subjectivity, then, is that these objects of consciousness have a certain property, the property of being apprehensible only from a certain point of view. Subjectivity, thus, derives from the nature of an object of consciousness. And this is, as we have seen, an objectualist rather than an actualist explanation of subjectivity. Therefore, it seems that Nagel is committed to an objectualist interpretation of what it is like» (Rowlands, 2001, p. 142). He criticizes in the same way also the explanatory gap argument and the knowledge argument (pp. 141-146).

I shall come back to this point about Sartre's and Rowland's conception of consciousness in the next chapter, for our purposes we only need to know that beyond being a phenomenon with peculiar characteristics investigated by both science (neural and behavioural correlations) and philosophy (intrinsic qualitative aspects), consciousness lies first and under every conceptualization or reification of itself. Consciousness is, first of all, a pre-reflective relation of acquaintance of something, and only through consciousness there can be an object of knowledge. This distinction between the actualist (Rowlands might say also "transcendental"; Rowlands, 2003) and objectualist conceptions (both qualitative and quantitative) of consciousness could help in understanding why consciousness seems so special. We may have problem reports because experience is not a mere phenomenon, but fundamentally is the condition of possibility of the acquaintance of every other phenomenon. Complex natural phenomena or ontologically vague concepts fall together into the category of objects of consciousness and they do not elicit problem reports, while consciousness itself is the very same act of directedness towards them. Therefore, actual consciousness lies on (or better yet, is) the epistemic access for the class of objects in general, and this may be the main reason why we have the meta-problem of consciousness, while we do not have meta-problems for other complex phenomena.

One might argue that this is a strong realist position, therefore it cannot be used to clarify the debate between realism and illusionism. However, this is not the case, since the actualist/objectualist conception of consciousness comes before any ontological characterization of consciousness – on a transcendental-epistemic ground – and can be accepted by both illusionists and realists. As said, all the objective/subjective, physical/non-physical, qualitative/quantitative, internal/external distinctions fall into the objectualist conception of consciousness, as attempts to abstract and objectify the wider context of conscious experience. All those defences of the qualitative character of experience – first-person descriptions, what it is like, subjectivity, etc. – are objectualist positions, since they refer to some abstract phenomenal feature, e.g. the raw feel, the subjective sensation, the thought, and forgetting that consciousness is structured and embedded in a deeper and wider worldly context (Ratcliffe, 2007). In this sense, as I will discuss in detail in the second chapter, even Chalmers' claim about the need of an "extra ingredient in the

explanation” (Chalmers, 1995, p. 210) and other realist position about the existence of qualia are quite objectualist.

1.3.4 The Illusionist position on Consciousness

On the other side, an illusionist account of consciousness may claim that consciousness as qualitative, subjective, intrinsic, private, non-physical does not really exist, but is only an illusion created by the neural networks of the brain or something like that (Dennett, 2018). This seemingly coherent argument about the physicality or non-physicality of experience has nothing to do with the actualist conception of consciousness, i.e. the contextual act as such that support the objectification of consciousness, either as qualia or neural states. In other words, consciousness epistemically precedes and exceeds both subjective and objective representations of itself, because it is also the very same act of acquaintance of those representations. Then, this experiencing act might elicit biases and illusions or solid and lucid conscious states, but it would exist whatever the case. Denying the existence of an act that bears the awareness of something, e.g. the recognition of the surrounding environment, the execution of complex equations or just the perception of a red apple, would lead to what Galen Strawson calls the “Denial”, the claim that there isn’t really any experience, although it genuinely seems that there is experience (Strawson, forthcoming). Even though we seem to be conscious, the Deniers (consciously) believe that there are no beliefs, think that there are no thoughts. These are incoherent claims that leads to performative contradictions which can be related to the so-called Ishmael effect: in Herman Melville’s novel *Moby Dick*, if the ship was really destroyed by the whale, Ishmael would not have survived in order to tell the tale. In order to explain better this argument, I would use another metaphor similar to the Ishmael effect but the other way around: I would call it the “Clockwork Orange effect”, from the famous Stanley Kubrick’s movie *A Clockwork Orange*. By inverting the Ishmael effect, which highlights the incoherence between story and narration, in *A Clockwork Orange* Alex shows that the story would not have started without the voiceover of the main

character⁴⁸. The same goes, metaphorically, for consciousness. Any qualitative properties of experience or any empirical findings about the ontogenetic emergence of consciousness, i.e. any objectualist definition of consciousness, depends on the actualist explanation of consciousness *per se*⁴⁹.

How can this distinction be useful regarding the illusionist positions on consciousness? Chalmers claims that the strong illusionist can avoid the performative contradictions by considering “illusions non-experientially” – judgments, intuitions, dispositions, etc. (Chalmers, 2018, pp. 52–53). This could be a way to avoid the performative contradiction and distinguish the act of seeing, judge, feel, etc. from the object seen, judged, felt, etc. However, I cannot see how an illusion could be non-experiential. Just take the “grand illusion” of visual consciousness he brought as example: the visual detail, e.g. colours in the peripheral field, does not really exist at all, while the judgment about the visual detail exists but nonetheless is wrong. First, the judgment is not extrinsic to the very same conscious act that elicits the visual detail, since producing a judgment is still part of that experience. Secondly, I would point out that this argument is biased by the materialistic ontology and it cannot be useful for defining the plausibility or coherence of illusionism. Indeed, claiming that colours in the peripheral field would exist only if there really were colours in the visual field refers to a distinction between the “real” thing (the thing “as such” according to the contemporary physical science) and the appearance (what it is not actually out there in the physical world, but only an internal product of our phenomenality)⁵⁰. If we remain within the limits of the alleged distinction between appearance and reality, the

⁴⁸ The movie is the story of the main character Alex, narrated by Alex himself. Towards the end, Alex attempts suicide, and the fade-out effect of the scene makes us think that he actually was killed in the fall. However, in the next scene, we see that Alex is still alive, even though severely injured, and his voiceover says: “I jumped, O my brothers, and I fell hard, but I did not snuff it. If I had snuffed it, I would not be here to tell what I told have!”.

⁴⁹ «Consciousness is not only what is revealed from the objectualist perspective, it is also the adopting of an objectualist perspective. That is, consciousness is also the adopting of the perspective from which consciousness is seen to comprise a collection of items and properties. To think of consciousness as the directing of awareness rather than that upon which awareness is directed is to think of consciousness as the adopting of an objectualist perspective, either towards itself or towards external items. And to think of consciousness as the adopting of an objectualist perspective, rather than as what is revealed from this perspective, is to think of consciousness from, what I called the actualist perspective» (Rowlands, 2003a, p. 207).

⁵⁰ I will define in detail this distinction in the fourth chapter, as the “appearance-reality fallacy”.

illusionist can easily claim that “we actually see less than we think we see”⁵¹; this argument might be used as evidence for illusionism only if we strongly believe that nothing but physical science can determine what really exists in the universe. However, it would not help in the realist/illusionist debate on MP, as the argument is already ontologically committed. Conversely, from a wider phenomenological (and not phenomenal)⁵² perspective we could argue that we actually see more than what we receive in terms of mere stimuli⁵³ (Gallagher & Zahavi, 2012). Then, this ‘surplus’ in vision might be illusory products of the brain or a real non-physical property. Whatever the case, the phenomenological act of consciousness would be there⁵⁴, as its essence is to be undeniable.

⁵¹ Or, moderately, someone may say that «although we see more than the handful of objects, claimed by prominent models of visual attention and working memory, we still see far less than we think we do» (Cohen et al., 2016, p. 324).

⁵² I used here “phenomenological” as a way of describing actual consciousness, and it must not be confused with “phenomenality”: «[...] the terms “phenomenology” and “phenomenological” are often used as synonyms for “phenomenality,” i.e., as labels for the qualitative character of experience. To discuss phenomenology in that context is consequently to discuss a certain dimension of experience and at best to offer first-person descriptions of what the ‘what it is like’ of experience is really like. This way of talking about phenomenology has, however, little to do with phenomenology understood as a specific method or tradition in philosophy. When thinkers like Husserl, Scheler, Sartre, Heidegger and Merleau-Ponty engaged in phenomenological philosophizing, they most definitely were not simply seeking to offer fine-grained descriptions of the qualitative character of different experiences; they were not simply seeking to attend to hitherto unreflected and unnoticed aspects and details of lived experience» (Zahavi, 2018, p. 6).

⁵³ Moreover, the idea of the “poorness” of visual consciousness is far from being solid and ultimately proven from the empirical approach (Haun et al., 2017; Knotts et al., 2019).

⁵⁴ «If I have an experience characterized by this what-it-is-like-ness then (a localized part of) the world cannot fail to appear bright red. It may be that my experience is illusory. If so, there is no postbox that appears bright red, but some other object does so. Or it may be that I am hallucinating. If this case, there is no object that appears bright red. However, it is still true that a certain region of the visual world appears bright red» (Rowlands, 2015, p. 45).

2. The Logical and Phenomenological Structure of Consciousness

In this chapter, I will propose a provisional blueprint of consciousness. In other words, in the midst of so much critical literature about the problems of consciousness, I will try to start the analysis of the phenomenon of consciousness from the very same definition that we generally give to the term “consciousness”. In this regard, I will use Kripke’s direct reference theory, in order to analyse the term “consciousness” in a non-descriptive way, as if “consciousness” were a rigid designator in the same way as the terms “gold” or “tiger” with their respective natural kinds. The Kripkean discussion will lead me to critically discuss the validity of the definition of “consciousness” as “neural activity x”, which is affirmed by many supporters of the mind-brain identity. Secondly, after offering the provisional definition of consciousness as “being aware of something”, I will try to determine the phenomenological structure of our experience, focusing on the issue of pre-reflective consciousness, developing a fruitful dialogue with various contributions from phenomenology (Sartre, Merleau-Ponty, Gallagher, Zahavi), and delineating some preliminary features, like phenomenality, intentionality and self-awareness. These features will then be tested in the last part of the chapter, through some peculiar thought and real experiments on the possibility of unconscious experiences (refrigerator light problem, long-distance truck driver, blindsight).

2.1 What Starting Definition for Consciousness?

Is there a definition that can be used as a starting point for a theory of consciousness? After having exposed all the variances of meanings, and all the issues raised by specific kinds of description, it seems quite problematic to give a univocal answer to this question, for many reasons. First, as I exposed here above, there are lots of definitions, from various perspectives (FPP, SPP, TPP) and approaches (methodological,

ontological, epistemological, phenomenological, etc.). Indeed, each theory is based on a set of core concepts, which derive from some (empirical, analytical, etc.) approach, perspective or method. Thus, each theory of consciousness refers to an ontological and metaphysical commitment about what consciousness is, which can compromise from the start any attempt to offer a complete description of its nature. Still, at the same time, it seems impossible to avoid any provisional ontological commitment about such a “universal” phenomenon. Second, it seems difficult to gather together all the aspects of consciousness within an objective or subjective category or definition, because of its specific hybrid structure. Third, even within any single approach there are sub-categories of definition, depending on the kind of method used (as in cognitive science) or on the variability of the meanings used (as in phenomenology¹). For example, Marie Guillot has highlighted the extreme variation of words used (seven) for the definition of “subjectivity” within a passage from Shaun Gallagher and Dan Zahavi². Surely, the richness of language offered sometimes by the phenomenological analysis is very important, principally in order to avoid any conceptual reductionism. However, this does not help us in the determination of a provisional but common starting point for the study of consciousness. Instead, it betrays a certain “conceptual blurriness” in the discussion (Guillot, 2017, p. 26), and this is a critical point often raised by scientific and reductionist positions against qualitative and phenomenological researches about consciousness. Fourth, almost all the definitions exposed so far can be considered analytic a priori propositions of the concept of consciousness, i.e. they give us some detailed descriptions of how consciousness is used, without however offering any information not already contained in the term “consciousness”³. For example, saying

¹ For example, regarding phenomenal consciousness, just consider all these concepts and meanings often used as synonyms: subjectivity, for-me-ness, me-ishness, me-ness, myness or mineness, first-personal character or first-personal givenness of experience, non-reflective or pre-reflective self-awareness, low-level self-consciousness, the sense of self, sense of ownership (see for details Guillot, 2017, p. 25).

² «The notion of pre-reflective self-awareness is related to the idea that experiences have a subjective ‘feel’ to them, a certain (phenomenal) quality of ‘what it is like’ or what it ‘feels’ like to have them. [...] [A]s I live through these differences [between various experiences], there is something experiential that is, in some sense, the same, namely, their distinct first-personal character. All the experiences are characterized by a quality of mineness or for-me-ness, the fact that it is I who am having these experiences. [...]. All of this suggests [...] that (phenomenal) consciousness consequently entails a (minimal) form of self-consciousness. In short, unless a mental process is pre-reflectively self-conscious there will be nothing it is like to undergo the process, and it therefore cannot be a phenomenally conscious process [...]» (Gallagher & Zahavi, 2016; see Guillot, 2017, p. 26).

³ «Still, the words ‘consciousness’ and ‘experience’, used imprecisely, can muddy the debate, so let me make as plain as I can what I mean by them. Unfortunately, the two most respectable avenues for definition appear to be closed. No analytic definition is possible, I suspect, because ‘consciousness’ or

that being conscious is something it is like to have an experience is already contained in “consciousness”, so that we can also overturn the sentence, saying that having an experience is to be conscious⁴.

I am not saying that we cannot start with a description or a set of descriptions of what we believe consciousness is, in effect we have to start somewhere in addressing this issue. On the other hand, any description we use is not necessarily a synonym of the word “consciousness”, that is, the meaning of “consciousness” does not lie on a set of descriptions. So far, I have talked about consciousness in many ways, without giving a description that fixes univocally the reference of the name, indeed, we competently mention and understand a lot of words without a complete knowledge of their conditions or with a partial description of the referent⁵. Moreover, things get complicated with consciousness, because it is difficult to describe its meaning avoiding any circularity, e.g. consciousness = phenomenal = qualia = what it is like = experience = consciousness, again. Epistemically speaking, we do not need the knowledge of the descriptive content univocally referred to “consciousness” in order to understand the word *per se*. Therefore – and this may seem quite obvious – in order to be acquainted with the term “consciousness” and consequently start an analysis of its conditions and features, we do not need beforehand to correlate to it any set of descriptions.

Accordingly, rather than introducing an already given empirical or theoretical account of consciousness, I would begin from the way we mention “consciousness” in our common sense, in order to see whether we can use it as a provisional starting point

‘experience’ (I use the words interchangeably) is a foundationally simple concept indivisible into component parts. It doesn’t seem to be like ‘bachelor’(a marriageable but unmarried man)or ‘quadrilateral’(a four-sided planar figure).Nor can we, at this stage of our knowledge, characterize consciousness functionally, by appeal to the role it plays in a system (a ‘heart’ is an organ that pumps blood, ‘currency’ is whatever physical tokens serve as the medium of exchange), since the functional role of consciousness is still very much in dispute. One characterization some people find helpful is this: Your conscious experience is whatever it is by virtue of which there’s ‘something it’s like to be you’, while there’s nothing it’s like(presumably!) to be a rock or a toy robot. But this characterization has all the disadvantages of definition by synonymy. Unless you’re already comfortable with ‘what it’s like’ talk (Nagel, 1974), it won’t help much. We can invoke other synonyms as well: ‘phenomenology’, ‘subjective experience’, Block’s (1995) ‘phenomenal consciousness’, Chalmers’ (1996) ‘qualia’» (Schwitzgebel, 2007, pp. 7–8).

⁴ This reminds us of Quine’s critique of the concept of analyticity as a circular definition. Analyticity refers to the concept of synonym while this one is determinable only starting from the former (Quine, 1951).

⁵ See the instances made by Kripke about Cicero and Feynman (Kripke, 1980, pp. 79–82).

for the discussion about its nature. In order to do this, I need to make use of Saul Kripke's direct reference theory, taken from *Naming and Necessity* (Kripke, 1980).

2.1.1 Kripke's Direct Reference Theory

Kripke argued that names and natural kinds are not synonymous with the descriptions associated with them. The description could be used initially to fix the reference of a name or kind, but this does not give the meaning of the term, and typically it needs not be known by users later in the history of the name (as often happens within our everyday vocabulary). Instead, a term designates an object without the need of a mediating meaning between language and reality, for example, the description "the clear drinkable liquid" does not represent the meaning of the term "water". According to Kripke's modal argument, a term designates x rigidly if the term designates x with respect to every possible world in which x exists and does not designate anything else with respect to worlds in which x does not exist (Salmon, 2005). Kripke principally holds that proper nouns are rigid designators, for example "Aristotle" designates the same person in every possible world, while the description "the most important Greek philosopher" is not an analytic statement related to Aristotle, as the descriptivist theory would sustain (Frege, Russell, Wittgenstein), rather it is a contingent one. Therefore, "the most important Greek philosopher" is not a synonym of Aristotle, because it is not necessary in every possible world, e.g. in some possible world Aristotle could have been a farmer. Also, certain general terms – including natural-kind terms, like "water", phenomenon terms, like "heat", and biological taxa, like "tiger" – are rigid designators, e.g. "water = H_2O " and "heat = molecular motion" are necessary in every possible world⁶. Of course, it is possible that these identities are mistaken, and maybe it will turn out that water is made of another different or new chemical element, i.e. the fact that water is H_2O cannot be determined a priori. However, this would not affect Kripke's theory, since that *if* "water is H_2O " is true, then water is necessarily H_2O . So,

⁶ Note that here the metaphysical concept of "necessity" is not the same as the gnoseological concept of "a priori". Stating that " x is necessarily true" means that x is true in every possible world, while stating that " x is true a priori" means that x is true independently from experience (Kripke, 1980, p. 38).

although our knowledge about water may not be complete (probably it is not yet), this will not change the fact that, if it is true, water is necessarily H₂O.

Kripke made also the example of gold: if one day it will turn out that gold, maybe due to a sort of hallucination, is not yellow but blue, this will not mean that gold never existed, but simply that gold has turned out not to be yellow, but blue. This because the description “being yellow” is not rigidly applied, e.g. in another possible world it might not apply to gold but to other stuff altogether, therefore the property “being yellow” of gold is not necessary in every possible world⁷. Instead, Kripke argues that “gold is the element with atomic number 79” is a theoretical – i.e. not contingent – identity involving two rigid designators from two different kinds, respectively, “gold” from the common-sense natural kind and “atomic number 79” from the scientific natural kind. This means that this statement – “gold = element with atomic number 79”, if true, is necessary in every possible world, i.e. we cannot conceive the term “gold” referring to something different from that metal with atomic number 79, as well as for “water” to H₂O, “heat” to molecular motion. Hence, if we find a metal that looks exactly like gold, but it is actually a different substance, then this would not be a variance of gold, but a completely different thing⁸. On the other

⁷ «Could we discover that gold was not in fact yellow? [...] Suppose there were an optical illusion which made the substance appear to be yellow; but, in fact, once the peculiar properties of the atmosphere were removed, we would see that it is actually blue. [...] Would there on this basis be an announcement in the newspapers: “It has turned out that there is no gold. Gold does not exist. What we took to be gold is not in fact gold.” [...] It seems to me that there would be no such announcement. On the contrary, what would be announced would be that though it appeared that gold was yellow, in fact gold has turned out not to be yellow, but blue. The reason is, I think, that we use “gold” as a term for a certain kind of thing» (Kripke, 1980, p. 118).

⁸ «Further, there might be a substance which has all the identifying marks we commonly attributed to gold and used to identify it in the first place, but which is not the same kind of thing, which is not the same substance. We would say of such a thing that though it has all the appearances we initially used to identify gold, it is not gold. Such a thing is, for example, as we well know, iron pyrites or fool's gold. This is not another kind of gold. It's a completely different thing which to the uninitiated person looks just like the substance which we discovered and called gold. [...] Now consider a possible world. Consider a counterfactual situation in which, let us say, fool's gold or iron pyrites was actually found in various mountains of the United States, or in areas of South Africa and the Soviet Union. Suppose that all the areas which actually contain gold now, contained iron pyrites instead, or some other substance which counterfeited the superficial properties of gold but lacked its atomic structure. Would we say, of this counterfactual situation, that in that situation gold would not have been an element (because pyrites is not an element)? It seems to me that we would not. We would instead describe this as a situation in which a substance, say iron pyrites, which is not gold, would have been found in the very mountains which actually contain gold and would have had the very properties by which we commonly identify gold. But it would not be gold; it would be something else. [...] (Once again, whether people counterfactually would have called it 'gold' is irrelevant. We do not describe it as gold) [...] Given that gold is this element, any other substance, even though it looks like gold and is found in the very places where we in fact find gold, would not be gold» (Kripke, 1980, pp. 118–119, 124–125).

hand, considering another possible world, suppose that we have in mind a situation which is like the actual world, but different in respect of how some things are, like gold or heat. In this possible world, we can try to imagine that gold have turned out not to have atomic number 79 (or that heat have turned out not to be molecular motion). However, according to Kripke these theoretical identities, if true, are always necessary true even if they are a posteriori, i.e. derived from empirical discovery. Therefore, a proposition which seemed clearly to be possible, e.g. that gold is a compound rather than an element with atomic number 79, turns out on closer inspection not to be possible, and therefore it «should be replaced (roughly) by the statement that it is logically possible that there should have been a compound with all the properties originally known to hold of gold» (Kripke, 1980, pp. 142–143). This because gold has been initially baptized as “the metal with the atomic number 79”, and this is a rigid designation, as well as the proper noun “Aristotle” refers to the same person in every possible world, regardless of the fact that Aristotle might have been a farmer, a king or a philosopher⁹. Hence, “gold has the atomic number 79”, if true, is a necessary statement, nevertheless it is a posteriori, because it comes from the empirical findings about the substance we initially baptized as “gold”. This leads to the possibility of necessary a posteriori statements (Kripke, 1980, p. 140).

2.1.2 A Kripkean-like definition of Consciousness

What about the term “consciousness”? Can we apply the direct reference theory to it? Consciousness is not a proper noun, but someone, like Searle, might consider it as a natural kind or a biological phenomenon, like digestion, photosynthesis, mitosis, etc¹⁰.

⁹ «The general answer to the objector can be stated, then, as follows: Any necessary truth, whether a priori or a posteriori, could not have turned out otherwise. In the case of some necessary a posteriori truths, however, we can say that under appropriate qualitatively identical evidential situations, an appropriate corresponding qualitative statement might have been false. The loose and inaccurate statement that gold might have turned out to be a compound should be replaced (roughly) by the statement that it is logically possible that there should have been a compound with all the properties originally known to hold of gold» (Kripke, 1980, pp. 142–143).

¹⁰ Our world picture, though extremely complicated in detail, provides a rather simple account of the mode of existence of consciousness. According to the atomic theory, the world is made up of particles. These particles are organized into systems. Some of these systems are living, and these types of living systems have evolved over long periods of time. Among these, some have evolved brains that are capable of causing and sustaining consciousness. Consciousness is, thus, a biological feature of certain organisms in exactly the same sense of “biological” in which photosynthesis, mitosis, digestion and

For the aims of this work, I am not going to discuss the validity of this position, nor to hold already any ontological commitment towards consciousness (I would contradict myself then). Also, even in case what Searle says was true, the statement that consciousness is a natural process is quite more complex than “water = H₂O”. But the question here is: since it appears quite problematic to start the analysis of consciousness through a set of descriptions coming from either empirical or conceptual domains, can we say something about consciousness in a Kripkean-like way? I say “Kripkean-like” because I do not mean to apply *verbatim* Kripke’s direct reference theory, but to use it as an inspirational tool for avoiding the swampy and murky waters of the ontological and epistemological battles between dualism and monism, FPP and TPP, a priori and a posteriori, etc. Moreover, instead of understanding the direct reference theory as a rigorous foundation of the scientific essentialism, i.e. some properties are necessary conditions for an object to be part of a given natural kind, I rather interpret it as an attempt to define the concept of reference in an intuitive, pre-theoretic way, in other words, as a way of explaining how names work according to common-sense. This would be useful as a starting point for the analysis of consciousness, at the very least, we would not be stuck immediately in the endless metaphysical or scientific diatribes. After that, we may also question the concept of consciousness and refine it through different methods (phenomenological, scientific, etc.).

Is the term “consciousness” a rigid designator? Can “consciousness” designate something rigidly with respect to every possible world? Generally, we use consciousness for referring to some behavioural and psychological state of someone, for example, taking the expression from Searle that state «of sentience or awareness that typically begin when we wake up in the morning from a dreamless sleep and continue throughout the day until we fall asleep again» (Searle, 1999: 40-41). Basically, we are able to determine if someone is conscious or if he lost his senses and we are able to do it pragmatically, without possessing specific philosophical or medical knowledge. Nevertheless, we can be wrong in some of our assignments, e.g. in case someone is pretending to be asleep or in severe situations, in which it is necessary to

reproduction are biological features of organisms (Searle, 1992, pp. 92–93). This position has been still sustained by Searle recently, see Searle, 2019).

use clinical scales in order to rightly assess the level of consciousness (coma, unresponsive wakefulness syndrome, minimally conscious state, etc.). On the other hand, it seems impossible to be deceived about our own conscious state when I refer to myself as “conscious” at a given moment¹¹, e.g. I can be deceived about the notebook in front of me which I am conscious of, but not of the fact that I am conscious of something, even of an hallucination. Anyway, there are situations that superficially resemble a conscious-like state to which consequently we would probably refer to them as conscious states, but that actually are not related with consciousness, and, vice versa, other situations that superficially resemble a non-conscious state, but where the person is actually conscious. Following the analogy with proper nouns and natural kinds, also consciousness may present some superficial features that are contingent, just as much as gold is yellow (as well as iron pyrite) or water is colourless (as well as vodka).

So far, I have discussed consciousness in a very general way, however in our daily lives we are not dealing with an empty and abstract concept, but with a full-fledged and lived experience of the world. First of all, there is never “consciousness as such”, in fact consciousness is always and intrinsically “consciousness of something”, i.e. if I am aware, I must be aware of something, and this something can possess different epistemic forms, e.g. it can be an object of perception or imagination, a pure sensation or an abstract thought, etc. This intrinsic aboutness or directedness towards an object that characterizes consciousness (and more generally mental states) is called “intentionality” (from Latin *tendere*: directedness towards, attending to, referring to something; see Jacob, 2019). Thus, when we perceive, believe, think, imagine, etc. we are perceiving, believing, thinking, imagining “something”, and through this the world is manifesting to us. The concept of intentionality was introduced into contemporary philosophy by Franz Brentano, who took inspiration directly from Aristotle’s theory of sensation and understanding (*De Anima*, chapter 12), Thomas Aquinas and the following scholastic tradition (Caston, 2008; Taieb,

¹¹ Of course, for an eliminativist or an illusionist the fact that we feel confident in knowing when we are conscious does not prove anything about the supposed reality of consciousness. As already seen, according to them we merely believe to be conscious, but there is nothing like this.

2018), and also by Edmund Husserl, who made intentionality one of the fundamental concepts of phenomenology.

There is a huge debate about this concept, for example whether intentionality could be considered the essential “mark of the mental” – as Brentano claimed – or only a necessary condition for mental states/events (Crane, 2009), or whether conscious states are intrinsically intentional (Sartre, 1956, 1970) or some of them can be non-intentional (like nervousness or anxiety, Searle, 1983) or even whether intentionality as the typical element of mental states (Voltolini, 2013) or phenomenal consciousness (Robinson, 2009) should be rejected in contemporary philosophy of mind. I cannot address here the question about the nature and legitimacy of the concept of intentionality (I will get back later to that, though). However, for the moment it shall be sufficient acknowledging that our conscious states are characterized from time to time by a different experienced object, so that we can reasonably say that our consciousness, in a broad sense, consists of a variety of “consciousness of something”, at least in the way we commonly intend the term consciousness, e.g. “I am (consciously) perceiving an apple”, “I have consciousness of this chair”, I am (consciously) aware of the consequences of my act”, etc. For the purpose of our argument and analogy, we may argue that in our phenomenal life we instantiate forms of “consciousness-of-something” that refer to the broad concept of consciousness, as well as the terms “Siberian tiger”, “white Bengal tiger” and even the “Bali tiger” (extinct) refer to the species of tiger (*Panthera Tigris*).

Regarding this last example, the term “tiger” refers to different species and instances of tiger, regardless of some contingent variations, e.g. colours, anatomical defects, different sizes and weights, etc. In effect, we could see a white pet tiger, deprived of the tail or a leg, but none of these characteristics would induce us not to refer to it as a tiger. In this case, the term “tiger” rigidly refers to its internal structure as the mark of the species of tigers and not simply to the appearances of the tiger as a big cat. Interestingly enough, we are keen to say that an animal that apparently differs from a tiger but with the same internal structure of a tiger is precisely a tiger, while a seeming tiger with the internal structure of a reptile is likely to be a reptile¹². This is

¹² «Just as something may have all the properties by which we originally identified tigers and yet not be a tiger, so we might also find out tigers had none of the properties by which we originally identified

not because the old appearance-based concept of tiger has been substituted by a new science-based one, indeed we are simply supposing – not knowing – that the internal structure of the tiger, e.g. DNA, can determine the criterion for a natural kind, and we are stating this by a posteriori investigation (Kripke, 1980, pp. 120–121). So, “tiger” is a rigid designator, which means that we do not understand it due to some identifying definition, or to a cluster of descriptions. Thus, we can track tigers across counterfactual situations, completely independent of any physical characteristics, e.g. tigers without legs or tail, or strangely coloured tigers. Moreover, some other species could have had the external appearance of tigers, e.g. some other cat-like animal, with orange fur and black stripes, etc., however, our term “tiger” picks animals out as they are in the actual world, and this is what fixes the species across possible worlds.

If the analogy holds, we understand the meaning of “consciousness”, as seen above, not necessarily because of any particular description, nor to a cluster of descriptions, rather we can use “consciousness” across extremely different situations (consciousness of pain, of pleasure, of a red apple), e.g. saying “he is conscious now”, “this is a conscious being”, “I am feeling something”, “she is in pain”, “she passed out” (i.e. “she has not consciousness”), etc. Also, and this is very important, our general term “consciousness” picks instances of consciousness in the actual world, fixing the reference in every possible world, so that we can conceive a possible world in which, for example, consciousness is revealed through a different cluster of behaviours (closed eyes, little movements, no observable reactions) or different physical activities (neural activity in the cerebellum, without any significant activity in the cerebrum, activity in other organs, like the liver, or even activity in an electronic circuit board), nevertheless it would still necessarily refer to that process of being acquainted/aware of something.

This can be used also as a counterargument against the illusionist position, as I exposed in the last part of the first chapter. The strong illusionists, i.e. the Deniers,

them. Perhaps none are quadrupedal, none tawny yellow, none carnivorous, and so on; all these properties turn out to be based on optical illusions or other errors, as in the case of gold. So the term 'tiger', like the term 'gold', does not mark out a 'cluster concept' in which most, but perhaps not all, of the properties used to identify the kind must be satisfied. On the contrary, possession of most of these properties need not be a necessary condition for membership in the kind, nor need it be a sufficient condition» (Kripke, 1980, p. 121).

hold that consciousness is an illusion and ontologically it does not exist at all, differently from the weak version of illusionism, which argues that we feel like we perceive, for example, a colourful world outside our self, but actually our experience is not like this. The solution to the meta-problem of consciousness is that there is nothing like qualia or experience to explain or reduce, the only thing to explain is how we have this illusion of being conscious, i.e. the illusion problem (Frankish, 2016). For argument's sake, we can conceive this situation as a counterfactual in our actual world or even as a possible world in which all the people live into the illusion of being conscious, while there is nothing like this at all (just consider the illusionary world illustrated in the movie *The Matrix*, where people live in a simulation, while they actually are “grown” into pods). In that case, any description or set of descriptions, as expected, would fail, because consciousness cannot be described here as some particular behavioural modification or brain activity (we can imagine that those people have a different brain, for the purpose of being grown in the foetus fields). Nevertheless, consciousness would still rigidly designate the fact that, even in that simulation, people are “conscious of something”, e.g. the illusionary chair, the hallucinatory pain, the illusionary pleasure, even conscious of absent objects¹³. The content of consciousness may be illusory (the chair can be a hallucination) and even the sensation may be illusory as well (the redness of an apple that is actually green¹⁴),

¹³ See, for example, how Sartre defends the possibility to positionally conscious of non-existent objects: «Every consciousness posits its object, but each in its own way. Perception, for example, posits its object as existing. The image also includes an act of belief or a positional act. The act can take four and only four forms: it can posit the object as nonexistent, or as absent, or as existing elsewhere; it can also ‘neutralize’ itself, which is to say not posit its object as existent. Two of these acts are negations; the fourth corresponds to a suspension or neutralization of the thesis. The third, which is positive, assumes an implicit negation of the natural and present existence of the object. These positional acts – this remark is crucial – are not superimposed on the image after it is constituted: the positional act is constitutive of the image consciousness. [...] The characteristic of the intentional object of the imaging consciousness is that the object is not there and is posited as such, or that it does not exist and is posited as nonexistent, or that it is not posited at all» (Sartre, 2004a, pp. 12–13).

¹⁴ We may say, however, that some qualitative sensations are a whole with the act of consciousness itself, like pleasure, pain or even colourful experiences. See the examples made by Sartre: «Pleasure can not be distinguished – even logically – from consciousness of pleasure. Consciousness (of) pleasure is constitutive of the pleasure as the very mode of its own existence, as the material of which it is made, and not as a form which is imposed by a blow upon a hedonistic material. Pleasure can not exist “before” consciousness of pleasure-not even in the form of potentiality or potency. A potential pleasure can exist only as consciousness (of) being potential. Potencies of consciousness’) exist only as consciousness of potencies. Conversely, as I showed earlier, we must avoid defining pleasure by the consciousness which I have of it. This would be to fall into an idealism of consciousness which would bring us by indirect means to the primacy of knowledge. [...] There is no more first a consciousness which receives subsequently the affect “pleasure” like water which one stains, than there is first a pleasure (unconscious or psychological) which receives subsequently the quality of “conscious” like a pencil of light rays.

however what cannot be contingent in every possible world is the fact that I am instancing something consciously, and I am referring to myself as “conscious of something”. Using the conceptual terminology exposed in the previous chapter, the objectual consciousness is made by the set of possible descriptions (both objective and subjective) that can characterize a conscious state, and therefore is contingent because it may be different in some possible world, while the act of consciousness that constitutes the experience as such and to which I refer when I say “I am conscious of this chair” or “she/he is conscious of what happened” – generally as “being conscious/aware/acquainted with something” – is necessary in every possible world, even where consciousness is an illusion.

For this reason, I would consider “consciousness” a rigid designator, because it necessary refers to the act of being aware of something in every possible world. In other words, if true, “consciousness = being aware of something” is necessary in every possible world. Hence, as Kripke did with “gold” and iron pyrite or “tiger”, suppose it turns out that our experience is a deceptive illusion, and that our phenomenal world is not full of sounds, colours, sensations, flux of thoughts and also that our conscious memories are completely fake a posteriori reconstructions; would we then say that there turn out to be no consciousness after all – as the Deniers would say? I think we would say that, due to some reason that deceived our intuitions, consciousness is not what we think it is, but not that consciousness does not exist. At the same time, suppose that in a possible world we discover a person who, though having all external appearances of a conscious being (e.g. behaviour), has a completely different internal structure from the human one (e.g. a synthetic skeleton with synthetic organs, or even better a metal skeleton with gears, pumps and electronic devices). In this case, I believe that the assignment of consciousness would at least be questioned.

2.1.3 A necessary a posteriori Statement for Consciousness?

So far, I have exposed a Kripkean-like direct reference theory, understanding the general term “consciousness” as if it were a natural kind or a proper noun, and I

There is an indivisible, indissoluble being—definitely not a substance supporting its qualities like particles of being, but a being which is existence through and through» (Sartre, 1956, p. liv).

claimed that “consciousness” directly refers to the experiential act of being aware of something in every possible world. At this point, one may argue that this merely is a synonym of the term, which simply paraphrases the supposed meaning of consciousness. Also, one may argue that, instead of this, I should offer a reliable statement about consciousness that is at the same time necessary and a posteriori (like H₂O for water). This argument partially misses the point, given that the preliminary understanding of “being aware of something” is only the empty structure of the direct reference of consciousness, because that structure must be fulfilled every time we refer to consciousness with concrete instances of conscious acts, in the same way as the term “water” directly refers not to some abstract H₂O chemical definition, but to “this” or “that” instance of H₂O, e.g. a glass of water or the water of a river, etc. In other words, saying that “consciousness” refers to being aware of something does not mean anything if we do not complete it – in an ostensive way – with instances of experience, hence consciousness of pain directly refers to “that” pain state, as well as consciousness of red directly refers to “that” red object (imagined or presented), etc. Thus, “consciousness” does not refer to some general consciousness act, as well as when we mention “tiger”, “gold”, “water”, we refer to instances of tiger, gold and water.

Moreover, our general understanding of the meaning of “consciousness” is not directly due to a set of descriptions of the phenomenon we are supposed to know, rather it is given by the reference to our own examples¹⁵. Anyway, it must be said that the level of knowledge we possess about water, tiger and gold is more detailed and deeper than about consciousness. Of course, we have a direct acquaintance of what it is like to be conscious, while we needed to start a process of knowledge in order to get acquainted with water, gold or tiger. With “level of knowledge” I mean what we know about what Kripke calls the “internal structure”, e.g. H₂O for water, the atomic number

¹⁵ «The crucial point is that the word “consciousness,” as pretty much everyone uses it, is defined largely by reference to our own example. We don’t have access to some separate, human-independent definition of consciousness, which would allow us even to frame the question of whether it’s possible that toasters are conscious whereas humans are not. By analogy, imagine 19th-century scientists built a thermometer that delivered the result that boiling water was colder than ice. The possibility that that was true wouldn’t even merit discussion—it would be immediately rejected in favor of an obvious alternative, that the thermometer was simply a bad thermometer, since it failed to capture our pre-theoretic notion of what temperature is even supposed to mean, which concept includes boiling water having a higher temperature than ice» (Aaronson, 2014).

79 for gold or molecular motion for heat. Instead of “internal structure”, which implicitly refers to the idea of something inside, internal and real opposed to something outside, external and apparent¹⁶, I would prefer using the terms “intrinsic” or “constitutive”, that is, something basic, necessary and fundamental, opposed to something extrinsic, contingent and not essentially related to the thing. Therefore, I will use “intrinsic or constitutive structure of consciousness” as those conditions that are necessary for consciousness in all possible worlds.

Following Kripke’s theory, if we want to define the intrinsic structure of consciousness, we should not settle for a mere analytic definition of consciousness, but we should consider the possibility of a necessary a posteriori statement, like “water is H₂O”. As I said above, the statement that consciousness means being aware of something is not sufficient for this aim, and constantly needs to be fulfilled by instances of conscious states. In this respect, we would need to identify those elements that are necessary for the existence of these various instances of consciousness, in a similar way as H₂O is fundamental for the existence of any instance of water (the water in the cup, the water of the river, etc.). Now, the first question is about whether it is possible to find a necessary a posteriori statement for consciousness. Consequently, the second question will be about what kind of necessary a posteriori truth we can find for consciousness, and, in effect, the intrinsic structure of the examples brought by Kripke are all related to physical or chemical conditions (atomic numbers, DNA, molecular dynamics, etc.). Accordingly, we need to understand which kind of conditions fits well with consciousness, given that, as seen in the first chapter, consciousness is a multi-layered phenomenon, with qualitative, quantitative, objective and subjective features, and therefore it seems reasonable to consider more dimensions of investigation, if we want to hypothesize some probable necessary a posteriori statement of consciousness.

2.1.4 The Consciousness-Brain Relation

¹⁶ I shall explain later how this distinction can be misleading.

This last point leads us to a critical question, that is, which is the intrinsic structure of consciousness? Despite being a good starting point, saying that consciousness is being aware of something is still extremely vague, is not informative, and needs huge empirical and phenomenological improvement, in order to be a proper necessary a posteriori statement. Even if consciousness is a common phenomenon that characterizes our daily life, on the empirical and phenomenological ground we definitely know little and we cannot make any comparison with what we actually know about water, gold or tigers. The example of the human-like robot seems to imply that what is important about the investigation of consciousness is our intrinsic structure, e.g. our body. We already have in the world instances of artificial intelligences with human appearances (including facial expressions), like the android called Sophia, and there already are debates about whether assigning consciousness or not to them¹⁷. Thus, it seems that something about the intrinsic structure of the system that behaves like a conscious being is more important than the mere appearance of being conscious, in the same way as H₂O is the essential referent of water, rather than adjectives like “colourless”, “tasteless”, etc.

Now, intuitively one might infer that not the body in general, but the brain or, even better, some particular neural network is the rigid designator of consciousness, e.g. the thalamocortical system (Baars, Franklin, & Ramsoy, 2013; Gerald M. Edelman, Gally, & Baars, 2011; Gerald M Edelman & Tononi, 2000; Rosanova et al., 2012). However, we may consider first that the statement “consciousness = neural activity *x*” could be a so-called “contingent identity”, i.e. two terms differing in sense but identical in reference (we can imagine some possible world in which one side of the identity exists without the other)¹⁸. Kripke goes against the supposed validity of a

¹⁷ Interestingly, Sophia became the first AI application to receive citizenship of any country (Saudi Arabia) in October 2017 (Pagallo, 2018)

¹⁸ A clear example of contingent identity is made by Gibbard: There is a statue, Goliath, and the clay, Lumpl, from which it is composed. If Lumpl and Goliath coincide spatiotemporally, it is reasonable to say that they are identical. But they might not have been. «For suppose I had brought Lumpl into existence as Goliath, just as I actually did, but before the clay had a chance to dry, I squeezed it into a ball. At that point, according to the persistence criteria I have given, the statue Goliath would have ceased to exist, but the piece of clay Lumpl would still exist in a new shape. Hence Lumpl would not be Goliath, even though both existed» (Gibbard, 1975, p. 191). Thus, the identity of Lumpl and Goliath seems merely contingent.

contingent identity¹⁹, arguing that identity can be only necessary according to Leibniz's Law, i.e. if an object a has different properties than an object b, then a and b are not identical (Kripke, 1980, pp. 97–105). Kripke made an example with “pain = C-fiber stimulation”. This identity is supposed to work in the same way as “heat = molecular motion”, indeed all the terms here in play – “pain”, “C-fiber”, “heat” and “molecular motion” – are rigid designators, i.e. they refer to the same phenomenon in every possible world (Kripke, 1980, pp. 148–149). About heat and molecular motion, it seems impossible that the latter might have existed without being the former, rather molecular motion might exist without producing the accidental property of “sensation of heat”, understood as an intermediary between “heat” and “molecular motion”²⁰. So – I would say – molecular motion could exist without producing the sensation of heat, e.g. when part of my skin is anesthetized, and it is also possible being in the same epistemic situation of another person feeling heat, even in the absence of heat, e.g. when we get freezer burns. However, this is not the case for “pain” and C-fibers, because there is no accidental intermediate between the two, and being in the epistemic situation of having pain means only that I have a pain, i.e. the sensation of pain is pain as such (just consider a qualitatively identical situation to one in which you are in pain. Could that situation be such that you are not in pain?)²¹. Also, we can imagine a situation in which a particular pain sensation exists without the neural activity *x* and vice versa. Note that Kripke is not arguing for some ontological mind-body dualism between pain sensations and physical biological activities of the body, he is showing how “C-fibers” and “pain” cannot be a theoretical identity, i.e. a necessary identity,

¹⁹ «I felt little sympathy for the dark doctrine of a relation of ‘contingent identity’. Uniquely identifying properties can coincide contingently, but objects cannot be ‘contingently identical’» (Kripke, 1980, p. 4).

²⁰ Similarly, also “water” can be picked out either by its internal structure H₂O.

²¹ «In the case of the identity of heat with molecular motion the important consideration was that although 'heat' is a rigid designator, the reference of that designator was determined by an accidental property of the referent, namely the property of producing in us the sensation S. It is thus possible that a phenomenon should have been rigidly designated in the same way as a phenomenon of heat, with its reference also picked out by means of the sensation S, without that phenomenon being heat and therefore without its being molecular motion. Pain, on the other hand, is not picked out by one of its accidental properties; rather it is picked out by the property of being pain itself, by its immediate phenomenological quality. Thus pain, unlike heat, is not only rigidly designated by 'pain' but the reference of the designator is determined by an essential property of the referent» (Kripke, 1980, pp. 152–153). Interestingly, this is quite similar to what Sartre means with “consciousness (of) pleasure” but also works with pain: «Pleasure can not be distinguished – even logically – from consciousness of pleasure. Consciousness (of) pleasure is constitutive of the pleasure as the very mode of its own existence, as the material of which it is made, and not as a form which is imposed by a blow upon a hedonistic material» (Sartre, 1956, p. liv), see also above.

but rather a correlation, correspondence or co-occurrence. Thus, following Kripke, the C-fiber-pain is not an identity at all, while for others like Gibbard or Schwarz is a case of contingent identity (Gibbard, 1975; Schwarz, 2013).

Now, we may put into question the relation between “consciousness” and “brain activity x ”, in order to see whether the relation constitute a theoretical identity, as “heat = molecular motion”, or an empirical covariation/correlation or, at best, a contingent identity, as “pain” and “C-fibers”. So, we may introduce the terms “consciousness” and “brain activity x ” respectively for “heat” and “molecular motion”, for “pain” and “C-fibers”, and also for “water” and H_2O . Then, I would introduce the “what it is like to be”, the phenomenal character for consciousness, as analogue term for “sensation of heat”, “feeling of pain” and “colourless”. In this way, we now have four triads of terms: “heat-sensation of heat-molecular motion”, “pain-feeling of pain-C-fibers”, “water-colourless- H_2O ” and “consciousness-what it is like to be-brain activity x ”. We already seen that, according to Kripke the intermediate “sensation of heat” is an accidental property of heat-molecular motion and this is even more evident with the intermediate “colourless” in the “water- H_2O ” identity. On the other hand, pain can only be picked out by the feeling of pain, because being in “that” pain is not an accidental property of pain, rather an essential or intrinsic one. Then, regarding consciousness and the brain, given some particular brain activation for argument’s sake, now the issue would be whether the relation “consciousness-what it is like to be-brain activation” is more like “pain-feeling of pain-C-fiber” or “heat-sensation of heat-molecular motion” and “water-colourless- H_2O ”. In effect, I might state that since the feeling of pain is an instance of conscious state, then the relation with consciousness and that brain activity is contingent. Nevertheless, I shall test it in another way, first through the Cartesian and Zombie intuitions (Dietrich & Hardcastle, 2005), then analysing the nature of the “what it is like” within the consciousness triad.

First, what about the “consciousness-brain activity x ” relation? The Cartesian intuition refers to the Part IV of the *Discourse of the Method*, in which Descartes addresses the possibility that our experience can be what it is regardless of how the world is²², i.e. it is not necessary for our consciousness to cohere with how the physical

²² «Thus, because our senses sometimes deceive us, I wished to suppose that nothing is just as they cause us to imagine it to be ; and because there are men who deceive themselves in their reasoning and

world actually is (Dietrich & Hardcastle, 2005, p. 27), like in *The Matrix*. The Zombie intuition instead derives initially from a Robert Kirk's argument made in 1974, later developed by Chalmers' thought experiment about the so-called "philosophical zombie" (Kirk, 2019)²³; the intuition suggests that, in some possible world, there might be creatures perfectly identical to us who behave in a perfectly normal way, e.g. with some conscious-like input-output behaviour, but who completely lack phenomenal consciousness, i.e. qualia, so there is "nothing it is like" to be a zombie²⁴. To sum up, the two intuitions claim that consciousness may vary while the world remains the same (Zombie intuition), and that the world may vary while consciousness remains the same (Cartesian intuition) (Dietrich & Hardcastle, 2005, p. 28)²⁵. So, transposing the "mind-world" argument into our "consciousness-brain activity x" relation, it seems that we can conceive some possible world in which consciousness exists without that particular brain activation (Cartesian intuition) and, at the same time, we can conceive the activation of that neural network without the necessary existence of consciousness

fall into paralogsms, even concerning the simplest matters of geometry, and judging that I was as subject to error as was any other, I rejected as false all the reasons formerly accepted by me as demonstrations. And since all the same thoughts and conceptions which we have while awake may also come to us in sleep, without any of them being at that time true, I resolved to assume that everything that ever entered into my mind was no more true than the illusions of my dreams. But immediately afterwards I noticed that whilst I thus wished to think all things false, it was absolutely essential that the 'I' who thought this should be somewhat, and remarking that this truth 'I think, therefore I am' was so certain and so assured that all the most extravagant suppositions brought forward by the sceptics were incapable of shaking it, I came to the conclusion that I could receive it without scruple as the first principle of the Philosophy for which I was seeking. And then, examining attentively that which I was, I saw that I could conceive that I had no body, and that there was no world nor place where I might be; but yet that I could not for all that conceive that I was not. On the contrary, I saw from the very fact that I thought of doubting the truth of other things, it very evidently and certainly followed that I was; on the other hand if I had only ceased from thinking, even if all the rest of what I had ever imagined had really existed, I should have no reason for thinking that I had existed. From that I knew that I was a substance the whole essence or nature of which is to think, and that for its existence there is no need of any place, nor does it depend on any material thing» (René Descartes, 2003, p. 23).

²³ Actually, also this intuition derives from a Cartesian argument we can find in the VI *Metaphysical Meditation* (it is not the same, tough): « if I consider the body of a man as being a sort of machine so built up and composed of nerves, muscles, veins, blood and skin, that though there were no mind in it at all, it would not cease to have the same motions as at present, exception being made of those movements which are due to the direction of the will, and in consequence depend upon the mind (as opposed to those which operate by the disposition of its organs)» (René Descartes, 2003, p. 117)

²⁴ Apart from any speculative thought experiment like the philosophical zombie's one, also in some unresponsive wakefulness syndrome patients (UWS) we may see some typical conscious-like activity (e.g. eyes movements, postural movements, startle responses, etc.) correlated with some residual cortical activity similar to those found in healthy controls. Nevertheless, it is often difficult to assign consciousness to these cases, and sometimes they are diagnosed then as minimally conscious (MCS).

²⁵ Note that it is not my intention to use these as strong ontological arguments against any physicalist theory of consciousness, indeed there are some concerns about the validity of these intuitions, e.g. regarding the logico-ontological passage from "conceiving" zombie to the "possibility" of them (Kirk, 2019). I am using the two intuitions just to put into question the alleged necessary identity between consciousness and some particular brain activity.

(Zombie intuition). Therefore, the relation between “consciousness” and “brain activity x ” is not necessary. As I will expose later, this is far from saying that consciousness and brain are not related, rather we can say that we need to be careful about the distinction between necessary and contingent identities. For this reason, at this stage I would not claim that the essential structure of consciousness is some part of the brain, and consequently that consciousness does not rigidly refer to some brain activity. In the following chapters I shall add further arguments against the “consciousness-brain” necessary identity and in favour of other alternatives of contingent identity. In particular, I will expose in the last chapter how the brain-consciousness relationship can be conceived as a necessary a posteriori identity, that is, only by shifting the perspective from a brain-centered approach to a neuro-ecological approach, where the brain as part of a broader body-environment relational system is indirectly related to consciousness in a necessary way (Northoff, 2018b).

Secondly, what about the intermediate “what it is like to be”, the phenomenal character of consciousness, i.e. the fact that we are aware of what is presented in front of us (or also imagined by us)? Is this intermediate an accidental property like “sensation of heat” for “heat” or “colourless” for “water” or is it an intrinsic property like “feeling of pain”, whereby pain is rigidly designated by the feeling of “that” pain. I would argue that consciousness is not picked out by some accidental property, rather by the intrinsic property of “what it is like to be”. In other words, we cannot conceive the separability between being conscious and “what it is like to be”, i.e. the awareness of being in that state, in the same way we cannot conceive the separability of pain and the feeling of pain. So, differently from the sensation of heat that is contingent with regard to the heat-molecular motion relation, the feeling of being conscious/the act of being aware of the state/the what it is like to be is not contingent. Once removed the phenomenal character of consciousness, there is no consciousness at all.

Against this idea that consciousness and what it is like to be are not liable to the feeling of separability, one may argue that in our daily experience we are not perpetually aware of every detail and that there are also cases in which we may perceive something unconsciously, i.e. without conscious awareness, still having significant impact on later behaviour and thought, like in the unconscious perception cases (Prinz, 2015). So, one might say that in people with blindsight but also during

visual masking in healthy subjects perception and the “what it is like to perceive that” are not intrinsically related as consciousness and “what it is like to be”, since these are cases of unconscious perception, i.e. perception without conscious awareness of it. This is an interesting issue which – I believe – lies on misconceptions of consciousness and awareness. In order to argue for this, I need first to analyse deeper the concepts I am using here.

2.2 What am I aware of when I am aware? On Pre-reflective Self-awareness

So far, I have argued that we can probably accept the provisional definition of consciousness as “being aware of something”, however I have not said anything yet about the meaning of awareness used here. As we still were at the beginning of this research, I preferred not to load the term “awareness” with peculiar meanings from specific theoretical positions. At the moment, I intend the term in the same way as I made the examples here above, in a primitive way, as the what it is like to see a chair, to see a red apple, to feel a pain, etc., i.e. the phenomenal character of conscious states that seems to be fundamental for the concept of consciousness. Even contemporary reductionist as well as eliminativist positions – except for strong illusionists/Deniers – hold that, although there is nothing in reality like the phenomenal character of consciousness, still it seems to exist (e.g. Dennett, 1991, p. 372). At the same time, we must be careful not to oversimplify the meaning of awareness, for example as mere responsiveness to the environment.

In this respect, even people in Unresponsive Wakefulness Syndrome (UWS, also called Vegetative State) can show some degraded but still preserved level of responsiveness towards environmental stimuli, detectable through scales like the Glasgow Coma Scales (GCS) or the Recovery Coma Scale-Revisited (CRS-R). Both the scales present indices that give a numeric value to the visual and auditory response, the verbal and communication response, the motor and arousal functions, and these values can go from the total absence to a consistent preservation of the function. Let us take the arousal, and the auditory and visual functions, some patients in UWS may

open the eyes in response to stimulation, they may present visual startle, and fixation or, some of them may show auditory startle or localization to sounds, also they may breathe on their own, and they have an almost normal sleep-wake cycle.

However, it is questionable whether considering these reactions as real awareness of the environmental stimuli²⁶, understood as the basic intentionality of consciousness, where by “basic intentionality” I mean the directedness towards something being phenomenally presented to a subject (Nida-Rümelin, 2014). In other words, the UWS patients may behaviourally show responsiveness to the environment in an intentional-like way, still they probably lack the phenomenal character of consciousness because they do not know what it is like to receive that stimulus, in a similar way to the Zombie intuition²⁷.

These examples of responsiveness without awareness open the question about whether and how being conscious of something would require also being conscious of us as being conscious of something. It is almost intuitive thinking that, in order to determine a state as conscious, one must be conscious of his/her state: how can I be conscious of the notebook in front of me without, at the same time, being conscious of my state of being conscious of the notebook?²⁸. Starting from Brentano to the classical phenomenological tradition (Husserl, Merleau-Ponty, Sartre, Heidegger, Henry, Gurwitsch, among others) and contemporary authors like (Horgan, Kriegel, Zahavi, Gallagher, etc.), it has been defended the idea that it is impossible to undergo an experience without being aware of undergoing it, and this seems intrinsic to the

²⁶ Indeed, according to the CRS-R, in order to diagnose a Minimally Conscious State (MCS) or an Emergence from a Minimally Conscious State, the evaluation must obtain higher scores that resemble responses like reproducible or consistent movement to auditory command, visual pursuit, object localization or recognition (Schnakers, Vanhaudenhuyse, et al., 2009).

²⁷ «A patient in the vegetative state appears at times to be wakeful, with cycles of eye closure and eye opening resembling those of sleep and waking. However, close observation reveals no sign of awareness or of a ‘functioning mind’: specifically, there is no evidence that the patient can perceive the environment or his/her own body, communicate with others, or form intentions. As a rule, the patient can breathe spontaneously and has a stable circulation. The state may be a transient stage in the recovery from coma or it may persist until death» (Bates, 2003, p. 249).

²⁸ «[...]the necessary and sufficient condition for a knowing consciousness to be knowledge of its object, is that it be consciousness of itself as being that knowledge. This is a necessary condition, for if my consciousness were not consciousness of being consciousness of the table, it would then be consciousness of that table without consciousness of being so. In other words, it would be a consciousness ignorant of itself, an unconscious-which is absurd. This is a sufficient condition, for my being conscious of being conscious of that table suffices in fact for me to be conscious of it. That is of course not sufficient to permit me to affirm that this table exists in itself-but rather that it exists for me»(Sartre, 1956, p. lii).

concept of “being conscious of something”. Nevertheless, a relevant issue is to understand the relation between “being conscious of something” and “being conscious of being conscious of something, i.e. consciousness and self-consciousness. For example, Brentano argued that a mental state becomes conscious by taking itself as an object, so that in every single conscious experience two intentional objects are available to us, i.e. the notebook I am conscious of and the act of being conscious of the notebook²⁹. According to the higher-order theories of consciousness there are two different mental states, i.e. the consciousness of the notebook and the awareness of seeing it. Against this idea and following the phenomenologists, I claim that there are not two intentional objects of consciousness nor two mental states for the same object, but only the experience of one object, e.g. the notebook, while the type of self-consciousness entailed by the consciousness-of-something is not intentionally structured (subject-object relation-based). Therefore, when I am talking about self-consciousness I certainly not referring to any kind of high-order or meta-cognitive theory of consciousness that interprets self-consciousness (higher level) as an intentional and reflexive act of taking consciousness as an object (lower level) (Rosenthal, Armstrong, Lycan, Carruthers among others; see for details Zahavi, 2008, pp. 17–20). I would argue for a “one-level account of consciousness” (Zahavi, 2008b), where the self-consciousness here is not a new kind of consciousness, nor a property added to the experience, a new meta-mental state, a reflexive and monitoring act of awareness upon our own states, rather an intrinsic and integrated feature of consciousness itself³⁰. Otherwise, we may require another new meta-meta state of self-consciousness in order to be conscious of the previous conscious state and so on, *ad infinitum*, but this would be absurd, as Sartre clearly explains³¹. Thus, the act of being

²⁹ «In the same mental phenomenon in which the sound is present to our minds we simultaneously apprehend the mental phenomenon itself. What is more, we apprehend it in accordance with its dual nature insofar as it has the sound as content within it, and insofar as it has itself as content at the same time. We can say that the sound is the primary object of the act of hearing, and that the act of hearing itself is the secondary object» (Brentano, 1973, pp. 127–128).

³⁰ This is also near to the “same-order monitoring theory of consciousness”: «[...] a conscious state arises, on this view, when a mental event [...] and the subject’s awareness of it are integrated into a single unit through the relevant sort of cognitive process» (Kriegel, 2006, p. 150).

³¹ «It does not seem possible for us to accept this interpretation of the consciousness of consciousness. The reduction of consciousness to knowledge in fact involves our introducing into consciousness the subject-object dualism which is typical of knowledge. But if we accept the law of the knower-known dyad, then a third term will be necessary in order for the knower to become known in turn, and we will be faced with this dilemma: Either we stop at anyone term of the series-the known, the knower known, the knower known by the knower, etc. In this case the totality of the phenomenon falls into the unknown;

aware of our own conscious states, i.e. self-awareness, must be non-intentional, i.e. it does not posit any new object, and it must be non-reflexive, i.e. it does not involve a new cognitive reflection upon the conscious state, like when I deliberately think of my awareness.

Sartre called this non-voluntary (spontaneous), non-observational (it does not take the body as an object, nor the mind as an entity), non-inferential (it is not a belief, a representation or a judgment) and non-intentional act (it does not involve a subject-object relation) (Eshleman, 2016) the “non-positional consciousness of itself”, what contemporary phenomenologists and some analytic philosopher as well (e.g. Goldman, Strawson, Nida-Rümelin) call “pre-reflective self-awareness or self-consciousness” (Gallagher & Zahavi, 2016).

«Every positional consciousness of an object is at the same time a non-positional consciousness of itself. [...] it is the non-reflective consciousness which renders the reflection possible; there is a pre-reflective cogito which is the condition of the Cartesian cogito. [...] This self-consciousness we ought to consider not as a new consciousness, but as the only mode of existence which is possible for a consciousness of something» (Sartre, 1956, pp. liii–liv)³².

that is, we always bump up against a non-self-conscious reflection and a final term. Or else we affirm the necessity of an infinite regress (*idea ideae ideae*, etc.), which is absurd» (Sartre, 1956, p. lii).

³² «Consciousness of self is not dual. If we wish to avoid an infinite regress, there must be an immediate, non-cognitive relation of the self to itself. Furthermore the reflecting consciousness posits the consciousness reflected-on, as its object. [...] all that there is of intention in my actual consciousness is directed toward the outside, toward the world. In turn, this spontaneous consciousness of my perception is constitutive of my perceptive consciousness. In other words, every positional consciousness of an object is at the same time a non-positional consciousness of itself. If I count the cigarettes which are in that case, I have the impression of disclosing an objective property of this collection of cigarettes: they are a dozen. This property appears to my consciousness as a property existing in the world. It is very possible that I have no positional consciousness of counting them. Then I do not know myself as counting. Proof of this is that children who are capable of making an addition spontaneously can not explain subsequently how they set about it. [...] Yet at the moment when these cigarettes are revealed to me as a dozen, I have a non-thetic consciousness of my adding activity. If anyone questioned me, indeed, if anyone should ask, “What are you doing there?” I should reply at once, “I am counting.” This reply aims not only at the instantaneous consciousness which I can achieve by reflection but at those fleeting consciousnesses which have passed without being reflected-on, those which are forever not-reflected-on in my immediate past» (Sartre, 1956, pp. lii–liii) See also J. P. Sartre, 2004b, pp. 4–6: «The existence of consciousness, indeed, is an absolute, because consciousness is conscious of itself; in other words, the type of existence that consciousness has is that it is consciousness of itself. And it becomes conscious of itself insofar as it is consciousness of a transcendent object. [...] We need to add that this consciousness of consciousness—with the exception of cases of reflective consciousness, which I will be examining in detail later—is not positional, i.e. consciousness is not its own object. Its object is outside itself by nature, and this is the reason why, in one and the same act, consciousness can posit and grasp its object. Consciousness as such knows itself only as absolute inwardness. I will call such a

In the words of Alvin Goldman:

«[Consider] the case of thinking about x or attending to x. In the process of thinking about x there is already an implicit awareness that one is thinking about x. There is no need for reflection here, for taking a step back from thinking about x in order to examine it...When we are thinking about x, the mind is focused on x, not on our thinking of x. Nevertheless, the process of thinking about x carries with it a non-reflective self-awareness» (Goldman, 1970, p. 96).

The concept of non-reflective self-awareness is in line and fits well with some of the notions I exposed in the previous chapter, e.g. the subjective/phenomenal character, for-me-ness, what-it-is-likeness, mineness, so that if a mental event is not pre-reflectively self-conscious there is nothing it is like to undergo that event (Zahavi, 2008b).

Thus, we reached from a phenomenological side the same result we had through the logical and metaphysical side through the direct reference theory, that is, the fact that “consciousness of something” and the phenomenal character of experience are intrinsically related, since the “what it is like to be” is a constitutive element of an intentional conscious act, without being in turn an intentional object of consciousness. Indeed, I would argue that the phenomenal character of consciousness is deeply related with the pre-reflective self-awareness structure of consciousness, since that every “what it is like to be conscious of something” is necessarily a “what it is like to be *for me*”³³, and this without involving any representational or high-order self-reflection upon the conscious state that explicitly takes the self as an object, e.g. mirror self-recognition, conceptual or narrative self-consciousness (I do not need to tell you that you are reading this paragraph in this very moment, and now that I told you do not need to check in order to confirm, saying something quite odd like “someone is reading this sentence, is it I”?). It is now clear that when I speak of “phenomenal” character I am not only referring to some qualitative property of the experience or to the first-person report of what means to feel something like this, but I am taking a more

consciousness ‘first order’ or ‘unreflective’ consciousness [...] In fact, all reflecting consciousness is in itself unreflected, and a new, third-order act is needed to posit it. Moreover, there is no infinite regress here, since a consciousness has no need of a reflecting consciousness in order to be conscious of itself. It merely does not posit itself to itself as its own object».

³³ This is already claimed by Nagel (see Nagel, 1974, p. 440n, and see above § 1.2).

“phenomenological” characterization of this concept³⁴, relating it with the fundamental structure of consciousness, i.e. intentionality, self-awareness, and – as we will see in the next chapters – time-consciousness, body-consciousness and the notion of being-in-the-world.

So far, we have seen that the term “consciousness” directly refers to intentional acts directed to something, in the terms of “being conscious of something”, and that this “consciousness-of-something” is not a general concept but is fulfilled from time to time by instances of experiences (consciousness of a chair, consciousness of an old memory, etc.), and that these instances possess an intrinsic element that we may broadly call phenomenal character, i.e. the fact we feel something to be in a particular conscious state or in another, and this feeling is not another object of consciousness, rather a “pre-reflective self-awareness of being conscious of something”. Without this – I would call it now – phenomenological character, there would be nothing like to experience something, that is, there would be no consciousness at all, as I showed, for example, in the intrinsic relation between “pain” and sensation of pain. Taking together, these are the phenomenological features of consciousness: phenomenality, i.e. the fact that there is something like to be conscious; intentionality, i.e. the fact that to be conscious is always to be conscious of something; self-awareness, i.e. the fact that any conscious state of something intrinsically involves a form of awareness of itself as being conscious (not necessarily a reflective, cognitive, representational or high-order self-consciousness).

2.3 Regarding the possibility of Unconscious Experiences

As mentioned above, one might argue that consciousness seems not to be always related with pre-reflective awareness or with the what-it-is-likeness of experience, rather we have some cases – normal and pathological – in which it seems

³⁴ «When speaking of phenomenology the vast majority of analytical philosophers are simply referring to a first-person description of what the ‘what it is like’ of experience is really like. Phenomenology is in other words identified with some kind of introspectionism. But for anybody familiar with Continental philosophy, this notion of phenomenology will appear as both tame and lame» (Zahavi, 2003).

possible to have unconscious experiences, i.e. being intentionally related to an object without being aware of it. Basically, we can develop this issue through three conceptual and empirical cases, the blindsight case, the long-distance truck driver and the refrigerator light problem³⁵.

2.3.1 The Refrigerator Light Problem and the Long-distance Truck Driver

The refrigerator light problem addresses the question about how much our conscious experience is pervasive in our waking life. If I reflect on my conscious state in this very moment, I obviously find myself as aware, however this does not help us to understand if I am truly aware when I am not intentionally focusing on my own conscious states. It would be like assuming that the light in the refrigerator is always on because every time I open the door the light is on. Similarly, I cannot assume that I am constantly aware in every single moment of my waking life merely because when I introspect on my experience, I find myself aware. Hence, this issue is also called the refrigerator light “illusion” or “fallacy”, it seems that our conscious experience is pervasive, but it is likely that it is not. If this is true, claiming that every consciousness of something must be at the same time non-intentionally and pre-reflectively aware of itself could be victim of the refrigerator light fallacy. It seems that our conscious states are always awareness of ourselves as being conscious of something, only because we introspectively reflect upon those states in those moments, as well as we see the light on whenever we open the refrigerator door, but we cannot know anything about all the other times when the door is closed, that is, metaphors aside, we cannot know whether we are pre-reflectively self-aware when we are not focusing on our own conscious

³⁵ Not only philosophy, but also empirical psychology has dealt with the so-called behavioural automatisms, i.e. those cases (more frequent than one might think) in which our actions are not accompanied by constant conscious monitoring. Indeed, these actions are first prepared and performed unconsciously, and the subject becomes conscious only later, after the brain activity and bodily skills have already begun (Bargh & Morsella, 2008; Dijksterhuis, 2010; Kihlstrom, 1987). In particular, see Libet’s pioneering experiments that challenged the concept of free will, showing the unconscious cerebral activity (readiness potentials through EEG) that precedes conscious decisions (Libet, 2004; Libet et al., 1983). Recently, these experiments have been replicated and refined with new techniques (e.g. fMRI) by Haynes and colleagues (Soon et al., 2008, 2013). For a general review and critique of the topic see Lavazza, 2016, 2019; Lavazza & De Caro, 2010.

states, i.e. when our “mind door” is closed. However, I shall argue that the structure of pre-reflective self-awareness is not undermined by the refrigerator light fallacy.

First, we need to distinguish between at least two different approaches to the issue. According to the rich version, consciousness seems pervasive because it actually is constant, through a complex flow of experience, but we often cannot access or report this consciousness. On the other hand, the thin version claims that consciousness seems pervasive, but this is just an illusion created by our introspection³⁶. Thus, if the rich version is true, what is the role of pre-reflective self-awareness, how and when does it work in our flux of experiences? Instead, if the thin version is true, what is the point of having pre-reflective self-awareness?

In order to develop better this point, we could add another issue related to the refrigerator light illusion, the paradigmatic case of the “long-distance truck driver”, invented by David Armstrong in 1981:

«After driving for long periods of time, particularly at night, it is possible to “come to” and realize that for some time past one has been driving without being aware of what one has been doing. The coming-to is an alarming experience. It is natural to describe what went on before one came to by saying that during that time one lacked consciousness. Yet it seems clear that, in the two senses of the word that we have so far isolated, consciousness was present. There was mental activity, and as part of that mental activity, there was perception. That is to say,

³⁶ «Do you have constant tactile experience (or ‘consciousness’ or ‘phenomenology’) of your feet in your shoes? Constant auditory experience of the hum of traffic in the background? Constant visual experience of the frame of your eyeglasses? Or, when you aren’t attending to such matters, do they drop out of consciousness, so that they’re in no way part of your stream of experience, no part of ‘what it’s like to be you’? We might think of consciousness as like a soup. Is it a rich soup, full of experience in a wide variety of modalities simultaneously — visual, auditory, tactile, olfactory, imagistic, proprioceptive, emotional — or is it a thin soup, limited to one or a few things at a time? [...] Ordinary people’s intuitions differ. Researchers disagree. James (1890/1981) and Searle (1992) endorse the rich view, according to which the stream of experience involves both a centre of attention and a broad periphery of consciously experienced but unattended objects and background feelings. Jaynes (1976), Dennett (1991), and Mack and Rock (1998) endorse the thin view: Consciousness is limited to only one or a few objects, modalities, topics, or fields at a time. The unattended hum of the traffic in the background is no part, not even a peripheral part, of your experience when you’re sufficiently absorbed in other things» (Schwitzgebel, 2007, p. 6). There would also be the “maximalist/minimalist” version for the refrigerator light problem, where the former claims that the contents of consciousness are rich, while the latter claims that the contents of consciousness are sparse (Bayne, Cleeremans, & Wilken, 2009, pp. 561–562). Despite the two versions can be used as synonyms, I prefer the rich/thin version because it is more focused on the presence/pervasiveness versus absence/rarity of self-awareness in consciousness, where the maximalist/minimalist version seems more related to the contents of consciousness.

there was minimal consciousness and perceptual consciousness. If there is an inclination to doubt this, then consider the extraordinary sophistication of the activities successfully undertaken during the period of “unconsciousness.” [...] Above all, how is it possible to drive a car for kilometers along a road if one cannot perceive that road? One must be able to see where one is going, in order to adjust appropriately. It would have to be admitted, at the very least, that in such a case, eyes and brain have to be stimulated in just the same way as they are in ordinary cases of perception. Why then deny that perception takes place? So it seems that minimal consciousness and perceptual consciousness are present. But something else is lacking: consciousness in the most interesting sense of the word»³⁷ (D. Armstrong, 1997, p. 721).

The case is quite clear, and we can imagine lots of similar situations of our daily life during which we seem not really aware of ourselves while doing something in an automatic way (“did I locked the door?”). Still it seems absurd to say that in those moments we are not conscious of what we are doing and perceiving. Are those cases of positional consciousness without pre-reflective self-awareness? Is the truck-driver “conscious of the road” in an automatic and – paradoxically – “unconscious” way? This would strongly undermine the claim that every consciousness-of-something needs to be at the same time a non-intentional consciousness of myself as conscious of something, i.e. pre-reflective self-awareness (Howell & Thompson, 2017; Schear, 2009; for details, see Zahavi, 2019). Hence, if this is true, we might claim that we are self-aware only whenever we introspectively reflect upon our own conscious state (I am driving and, at some point, I find myself introspecting my own mental states), i.e. when I open the refrigerator/mind door.

³⁷ The case is also called the absent-minded driver, see for example (Schwitzgebel, 2007, p. 6,8): «Suppose you’ve driven the same route to work a thousand times. Today, you’re deeply absorbed in thinking about an unpleasant interaction with your department head. Traffic is light; no dangerous situation occurs; you drive habitually. You arrive at the parking lot and seem to ‘wake up’ — ah, I’m at work already! — with virtually no explicit memory of having driven there. Now consider this: Did you have visual experience, visual consciousness, while you were driving, or not? You responded to events on the road: You stopped at the red light, you stayed in your lane. Visual input obviously had some regulatory influence on your behaviour. But maybe visual input can influence behaviour without the involvement of consciousness. [...] In absent-minded driving, I’m clearly ‘aware’ of the road in the sense of being responsive to stop lights and turns. In some sense, I’m also ‘aware’ of subliminal stimuli if they influence my behaviour in the right kind of way. To equate consciousness with this purely response-dependent sense of ‘awareness’ misses the central issue».

In order to solve this problem, I shall turn back again to Sartre's analysis of pre-reflective self-awareness. In *The Transcendence of the Ego* (1936), Sartre's main argument was against the idea that every intentional consciousness is founded or structured by the Ego, expressed by the active ("I") or passive ("me") way. He explained this point through some instances of acts of consciousness in which there is no presence of the Ego, e.g. the act of reaching a departing tram:

«[...] since all the non-reflective memories of unreflected consciousness show me a consciousness without me, and since on the other hand theoretical considerations based recognize that the I could not be part of the internal structure on consciousness's intuition of essence have obliged us to of *Erlebnisse*, we are forced to conclude: there is no I on the unreflected level. When I run after a tram, when I look at the time, when I become absorbed in the contemplation of a portrait, there is no I. There is a consciousness of the *tram-needing-to-be-caught*, etc., and a non-positional consciousness of consciousness. In fact, I am then plunged into the world of objects, it is they which constitute the unity of my consciousnesses, which present themselves with values, attractive and repulsive values, but as for me, I have disappeared, I have annihilated myself. There is no place for me at this level, and this is not the result of some chance, some momentary failure of attention: it stems from the very structure of consciousness» (Sartre, 2004b, pp. 7–8).

So, when I am running towards the departing tram I am not accessing to my own consciousness in a way like "I am running towards the tram", rather there is only consciousness of the tram that must be reached, and I may not even remember which obstacles I avoided and whether there were other people along the sidewalk³⁸. In reference to the truck-driver case, during the long and exhausting drive, there are some moments when there is only "consciousness of the point of delivery-needing-to-be-reached", without any positional consciousness of which roads and turns the driver took. Alternatively, regarding the example made by Schwitzgebel, while you are on our way to work, thinking about how to deal, step by step, with that particular task you have to manage by end of day, there is not a state like "I am conscious of me as the

³⁸ «When I reflect on my experience of running to catch up with a bus, I appear to myself as an agent (I choose to run), as being in a certain state (I desire to catch the bus), and as possessing a certain quality (I am late "again"). But Sartre denies that such reflection captures experience as it is lived, where, in contrast, there is only "consciousness of the streetcar-as-having-to-be-overtaken"» (Crowell, 2012, pp. 203–204).

person that needs to deal with that task” but there is only “consciousness of an errand-to-run”. Also, once we get out of the car, we may not easily recollect how the trip was and what happened along the way.

To sum up, when I am absorbed or focused into some intentional object that completely captures my attention or, on the other side, when I am doing something that I perfectly know how to perform it, there seems to be no presence of myself as a subject that does/perceives/thinks this things and there seems to be no presence of self-awareness. Sartre says in *The Transcendence of the Ego* that the «transcendental field of this unreflected consciousness» is «impersonal, or, if you prefer, ‘prepersonal’, it is without an I» (Sartre, 2004b, p. 3), the things of the world made the person, i.e. are the conditions for the emergence of an “I”, and not vice versa³⁹. The “I” is a product of the reflection and therefore it counts as another transcendent object of consciousness – useful, necessary, etc. but still a consequence of the conscious act, not the foundation of it. This could be called the “anonymity objection”: «[a]ccording to what might be termed the *anonymity objection*, there is no for-me-ness or mineness on the pre-reflective level, nor is there any I or self» (Zahavi, 2014, p. 26; Schear, 2009).

This seems in line with the thin version of the refrigerator light problem, since Sartre seems to be saying that only when we introspectively reflect upon ourselves we are conscious of a person that is the subject of a collection of conscious acts, in the same way as the light is on only when we open the refrigerator door. However, that is not the case. How can this go with the pre-reflective self-awareness? First, saying that when we are not conscious of ourselves as an “I” or “me” when we are immersed in our conscious life does not mean that we are not aware of ourselves at all. Indeed, after having driven for a long time, despite not being able to recollect what we actually did, we do not need to ask something like “someone drove all the night, was it me?”. So, from the fact that sometimes we can do something while conscious but in an apparent state of unawareness, we cannot infer that we are not aware at all of ourselves. Secondly, the fact that our consciousness can possess degrees of anonymity, spontaneity and impersonality does not mean that there is no room for self-

³⁹ «The ‘I think’ can accompany our representations because it appears against the background of a unity that it has not contributed to creating, and it is this pre-existing unity which, on the contrary, makes it possible» (Sartre, 2004b, p. 3).

consciousness, first-person givenness or for-me-ness⁴⁰. Despite the absence of a robust “I-subject”, i.e. the Ego as pivot and collector of all the lived experiences, still all consciousnesses are structured by a presence to itself, i.e. for-me-ness, selfhood or “selfness” with the words of Sartre⁴¹. This condition of self-consciousness (what Sartre defines as “*conscience de soi*”, the pre-reflective cogito) is not necessarily and automatically manifested through every single moment of our daily life, nor, on the other hand, is a cognitive reflection upon our own states of consciousness, rather it is the non-intentional and non-objectifying condition for the subsequent high-order self-consciousness (what instead Sartre defines as “*connaissance de soi*”, the Cartesian cogito) (Sartre, 1948). Without this precondition – necessary non-sufficient – involved in any world-immersed consciousness, we would not be able to understand how we can reflect upon our own conscious state in any moment:

«[...] from the fact that pre-reflective self-consciousness isn’t sufficient for first-person knowledge, one can obviously not conclude that it is therefore also unnecessary if such knowledge is to obtain. The former provides for an experiential grounding of any subsequent self-ascription, reflective appropriation, and thematic self-identification, and is to that extent epistemically enabling. To rephrase, any account of first-person knowledge that

⁴⁰ The entire absence of self-consciousness is sustained by the phenomenal externalism as exposed by Zahavi: «We can deny that phenomenal consciousness entails self-consciousness and endorse a radical version of the transparency thesis. We can then argue that phenomenality is strictly and exclusively world-presenting; that is, when consciously scrutinizing the lotto coupon or when consciously admiring a portrait, we are de facto self- and mind-blind» (Zahavi, 2014, p. 37). Phenomenal externalism is sustained by Fred Dretske and Michael Tye among others, and, as I shall argue, is partly right regarding the nature of phenomenal features of consciousness.

⁴¹ «Yet we need not conclude that the for-itself is a pure and simple “impersonal” contemplation. But the Ego is far from being the personalizing pole of a consciousness which without it would remain in the impersonal stage; on the contrary, it is consciousness in its fundamental selfness which under certain conditions allows the appearance of the Ego as the transcendent phenomenon of that selfness» (Sartre, 1956, p. 103). «But even Sartre himself came to see the distinction as being too crude. Whereas Sartre in his 1936 work had characterized non-egological consciousness as impersonal, he described this view as mistaken in both *L’Être et le néant* and his important article ‘*Conscience de soi et connaissance de soi*’, from 1948. In both cases, Sartre’s crucial move is to distinguish self and ego (Sartre 2003: 263). Although no ego exists on the pre-reflective level, consciousness remains characterized by a basic dimension of selfhood precisely because of its ubiquitous self-consciousness. This is why Sartre could write, ‘pre-reflective consciousness is self-consciousness. It is this same notion of self which must be studied, for it defines the very being of consciousness’ (Sartre 2003: 100). Thus, rather than defining self-consciousness on the basis of a preconceived notion of the self, Sartre’s idea was to let our notion of the self arise out of a correct understanding of self-consciousness. It is certainly true that pre-reflective self-consciousness shouldn’t be understood as consciousness of a separate and distinct self. But pre-reflective self-consciousness is ineliminably first-personal, that is, characterized by for-me-ness. And that is all that is needed in order to warrant the notion of an experiential self» (Zahavi, 2014, p. 48). For a neo-Sartrean position about selfhood and pre-reflective self-awareness, see Zahavi, 2008.

denies or ignores the first-personal character of our experiential life is a non-starter. We need to explain how I can appropriate an experience as my experience. Had the experience been completely anonymous and impersonal when originally lived through, had it lacked first-personal givenness altogether, such an appropriation would be rather inexplicable. This is also why the obvious reply to the refrigerator objection is that it leaves it quite mysterious how our reflective gaze or monitoring stance is supposed to have that kind of effect, that is, offer that kind of illumination» (Zahavi, 2014, p. 37).

So, if pre-reflective self-awareness is always involved in every positional consciousness as a precondition for any other self-knowledge of our own conscious states, how do we recognise its fundamental presence without confounding it with introspection, meta-cognition, high-order representation, etc.? I shall outline some interesting phenomenological cases. The first concerns body awareness. In this very moment you are reading this sentence, but now I ask you to tell me whether you feel the shoes on your feet (or the clothes on your back). The answer is obviously yes, but, again, this is not a solution for the refrigerator fallacy, because the main point to understand is whether you were aware of your shoes before I asked to check. However, we may find here the first hint about the pervasiveness of the pre-reflective self-awareness due to the fact that, when you checked your feet, you did not need to actually “re-flect” upon your body state, rather the apprehension is immediate and natural, like it was already there waiting for a cognitive update. The experience seems not to be “created” by the introspection but emerges as already cut-and-dried.

2.3.1.1. Bodily Awareness as Pre-reflective Self-consciousness

In order to demonstrate this point, we would need a situation in which pre-reflective self-awareness is disrupted or weakened, and thus I shall now consider the case of some people who do not possess the background awareness of their own body. Before discussing these cases, I need to briefly explain the difference, brilliantly defined by Shaun Gallagher (Gallagher, 1986), between “body image” and “body schema”, in order to understand what we are referring to here. The body image

«consists of a complex set of intentional states – perceptions, mental representations, beliefs, and attitudes – in which the intentional object of such states is one's own body. Thus the body image involves a reflective intentionality» (Gallagher & Cole, 1995, p. 371)⁴². Instead, the body schema «involves a system of motor capacities, abilities, and habits that enable movement and the maintenance of posture», so contrarily to the body image, it does not consist of perceptions, conscious representations, beliefs, or attitudes, rather it operates «below the level of self-referential intentionality, although such functions can enter into and support intentional activity. The preconscious, subpersonal processes carried out by the body schema system are tacitly keyed into the environment and play a dynamic role in governing posture and movement» (Gallagher & Cole, 1995, p. 371), also it can product specific effects on cognitive experience (Gallagher, 1995).

According to this conceptual and phenomenological distinction, the ability to immediately refer to the state of our body parts, like the feeling of your feet in shoes exposed above, depends pre-intentionally to the body schema that gives us a constant but covered, marginal feedback about our position in the environment, without positing the body as an object of consciousness⁴³. Only then and through the body schema, we are able to intentionally relate with our body parts, creating a conscious representation of our body parts and their relations with other objects, e.g. clothes, the chair, various objects around us, other people, etc. We thus may define the body schema as a form of the pre-reflective self-awareness. What happens when we partially lose the body schema?

Under normal conditions, we constantly receive information from the environment and our internal states through the body – skin, muscles, tendons, organs,

⁴² «On the pre-reflective level, when consciousness is engaged in the world, the body is not explicitly present to consciousness. The body does not appear (the body image does not manifest itself) until there is a forced reflection brought on by pain, discomfort, fatigue. However, when the body is "in tune" with the environment, when events are smoothly ordered, when the subject is engaged in a task that holds the attention of consciousness, then the body is anonymously performing with its environment on a behavioral level that escapes consciousness» (Gallagher, 1986, p. 549).

⁴³ «The body schema, understood in this way, is not the perception of "my" body; it is not the image, the representation, or even the marginal consciousness of the body. Rather, it is precisely the style that organizes the body as it functions in communion with its environment. It is not the object of consciousness, nor the result of a conscious project; it is the non-conscious, pre-reflective performance of the body as it lives its environment. The body performs in rapport with its environment for the most part without a perceptual image or awareness of the body being produced, even in a marginal consciousness» (Gallagher, 1986, pp. 549–560).

etc. However, in rare cases, some neurological pathologies lead to the selective de-afferentation of the body, e.g. to the loss of touch and proprioception, depriving the haptic system of peripherally originating feedback (J. D. Cole & Paillard, 1995)⁴⁴. Proprioception refers to non-visual input that tells us where our body is in space, and plays an important function in normal motor control and motor learning as well as in apprehension of internal and external stimuli (Yousif et al., 2015). In other words, without proprioception we would not be able to perceive where our body is located in space and what it is in contact with, and also to perform lots of ordinary movements, except when we look at the body, focusing on the movements and positions of each part of it⁴⁵. In this respect, people affected by this kind of de-afferentation are initially overwhelmed by loss of all coordinated movement of their limbs and trunk, but with training, mental concentration and visual supervision, useful and near normal function for most motor tasks can be recovered (Yousif et al., 2015)⁴⁶. In terms of the

⁴⁴ I say “selective” because it depends on which fibres are affected by the neuropathology. In the cases I am going to discuss, only the large fibres has been affected (impairing touch and proprioception), while small fibres has been spared, leaving perception of pain, heat and cold intact (J. D. Cole & Paillard, 1995).

⁴⁵ «Although the control and coordination of movements are normally accomplished by a body schema without the help of conscious attention, and in some cases even without conscious awareness of what our body is doing, there are instances in which the perception of one's own movements can be complexly interrelated to their accomplishment. Conscious perception of my own body can be used to monitor and control my posture and movements. The visual, tactile, and proprioceptive awareness that I have of my body may help me to learn a new dance step, improve my tennis game, or imitate the movements of others. Ordinarily, however, in walking I do not have to think about putting one foot in front of another; I do not have to think through the action of reaching for something. I may be marginally aware that I am moving in certain ways, but it is not usually the center of my attention. And marginal awareness may not capture the whole movement. If I am marginally aware that I am reaching for something, I may not be aware at all of the fact that for balance my left leg has stretched in a certain way, or that my toes have curled against the floor. Posture and the majority of bodily movements usually operate without the help of a body image» (Gallagher & Cole, 1995).

⁴⁶ A paradigmatic example of this cognitive supervision upon normal movements is the act of walking, where people de-afferented in proprioception and touch must follow with vision every single step, focusing on the coordination of feet and legs. They cannot talk too much or let the mind wander, otherwise they may lose control of their body. This is one these cases, described by Gallagher and Cole: «At the onset of the neuropathy IW's initial experience was of complete loss of control of posture and movement. He could not sit up or move his limbs in any controllable way. For the first three months, even with a visual perception of the location of his limbs, he could not control his movement. In the course of the following two years, while in a rehabilitation hospital, he gained sufficient motor control to feed himself, write, and walk. He went on to master everyday motor tasks of personal care, housekeeping, and those movements required to work in an office setting. To maintain his posture and to control his movement IW must not only keep parts of his body in his visual field, but also conceptualize postures and movements. Without proprioceptive and tactile information he neither knows where his limbs are nor controls his posture unless he looks at and thinks about his body. Maintaining posture is, for him, an activity rather than an automatic process. His movement requires constant visual and mental concentration. In darkness he is unable to control movement; when he walks he cannot daydream, but must concentrate constantly on his movement. When he writes he has to concentrate on holding the pen and on his body posture. IW learned through trial

conceptual distinction outlined, this people show an impairment in the body schema and need to fill the gaps in their body awareness through an overloaded usage of the body image (J. D. Cole & Paillard, 1995; Gallagher & Cole, 1995)⁴⁷. Accordingly, this is what happens when our pre-reflective, non-intentional consciousness of our own body is partly damaged or disrupted⁴⁸.

By the way, someone might say that this only applies to pre-reflective awareness towards the body, not to the environment, nevertheless I can make a case also for external stimuli⁴⁹. Imagine that you are in a silent library, concentrating on these words, when suddenly the air conditioner turns off. At that moment, you immediately notice the lack of that hum, which you did not even consciously notice before, but which you probably have encoded in the background of your conscious experience. The example is interesting, because this is not a new sound appearing (in

and error the amount of force needed to pick up and hold an egg without breaking it. If his attention is directed toward a different task while holding an egg, his hand crushes the egg» (Gallagher & Cole, 1995, pp. 374–375).

⁴⁷ By the way, just like pre-reflective self-consciousness influences positional and reflective consciousness, in the same way the disruption of the body schema carries consequences also for the body image: Just as the loss of proprioceptive information impairs the body schema and disrupts the control of posture and movement, so the loss of proprioceptive awareness results in an impoverishment of the perceptual aspect of the body image. Although IW still had a visual perception of his body, this was not enough at first to gain control over his posture or movement. Because of the absence of proprioception, another important part of the body-schema system, a capacity for intermodal communication, failed. The normal intermodal communication between proprioception and vision had been disrupted. [...] In IW, the language of proprioception was missing, and he had to learn to use vision (constantly updating information on present posture and movement), as a partial substitute for proprioception, to drive motor processes directly» (Gallagher & Cole, 1995, p. 380).

⁴⁸ This does not mean that proprioception represents the only condition for pre-reflective self-awareness, otherwise it would be difficult to explain how patients without proprioception still can be positionally conscious towards external objects and their own body: «Christina's loss of proprioception [a case of selective de-afferentation illustrated in Sacks, 1985] diminishes her capacity to perceive and engage with her environment under body-relative modes of presentation, but it obviously does not undermine this capacity altogether. This stands to reason: although proprioception is a vitally important dimension of our bodily awareness, it does not exhaust such awareness. Christina retains the ability to perceive things as on her left, on her right, above her, below her, etc. simply on the basis of vision, which presents the world around her in an egocentric framework that has her body at its origin. She also retains the ability to initiate voluntary movements whose aims must be characterized in body-relative terms: stepping forward, reaching for the glass over there, etc. Her loss of proprioception diminishes her ability to execute these movements successfully, but her ability to represent her environment in such terms is supported by a lifetime of perceptual and practical engagement with the world around her, and does not simply disappear when somatic feedback is cut off. That she remains able to engage with her environment in these ways does not show that this ability is independent of bodily awareness; it simply shows that this awareness is more robust, and more diverse in its sources, than we might at first suppose» (M. Boyle, 2018).

⁴⁹ Even though I am now using this “external/internal” distinction, I believe that it can be useful only for an explanatory purpose, and it does not refer to any strong ontological distinction, rather it may produce in some cases an implausible clear-cut separation between body and environment. I will try to clarify better this point later.

that case, it would be merely a new positional consciousness of the sound), but an already present stimulus that abruptly disappears. How could you notice that the air conditioner turns off, if you did not possess some degree of pre-reflective consciousness of you being surrounded by that monotonous humming? You did not merely notice a vague perceptual difference between before and after, but now you are positionally conscious of the absence of that particular sound, this means that prior to being intentionally aware of the hum, you need to be non-intentionally and pre-reflectively aware of yourself as surrounded by that hum. The fact that you did not “hear” the sound for the first thirty minutes you were in the library does not mean that you were not aware of it, at least non-intentionally.

Now, it should be clearer how pre-reflective self-awareness is an essential condition for any intentional consciousness, e.g. taking a departing tram, going to work, etc., and this brings me back to the long-distance truck driver. The alternatives seem to be three: the driver is unconscious while driving, the driver is conscious of the road but not self-conscious, the driver is conscious of whatever he is thinking and pre-reflectively conscious of himself driving the truck. As Armstrong himself said, it seems inappropriate to say that the truck driver is unconscious, like we would say for a sleep-walker, of course there may exist specific habits or gestures we perform unconsciously, but driving a truck along the road is not likely to be one of these, so the driver must be conscious somehow. Then, the driver might be conscious of the road but not conscious of himself driving the truck, this would be also a good answer for the thin version of the refrigerator light problem, because it means that our conscious experience is not rich but sparse. However, this possibility is problematic, since it confuses the difficulty in reporting with the absence of conscious experiences. As well as the fact that we are not constantly aware in an intentional way of our feet on shoes or of the humming of the air conditioner does not mean that we are completely unconscious of them, in the same way as the fact that the truck driver, at some point, “comes to” after a long period in which he apparently has been unaware of what he has been doing does not mean that he was completely unaware of it. Probably, asking him what he did at some specific point of the road will help him in reporting the details of the journey.

Similar techniques are used in contemporary phenomenology for inducing the subjects to bring out details from experience, e.g. the Explication Interview (see chapter 1) and the Descriptive Experience Sampling methodology (DES) (Heavey, Hurlburt, & Lefforge, 2010; R. Hurlburt & Heavey, 2004; R. T. Hurlburt & Akhter, 2006). Regarding the latter technique, «DES subjects carry a random beeper in natural environments; when the beep sounds, they capture their inner experience, jot down notes about it, and report it to an investigator in a subsequent expositional interview. DES is a fundamentally idiographic method, describing faithfully the pristine inner experiences of persons» (R. T. Hurlburt & Akhter, 2006, p. 271). The beeper works kind of like the air conditioner that turns off in the library, leading the subject to identify experiential samples of what they were living just before the input, through some marginal, peripheral or pre-reflective consciousness. Of course, we do not have to consider these descriptive samples as pure and untouched instances of consciousness, since every time we come back to our own experience, we transform it. However, the point is another: if our experience was really sparse, limited to only one or a few objects, modalities, topics, or fields at a time, the immediate body-awareness, the air conditioner case and the Descriptive Experience Samples should be extremely hard to explain. Therefore, I would rather accept the third option, i.e. the truck driver is positionally conscious of whatever he is mind-wandering (the dinner after work, the coveted point of delivery, his last discussion with his boss, etc.) while having at the same time pre-reflective consciousness of himself driving along the road, as a sort of background that never goes away⁵⁰, but that can immediately emerge as a new positional consciousness, e.g. a vehicle drifts into the lane⁵¹.

⁵⁰ «[...] conscious perception in the absence of attention. A typical example of this is 'gist perception' exemplified by the absent-minded truck-driver who does not attend to the road in front of him but nevertheless manages to drive safely» (Schlicht, 2012, p. 326).

⁵¹ Of course, Sartre did not have the opportunity to read and address these issues, for obvious chronological reasons, however someone tried to analyse whether his concept of pre-reflective consciousness could resist the criticism of the long-distance truck driver. Kathleen Wider argues that Sartre is unable to give a clear answer to Armstrong's example of the long-distance truck driver, because he would have no choice but to claim the driver is conscious in a relevant way (and therefore also pre-reflectively self-conscious), but this would go against the fact that the driver cannot report spontaneously what he has been doing (Wider, 1993). Wider also argues that the only way to save Sartre's pre-reflective self-consciousness is to strongly relate it to his theory of bodily self-consciousness. The driver behaves through the movements of his body, bodily self-awareness seems implicit in the driver's behaviour, therefore the driver is in some way pre-reflectively conscious, he simply has a kind of "consciousness-plus-quick-forgetting" (Wider, 1997). Another potential Sartrean answer to the long-distance truck driver - which I prefer - is given by Katherine Morris, using an example from *Being and Nothingness*: «Here we need to point out that the picture of perception sketched

So, I would not accept neither the thin nor the rich versions as absolute answers to the refrigerator light problem (I am nearer to this last one though). As I just argued, saying that experience is sparse and modular means confusing one important element of consciousness – positional consciousness – with the broader structure of consciousness in general. On the other hand, saying that our experience is pervasive and rich of contents does not explain how we actually have different layers of conscious states – focal attention, intentional consciousness, foreground and background awareness, etc. – some of which are truly peripheral and could pass in silence during the report, if they were not elicited by something.

2.3.2 Unconscious Perception and the Blindsight case

After having examined the refrigerator light problem and the long-distance truck driver problem and after having exposed through them the idea of pre-reflexive self-awareness, it will now be easier to tackle the question of unconscious perception. As I just mentioned above, in our daily experience we are not clearly aware of all the details that meet our receptors, all together, at the same time and with the same priority. There seems to be also cases in which it we can actually perceive something unconsciously, i.e. without conscious awareness, which might have significant impact on later behaviour and thought (Faivre et al., 2014; Schnuerch et al., 2016), like in unconscious perception:

«[U]nconscious and conscious perceptual states may be alike in terms of content, but they implicate very different processes. Unconscious perceptual states are unavailable for deliberation and report. They are representationally analogous to conscious mental states, but functionally different. Thus functional difference may result from the fact that unconscious perceptual states are not modulated by

above [the long-distance truck driver], complex as it was, was still of the simplest kind of perception. Sartre describes a more complex case involving as it were layers of foreground and background when he discusses reading while having a painful finger: although 'my body does not cease to be existed in totality', in this case 'there is the presence of a particular structure of the body on the total ground of corporeality'; if a little later I turn the page with my painful finger, 'the pain in my finger, without becoming thereby an object of knowledge, will pass to the rank of ... a new organization of my body' (BN 334-335). We can imagine similar layers of figure-ground constructions in (iia) [the long-distance truck driver]: should a car in the next lane try to cut in front of the driver's truck too tightly, that car will momentarily become a figure on new organization of the road» (Katherine Morris, 2010, p. 154).

attention, and, without attention, they can access semantic networks, but not executive centres. Semantic networks are local, nestled within the temporal cortices within which perception occurs. Executive centres are far off in frontal context. Thus, unconscious perception may be distinguished from conscious perception by distance travelled. It is perception without the benefit of a frontal cortex. Such perception is not useless. It might be what reptiles and insects rely on, but it is not available for the higher cognitive functions that undergird human uniqueness. [...] Indeed, such states are very much like conscious perceptual states in terms of the information they carry, but they present that information without qualitative character, and they don't let it travel very far» (Prinz, 2015, pp. 20–21).

This is particularly evident in cases like blindsight, change blindness, inattentional blindness and neuropsychological syndromes, e.g. visual form agnosia, optic ataxia, neglect, Capgras delusion and split-brain cognition (Jensen et al., 2011; Naccache, 2009)⁵². Regarding blindsight, people with damages to the primary visual cortex are not able to see stimuli presented to certain areas of the visual field. At the same time, they possess wide-ranging residual visual capacities, “blindsight”, in the blind part of their visual field. Indeed, people with blindsight, even reporting no awareness of the target area and claiming to be guessing about what there is there, they actually are significantly better than chance on forced-choice visual and visuo-motor tasks such as stimulus discrimination, stimulus detection or orientation to stimulus spatial source by visual saccades. In other words, despite claiming not to be able to see stimuli presented in the blind field, patients are nevertheless better than chance at guessing the stimulus properties or location (Ramsøy & Overgaard, 2004)⁵³.

⁵² It must be said that concepts like “blindsight” and “unconscious perception” are literally controversial: «Perception is paradigmatically conscious, which is to say that the examples of perception that are most familiar take place consciously. Such examples include seeing a rose in full light, straight on, for a few seconds, or inhaling and enjoying its fragrance. When asked to provide examples of perception, we don't think of cases where a stimulus is too brief or peripheral to be fully and clearly experienced. We learn the concept of perception from conscious cases, and this raises the question: what would it even mean to say that perception can take place unconsciously?» (Prinz, 2015, p. 2).

⁵³ Unconscious processing of information is present also in healthy subjects, e.g. visual masking: «In fact, this unconscious visual process discovered in blindsight subjects is also active in healthy human subjects free of any visual cortex lesions. One way to observe it consists of using paradigms of masked or ‘subliminal’ visual stimulation in which a stimulus is briefly flashed foveally for tens of milliseconds, then immediately followed by a second stimulus, suppressing conscious perception of the former» (Naccache, 2009).

So, if it is true that every conscious state like perception states must involve at the same time pre-reflective self-awareness, and if it is true that consciousness and the what-it-is-likeness are necessarily related (like pain and sensation of that pain), what is unconscious perception and what happens when we unconsciously perceive something? The possibilities are the following: unconscious perception is actually a case of unconscious processing, i.e. there is no consciousness at all⁵⁴, therefore it is not a problem for our research on consciousness. Or, unconscious perception has some degree of awareness, therefore it is a particular kind of consciousness. The first possibility has problems in explaining how it is possible for the subject to guess the position of the object – as in blindsight – without awareness of it since that the whole process is unconscious. For the second one, on the other hand, it is hard to explain how there can be unconscious awareness.

Before discussing this issue, I need to give some other information about the concept of unconscious perception, focusing particularly on the case of blindsight. Even though unconscious perception is commonly considered as a well-established concept, there are still concerns about what it exactly means and in which measure one may be unconsciously perceiving something, etc., so that the debate is far from being closed (e.g. see Peters et al., 2017; Phillips & Block, 2017)⁵⁵. First, there are doubts about the meaning of unawareness or unconscious perception as such. If awareness of a stimulus means that «the subjective experience of that stimulus is different from the subjective experience of the stimulus' absence» (Peters et al., 2017, p. 2), then there is an actual difference during the unconscious perception tests, since the participants can discriminate the “unconscious” target above chance (i.e. they are perceiving it) and in some cases (like in Peters & Lau, 2015) they can also tell the interval that contains the target, suggesting that they are in some way conscious of them (Peters et al., 2017). This means that we may be confusing the ability of report a subjective experience with the phenomenal character as such of that experience: «Just because an observer reports he did not see a stimulus or has low confidence does not mean he had zero subjective

⁵⁴ Ned Block sustained that the process remains completely unconscious, i.e. it is a case of subliminal perception, and blindsight patients do not possess neither phenomenal nor access consciousness, i.e. there is nothing it is like for them to see the stimuli and they cannot report any object voluntarily from the blind field of their vision (Block, 1995). For more information about his new position on unconscious perception, see Peters et al., 2017.

⁵⁵ «Most of us in the field do believe the unconscious version of perception exists, but many of us don't believe it has been convincingly demonstrated» (Peters et al., 2017, p. 3).

experience of it, only that his experience fell below some (potentially very) arbitrary threshold for reporting “seen” or “high confidence”» (Peters et al., 2017, p. 2)⁵⁶. So, holding blindsight as an example, we may consider the possibility that it is not as a genuine case of perception without phenomenal consciousness, rather as abnormal or degraded – but nonetheless conscious – vision (Phillips, 2016), leading us to think that awareness could not be an all-or-nothing but a multi-layered or multileveled activity⁵⁷.

Anyway, assuming that unconscious perception is a genuine phenomenon of receiving stimuli relevant for our behaviour and cognition without awareness of them, for example in blindsight patients, I shall argue that this does not enter into contradiction with the idea of pre-reflective self-awareness as a necessary condition for intentional consciousness. How can this be consistent with the idea the a blindsight patient is perceiving something without being aware of doing that? In order to answer this question – this time we will go beyond the thin-rich controversy – I need to turn

⁵⁶ Anyway, it seems that unconscious perception cannot help us in a decisive way about the controversy between thin and rich view of experience: «Advocates of the thin view often offer empirical arguments for their position, but these arguments, too, are badly question-begging. A favourite argument is this: Absent attention, we fail to parse, respond to, notice, or remember what one might ordinarily think would be salient stimuli — a stream of speech coming in one ear (Cherry, 1953; Moray, 1959) or a woman in a gorilla suit walking through a ball game (Simons and Chabris, 1999). Therefore, it’s said, we are ‘blind’ (or ‘deaf’ or ‘numb’) to these stimuli; we don’t experience them (e.g., Dennett, 1991; Mack and Rock, 1998; Wright, 2005). Here’s the flaw in that argument: It’s one thing (indeed a very important and interesting thing) to show that we don’t do much processing of unattended stimuli; it’s quite another to say we have no experience whatsoever of those unattended objects. The conclusion simply doesn’t follow (and many psychologists refrain from drawing it). We may not parse the speech semantically (very much) or represent the black blob in the middle of the crowd as an ersatz gorilla, but we may still experience that unattended speech and gorilla in some more inchoate way (Simons, 2000). Furthermore, unless we really are blind, or deaf, or numb, we do process the unattended stimuli to some extent — as Searle points out, and as is universally acknowledged on all sides of the debate. We are drawn to the unexpectedly approaching object, the unanticipated call of one’s name, the familiar phone ring or doorbell that others can barely hear, the gentle tap on the shoulder. Such things must first register pre-attentively in some way to call our attention. The question is whether whatever processing or responsiveness or preparedness to respond we have prior to attention is enough to underwrite actual sensory consciousness. The present argument (and similar arguments involving ‘change blindness’, e.g. Rensink, 2000; 2004) doesn’t address that question. [...] Does mere behavioural responsiveness, for example, or above-chance responding on forced-choice questions about the presence of stimuli, demonstrate ‘consciousness’ of those stimuli? Not in any sense that should move an advocate of the thin view. Does failure to report stimuli outside of attention show that they weren’t consciously experienced? Not in any way that should satisfy an advocate of the rich view» (Schwitzgebel, 2007, pp. 10–11, 32).

⁵⁷ «[...] some recent studies cast doubt on the absence of phenomenology in these [blindsight] patients. For example, a recent study that used a graded measure of awareness, suggests that stimuli in blindsight patients might in fact be weakly consciously perceived» (Persuh, 2018, p. 1. For the study, see Mazzi, Bagattini, & Savazzi, 2016). «One conclusion that seems in line with current empirical findings is the suggestion that ‘blindsight’ is a condition in which some level of preserved visual abilities correlate with severely degraded, abnormal visual experiences» (Overgaard, 2012, p. 612).

back to Sartre once again. As Morris has faithfully exposed, there are no grounds for interpreting the cases of unconscious perception as a well-rounded act of perception (Katherine Morris, 2010, p. 153). If we accept that intentional consciousness is an act directed towards an object posited within a temporal and spatial worldly background⁵⁸, exhibiting constancies of shape, size, colour, etc., and which we pre-reflectively constitute, precisely, as a transcendent object other than me⁵⁹, i.e. through pre-reflective self-awareness, then that unconscious perception that happens during blindsight is not a concrete case of perception, rather a case of stimulus discrimination. The stimuli we constantly receive from our sensory receptors must be distinguished from consciousness of objects organized within a spatio-temporal, practical and meaningful world (as it was in the long-distance truck driver problem, in which there were meaningful objects to be conscious of, like the road-to-be-traveled, the exit-sign-to-be-find, the point-of-delivery-to-be-reach, etc.). Moreover, the blindsighted subjects report themselves as “guessing” about the stimulus, however the act of guessing is not even close to the conception of awareness discussed in this chapter, that is not merely the subjective difference between stimulus’ presence and stimulus’ absence (so, even though I agree with those who say that awareness is a multi-layered activity, I would not include the “guessing” activity within the category of “being aware of something”). Therefore, the stimulus discrimination involved in cases of unconscious perception does not seem to fulfil the conditions for being considered as a perceptive act in a very real sense⁶⁰.

Taken together, I tried to challenge the concept of consciousness as intentional act directed to something that must be structured, necessarily and jointly, within a non-

⁵⁸ «In the first place let us note that orientation is a constitutive structure of the thing. The object appears on the ground of the world and manifests itself in a relation of exteriority with other “theses” which have just appeared. Thus its revelation implies the complementary constitution of an undifferentiated ground which is the total perceptive field or the world. The formal structure of this relation of the figure to the ground is therefore necessary. In a word, the existence of a visual or tactile or auditory field is a necessity» (Sartre, 1956, p. 316).

⁵⁹ «In the case of the perception of the chair, there is a thesis—that is, the apprehension and affirmation of the chair as the in-itself which consciousness is not. What consciousness has to be in the mode of being of the For-itself is not-being-the-chair. For its “not-being-the-chair” is, as we shall see, in the form of the consciousness (of) not-being (i.e., the appearance of not-being) for a witness who is there only to bear witness to this not-being» (Sartre, 1956).

⁶⁰ «“[S]ense perception is in no way to be distinguished from the practical organization of existents into a world” (BN 321); but the visual stimuli are not practically organized in the ‘world’ of blindsighted individuals. They are not confronted with a perceptive-cum active field. Hence blindsight is not a counter-example to Sartre’s conception» (Katherine Morris, 2010, p. 153).

intentional, pre-reflective act of self-awareness. I have raised some issues related to the possibility of unconscious experiences, respectively, the refrigerator light illusion, the long-distance truck driver problem, and the unconscious perception, in particular through the case of blindsight. Generally, the main argument against this concept of pre-reflective self-awareness exposed here above was the following: since that there are cases – fictional or real – in which it is possible to imagine or detect presence of consciousness without self-awareness, the Sartrean conception of positional consciousness as also pre-reflective consciousness of itself, i.e. intentional consciousness as also pre-reflective self-awareness, is being questioned. However, as I showed, in all the cases exposed and discussed, the phenomenological structure of consciousness as pre-reflective self-awareness is preserved and also distinguished from the reflective act of positing consciousness as an object.

3. Neuroscience of Consciousness and Neuro-centrism

In this chapter, I will look into the contemporary scientific context, and more specifically into the neuroscientific studies on consciousness. If the purpose of neuroscience is to provide an explanation of consciousness, we must discuss about what a science – and, in particular, a neuroscience – of consciousness would be. A scientific account, in a very general way, should offer a theory that could explain, predict and describe the phenomenon of consciousness in all its features. In this regard, I expose and discuss various approaches and operationalizations in neuroscience (brain-based, mind-based, brain-reductive, mind-reductive), in particular the typical method of description of consciousness in terms of “contents” and “levels”. Then I critically examine the core theme of the neurosciences of consciousness, that is, the search for the so-called neural correlates of consciousness (NCC), which is often considered the holy grail that will reveal us all the secrets of experience, but that nevertheless is stuck into a mind-based or brain-reductive approach. More generally, most of contemporary scientific researches on consciousness maintains that consciousness can be now explained through all the new methods of neuroimaging available, and that in this way a complete description of how we experience can be reduced to the mechanisms of the brain. This kind of attitude towards consciousness – neuro-centrism – is based on the idea that our personality, our mental life, our consciousness itself are all products or equivalent to our brain activity (the brainhood paradigm), which finds its roots in the modern scientific revolution and philosophy. About that, I will delineate an historical-epistemological analysis of the brainhood paradigm, trying to establish a comparison between, on the one hand, Galileo and Descartes and, on the other hand, the contemporary reductionist and eliminativist approaches to consciousness.

3.1 The Neuroscience of Consciousness

In the view of the above, we have exposed the problems of consciousness and argued for a provisional definition, starting from the meaning of the term itself, in the way it is used in our daily language, to a preliminary – and thus non-exhaustive – phenomenological analysis of some of its fundamental features, like intentionality, pre-reflective self-awareness and what-it-is-likeness. What has been lacking until now is a concretization of the concept within a scientific contest, since so far, I only give formal definitions and conceptually rich descriptions without however any empirical support, except for the introduction of the bodily awareness. Thus, after the provisional analysis given in the previous chapters there still are several questions about consciousness, from both the philosophical and scientific levels. If the purpose is to provide an explanation of it, as exposed in the first chapter, we must discuss about what a science – and, in particular, a neuroscience – of consciousness would be. A scientific account, in a very general way, should offer a theory that could explain, predict and describe the phenomenon of consciousness in all its features. In other words, the theory must define what consciousness is (description), how it comes to existence (explanation), which its functions are (functionality) (Van Gulick, 2018).

One of the aims of this chapter is to analyse the concept of consciousness in our contemporary scientific context, hence we need to deal with the scientific discipline that, along with philosophy of mind and phenomenology, is investigating consciousness, i.e. neuroscience. Neuroscience – in particular some branches of cognitive and computational neuroscience – has become progressively relevant in recent studies on consciousness. During the nineteenth century, psychology and neuroscience provided new methodologies to study the mind and nowadays consciousness is of interest not only for philosophers but also for neuroscientists. Even more, with the recent neuroscientific revolution, it seems possible to study the mind by looking into the brain. Thus, for a philosopher it is now problematic or even impossible to study consciousness without considering also empirical findings from other disciplines. As I have already showed in the previous chapter, the relation between brain and consciousness – or more generally with the mind – is far from being a firm achievement of contemporary science and philosophy. If we follow the logical and phenomenological structure of consciousness exposed above it is impossible to

accept the identity between the brain and consciousness. Indeed, from a logical perspective, the sentence “consciousness is the brain activity x ” is not the same kind as “water is H_2O ” or “heat is molecular motion”; from a phenomenological perspective – at least from what we have seen so far – consciousness involves the body and its immersion within the environment, therefore it is quite reductive to indicate the brain as the only candidate for a possible relation of identity. However, even though the necessary identity relation seems impossible from a Kripkean perspective, this does not mean that there is no relation at all between brain and consciousness. From a logical perspective, that relation might be interpreted in an anti-Kripkean way, as a contingent version of identity, i.e. two terms differing in sense but identical in reference so that we can imagine some possible world in which one side of the identity exists without the other (see the example of Lump and Goliath in the previous chapter; Gibbard, 1975; Schwarz, 2013), or as an asymmetric relation of composition, i.e. “lightning is (composed by) electrical discharge, but not vice versa; clouds are (composed by) water molecules, but not vice versa” (Place, 1956)¹.

3.1.1 Approaches and Operationalization of Consciousness

Methodologically speaking, in the debate on the relationship between mind and brain, it is possible to distinguish four different approaches currently in use, i.e. mind-based, brain-based, brain-reductive, mind-reductive (for the first three, see Northoff, 2014). The first approach is the mind-based approach, which puts the mental at the core of the analysis by addressing the question “how do the mind and its mental features relate to the brain?”. The second approach is the brain-based one, which overturns the former question into “how do the brain and its neural features relate to the mind?”. The former approach entails the assimilations of mental concepts into neuronal structures, while the latter does not establish the correlation of any mental feature to a neural one; rather, a brain-based approach starts from considering the brain as such and its relationship

¹ Similarly, I used the asymmetric composition as a key to understand the relation between integrated information in the brain and consciousness supported by Tononi, Koch, Massimini and colleagues (Oizumi et al., 2014; Tononi, 2017b, 2017a): conscious experience may be composed by integrated information in a complex, but at the same time there is no sense in saying that integrated information in a complex is composed by conscious experience (Zilio, 2019).

with the body and the environment, and it only secondarily tries to find the correlation between some neural mechanism or activity and some mental feature or cognitive function. Both methodological approaches can lead to reductive consequences, such as the brain-reductive approach – i.e. “how to reduce the mind to the brain as the final goal” – where the mind is addressed as a negative template, with the aim of reducing it to some cerebral features – e.g. epiphenomenalism, identity theory and eliminativism. On the other side, from the same premises it may be possible to develop a mind-reductive approach – i.e. “how to reduce the brain to the mind as the final goal”, where the mental life, with all its mental features, is taken as the foundation of every possible epistemological and ontological examination, instead of the brain, e.g. some anti-naturalist, panpsychist or strong dualist positions (for example, Pelczar, 2015). This classification is merely methodological, with no ontological commitments, indeed in the mind-based approach we may find dualisms as well as functionalism and non-reductive physicalism, while non-reductive neurophilosophy and the 4E cognition (extended, embedded, embodied, enactive) could be found in the brain-based one.

Maybe surprisingly, I argue that contemporary neuroscience relies principally on the mind-based approach and its consequent radicalization, i.e. brain-reductive. Indeed, even though the main focus is the brain, its neural activity is considered and then investigated through the lenses of the concept of mind we have inherited from philosophy and our western culture in general. The brain is studied as the seat or the store of various mental features, even when the aim is to eliminate the mind since the brain is interpreted as the new mind. Indeed, it does not matter what ontological commitment is endorsed by some reductive/eliminative position, because, despite being brain-reductive in its final goal, it still lies on a mind-based starting point, i.e. leaning towards – and obsessed by – the enterprise of bringing back every mental element to its brain equivalent. Accordingly, even eliminative materialism may be considered as a radicalization of the mind-based approach into the brain-reductive one, because the mind still remains the (negative) starting point of the research project².

² Differently, the brain-based approach does not provide for the “brick by brick” neural reconstruction of the mind in the brain, rather it involves several and partially independent perspectives towards the brain, e.g. epistemological, metaphysical, phenomenological, and of course empirical, in order to connect – and not reduce or eliminate – phenomenal concepts with neural concepts (Northoff, 2014b).

Within the cognitive and neuroscientific context, the concept of consciousness is primarily subdivided into two dimensions, the level of consciousness (or arousal) and the content of consciousness. The former «applies to a whole organism and refers to a scale ranging from total unconsciousness (e.g. death and coma) to vivid wakefulness. A ‘conscious organism’ is one that is capable of having non-zero conscious levels. An organism that is dreaming has some conscious level, although the conscious level is reduced in dreamless sleep. A non-zero conscious level indicates the presence of some conscious content». The latter refers to «the continually changing phenomenal content (e.g. ‘qualia’ such as redness and warmth) and intentional content (e.g. explicitly held beliefs, conscious knowledge) that is present to varying degrees at non-zero conscious levels» (Anil K. Seth et al., 2008, p. 9)³. It is possible to draw a Cartesian plane with these two dimensions, in which we can put different disorders of consciousness depending on their supposed level and content of consciousness (fig. 1)⁴.

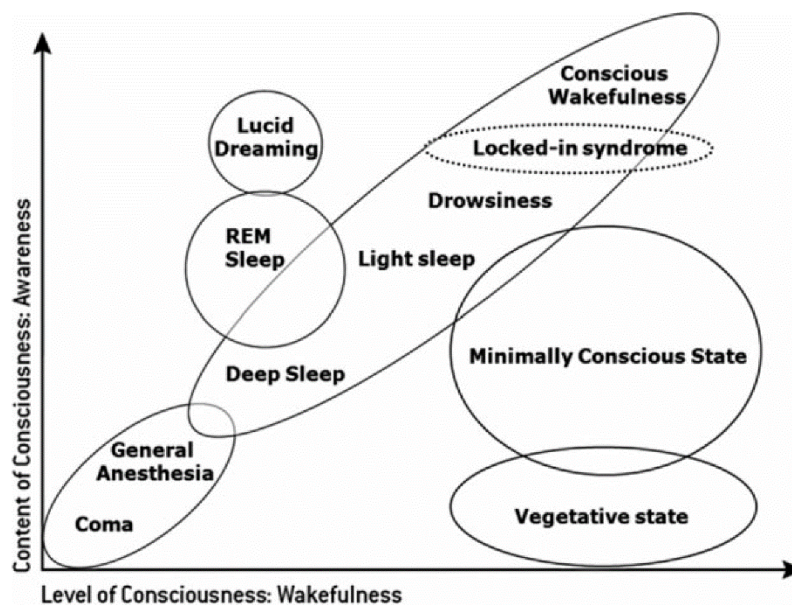


fig. 1

³ There are also representations of consciousness developed through three-dimensional space. For example, Northoff, 2014b describes consciousness with level, content and also form, understood as the ability to organize the phenomenal contents within a structure of spatiotemporal continuity (e.g. a schizophrenic patient may show both high level and content but low form of consciousness). Hanoğlu et al., 2014 represents instead consciousness within a three-dimensional space constituted by level, content of consciousness and self.

⁴ The figure is taken from Gosseries et al., 2011, but it is possible to find slightly different versions of it. For example, in this figure the unresponsive wakefulness syndrome is referred as a state of high level, i.e. wakefulness and low content, i.e. awareness, while in Squire et al., 2012 is indicated as a state with both low level and content, and thus located between coma and general anaesthesia.

The most common conditions are located along the bisector of the Cartesian plane, starting from the normal conscious wakefulness (higher awareness and wakefulness) we can go diagonally through drowsiness (in the last chapter I will talk about locked-in syndrome put in that place in the plane), light sleep (NREM 1-2), deep sleep (NREM 3), and further down some altered states of consciousness, like general anaesthesia and coma (the origin point is supposed to be brain death). Particularly, other altered states or disorders of consciousness like vegetative state/unresponsive wakefulness syndrome and minimally conscious state show the asymmetric relationship between the two dimensions. Indeed, while along the bisector wakefulness and awareness seem directly and bilaterally proportional, in other cases we may find discontinuity, e.g. UWS and MCS (high wakefulness, low-medium awareness), REM sleep and lucid dream state (low wakefulness, high awareness) and others (Nani, 2018; Zilio, 2018b)⁵. The dimensional representation, of course, merely provides a didactic function, after all we cannot expect to complete the description of the characteristics of consciousness through a graph or a figure. The main point is that, in the last years, neuroscience is aiming for an empirical fulfilment of these dimensions (awareness, wakefulness, self, etc.).

In this respect, consciousness is seen as one of the most challenging subjects of the mental domain, and over the last decades there has been a progressive in number of neuroscientific researches on it. As said in the first chapter, the difficulties faced in the empirical definition is one of the problems that characterizes consciousness, both epistemologically and ontologically. For this reason, since it is not possible to directly frame the phenomenon of consciousness, it is common to identify first the phenomenon of consciousness through some indirect representations of it, i.e. through an operationalization of consciousness. By “operationalization” I mean the process of definition of a scientific concept through the operations by which we measure them,

⁵ For example, sleep inertia is a case of high wakefulness with low awareness that everyone has probably experienced at least once in life. Sleep inertia is a state of drowsiness, inattention and confusion sometimes felt during the first 10-20 minutes after a phase of deep sleep, despite having already the ability of hearing, seeing, and generally receiving external stimuli (Marzano et al., 2011). See also Tononi, 2009, pp. 94-95: «Upon awakening, blood flow is rapidly re-established in brainstem and thalamus, as well as in the anterior cingulate cortex. However, it can take up to 20 minutes for blood flow to be fully re-established in other brain areas, notably dorsolateral prefrontal cortex. It is likely that this sluggish reactivation is responsible for the phenomenon of sleep inertia – a post-awakening deficit in alertness and performance that can last for tens of minutes».

i.e. if something is observable – there are causal interactions detectable – then it is object of an operationalization process, via instrumental and methodological mediations⁶. A theoretically complex phenomenon can be translated into a new single operational concept or divided into various operational concepts which are observable and measurable in a quantitative manner. If a concept passes this theoretical-operational test it can be considered a genuine scientific concept⁷. For example, ether and phlogiston has been excluded from science because they did not pass the operationalization of their theoretical concepts.

Now, there are several ways in which researchers tried to operationalize consciousness, and this process can start from the subjective report of conscious experience as well as from observable behavioural reactions or neural activity to stimuli and tasks. About the subjective reportability, there is the translation of the lived experience into psychological scales, e.g. the Verbal Rating Scale (VRS), the Numerical Rating Scale (NRS), the Perceptual Awareness Scale (PAS), the Visual Analog Scale (VAS), etc. (Irvine, 2013a; Overgaard, 2015), a process that leads to the passage from first-person perspective to a third-person methodology. Generally, the necessary condition for a reliable use of these methods is the ability of the subject to cognitively represent and verbally report his/her conscious state – and that is not always the case, e.g. as we have seen in the pre-reflective self-awareness, where if something is not explicitly elicited by some question or by self-reflection, then it is not immediately reportable, even though we have an implicit awareness of it. Anyway, there also are some scales and methods that focus on the tacit behavioural reactions

⁶ «[...] operationalizing a concept really amounts to determining what the causal links between it and the world are; even more striking is that in suggesting the meaning of a term lies in the procedures operationalizing it, proponents of the doctrine were really presenting a kind of causal theory of reference» (Azzouni, 2013, p. 197).

⁷ «Theoretical constraints are (usually) a set of (axiomatizable) sentences fixed in truth-value and containing the terms in question. These constraints [...] can do little by themselves to fix reference. What about operational constraints? These are assertability conditions (however loosely) associated with our (empirical) sentences. In various places Putnam writes of the observations associated with such sentences, or “experiential conditions,” or what can be measured or observed. The picture he gives is this: one finds associated with sentences certain test conditions, and although such test conditions are not stipulated to give the meaning of such sentences, how they turn out, all things being equal, constrains when we can take such sentences to hold. At this point, Putnam’s argument is very simple. Consider permutations of all our terms that hold constant such operational and theoretical constraints. Such permutations are easily available, and since the operational and theoretical constraints exhaust all our resources for fixing the references of the terms of the language, any position (for example, what Putnam calls “metaphysical realism”) which holds that our terms refer to external objects (of whatever sort) must be wrong» (Azzouni, 2013, p. 146).

correlated with supposed and specific conscious states, which are also used for the study of consciousness in animals (Le Neindre et al., 2017). About the behaviour, as we already seen in the previous chapters, there are scales, mainly used within the clinical context, that give a numeric value to the visual and auditory response, the verbal and communication response, the motor and arousal functions, and these values can go from the total absence to a consistent preservation of the function, e.g. the Glasgow Coma Scale, the Sensory Stimulation Assessment Measure (SSAM), the Coma Recovery Scale Revisited among others, (Seel et al., 2010). However, these scales are often useful for the pragmatic and immediate purpose of assessing the “level” of consciousness within the clinical context, but they say nothing about the quality, the structure and the content of consciousness.

Moreover, they may lead to wrong assessments of consciousness in patients with brain trauma that resemble complete unresponsiveness but are instead cases of covert consciousness, as in cognitive motor dissociation (CMD), locked-in syndrome (LIS)⁸ or unresponsiveness due to anaesthesia (e.g. Huang et al., 2018). Consequently, the relevant rate of misdiagnosis (37-43%) between minimally conscious state (MCS) and unresponsive wakefulness syndrome (UWS) (Bender et al., 2015; Schnakers, Vanhaudenhuyse, et al., 2009) elicits methodological, epistemological and ethical concerns about the reliability of these methods taken alone (Evers, 2016; Farisco & Petrini, 2014; Wade, 2018). These issues have pushed research towards the development of new techniques of assessment of consciousness characterized by the independence from behaviour and reportability (Tsuchiya et al., 2015)⁹. In the last decades, technological development in clinical and scientific context has permitted to focus on the activity of the brain by establishing an increasingly precise framework of the relevant neural mechanism of various mental activity, including consciousness. The main goal is not to determine what consciousness as such is, the question here is

⁸ Locked-in syndrome is usually described as a situation in which the body is paralyzed but consciousness and cognition are fully preserved or intact, generally because the thalamo-cortical tracts are spared, (Gosseries et al., 2009) while cognitive motor dissociation is characterized by concomitant injuries across the cerebrum that may significantly limit ability to harness a potential communication channel (Schiff, 2015).

⁹ «In recent years, the arsenal of indirect measures has been supplied with what is named “no-report paradigms.” [...] experiments of this kind attempt first to associate a particular objective measure (e.g., a behavior or a brain activation) with conscious experience, and then to apply this measure as a measure of consciousness so that no direct report is needed» (Overgaard, 2017, p. 2).

more pragmatic and referred to the definition of the biological mechanisms correlated with specific features of consciousness, e.g. responsiveness, awareness, wakefulness, arousal, etc.¹⁰, so that it could be possible to interpret neural activity in indices and markers of consciousness, i.e. operationalizations¹¹.

3.1.2 The Neural Correlates of Consciousness

In the last years, neuroscience is aiming for an empirical fulfilment of the theoretical dimensions of consciousness (awareness, wakefulness, self, etc.) by analysing the relevant activity of the brain. The fundamental postulate is that the brain is necessary – or even better sufficient – for the emergence of experiential states and the main goal has over time become the search for the so-called neural correlates of consciousness (NCC), the minimal neural activations that are sufficient for a specific content-consciousness, e.g. the NCC of a percept, a mental image, a memory, but also for state-consciousness as such, i.e. the NCC that mark the difference between background states and alterations of consciousness, e.g. being asleep, awake, dreaming, etc. (Chalmers, 2000; Hohwy, 2009; Koch, 2004; Overgaard, 2017)¹². In other words, the

¹⁰ As exposed in the first chapter, in this context the “real problem” is preferred over the “hard problem” of consciousness: «[H]ow to account for the various properties of consciousness in terms of biological mechanisms; without pretending it doesn’t exist (easy problem) and without worrying too much about explaining its existence in the first place (hard problem)» (Seth, 2016). As I will argue later, sometimes empirical theories of consciousness mistaken indirect operationalizations of consciousness for phenomenological characteristics of it, sliding from an empirical-ontological plausibility to an empirical-ontological fallacy, i.e. when one uses one and the same concept on both the empirical and ontological levels in the same way, without distinguishing the different domains (Northoff, 2018b).

¹¹ In this respect, in the last years it has been possible to assess consciousness to patients resembling complete unresponsiveness, behaviourally speaking. In the last 13 years Owen and colleagues have modelled and refined a method for assessing residual consciousness through mental imagery in patients with brain trauma: «In 2006, we put a young woman who had been diagnosed as being in a vegetative state into a functional magnetic resonance imaging (fMRI) scanner and asked her to imagine she was playing a game of tennis [...] We had seen this pattern many times in studies of healthy participants, who we had also asked to imagine playing tennis in the scanner. We then asked the patient to imagine moving from room to room in her house and a very different pattern of fMRI activity emerged. This included the parietal cortex and part of the parahippocampal gyrus, two brain regions that are known to be involved in spatial navigation. Again, this pattern of fMRI activity was indistinguishable from that seen in healthy participants. On the basis of these fMRI findings, we concluded that our patient was not vegetative at all, but conscious and aware, despite the fact that she had been entirely physically non-responsive for more than 5 months at that point» (Owen, 2019, p. 256; Owen et al., 2006; Monti et al., 2010).

¹² «An NCC is a minimal neural system N such that there is a mapping from states of N to states of consciousness, where a given state of N is sufficient, under conditions C, for the corresponding state of consciousness»

NCC is a neural system, without non-necessary components (“minimal neural activations”), that directly correlates with a state of consciousness, that is, when there is the NCC x, the subject is awake, when there is the NCC y, the subject is dreaming, when there is NCC z, the subject is seeing an apple instead of a chair, etc¹³. Thus, the aim is precisely to isolate these neural activations and respectively correlate with background states and specific contents of consciousness. This research field has quickly become one of the most important in cognitive sciences and does not seem to show signs of stopping. Technological progress has radically changed the way in which we investigate the brain in the last decades, thanks to the brain imaging techniques that permits to analyse our neuronal activity and structure in several ways¹⁴. Through different techniques of brain imaging¹⁵ is now possible to look into specific neuronal structures of the brain, observe the reactivity of certain areas during some cognitive task, compare the activity of the healthy brain between various groups of subjects (anaesthesia, disorders of consciousness, sleep, etc.), providing a large amount of empirical findings useful for both clinical and theoretical purposes (Gosseries et al., 2016; Gosseries, Schnakers, & Laureys, 2019).

There are many candidates running for being the most reliable and central NCC, as well as there are many broader neuroscientific theories that aim to account for a complete explanation of how consciousness is related with the brain¹⁶. Despite

¹³ «a [...] content-specific NCC can be characterized as the part of the brain in which a change in activity reliably and specifically predicts a particular change in experiential content» (Boly et al., 2017). Moreover, following Block 1995, there is also another differentiation between the phenomenal NCC and access NCC (Block, 2005), the former relating to phenomenal consciousness (the “early” activations), and the latter relating to access consciousness (the “late” activations).

¹⁴ «Over the last two decades, the development of new scientific techniques has provided powerful tools for studying the human brain in a way that was unattainable before. For the first time, high-resolution pictures of the brain in vivo offered a “physiological window into the human mind” (Dolan 2008). This, in turn, contributed to the current renaissance of neuroscience: crossroads of interconnected disciplines that aim to study the nervous system under specific but also converging aspects (molecular, cellular, developmental, structural, functional, evolutionary, computational, and clinical)» (Nani, Seri, & Cavanna, 2013).

¹⁵ There currently are several techniques of brain imaging that differ for the kind of technology used and for the degree of invasiveness of these devices, each of which serves to detect some specific property of the brain, e.g. anatomical, structural, temporal, spatial, complexity, functional, metabolic, etc. The most important neuroimaging technologies are the functional Magnetic Resonance Imaging (fMRI) the Positron Emission Tomography (PET), the Computed tomography (CT) and the functional Near-Infrared Spectroscopy (fNIRS), the most important neuroelectric technologies are the Electroencephalography (EEG), the Electrooculography (EOG) and the Deep Brain Electrical Stimulation (DBS), while the neuromagnetic technologies are the Magnetoencephalography (MEG) and the Transcranial Magnetic Stimulation (TMS).

¹⁶ Some of the first candidates (during the '90) for the NCC were: «40-hertz oscillations in the cerebral cortex (Crick and Koch 1990), intralaminar nuclei in the thalamus (Bogen 1995), re-entrant loops in

being a neutral term, “correlation” is often interpreted as a synonym of causation, reducibility, dependence, or even identity relation¹⁷. Indeed, correlation means that there is a relationship between two variables, which could refer to a case of dependency, nevertheless this is far from being immediately defined as causation or identity. On the other hand, when NCC is described as the «minimal neural mechanisms jointly sufficient for any one specific conscious experience» (Koch et al., 2016a, p. 307), it seems that they are looking for something epistemologically and ontologically stronger than a simple correlation (Wu, 2018). Rather, it is clear that no one within the neuroscientific context is doubting about the sufficiency of the brain for the emergence of consciousness, indeed the problem now is about where the NCC is/are and which particular NCC activates/causes/gives the foundation to that particular conscious content (e.g. see the debate about frontal vs posterior NCC, Boly et al., 2017; Koch et al., 2016b; Sandberg, Frässle, & Pitts, 2016; Mashour, 2018). In some cases, the correlation between neuronal activity and conscious experience is interpreted as an identity relation, such that consciousness of an apple is not simply caused by or dependent on the NCC x, rather ontologically is that NCC (Polák & Marvan, 2018), in the same way as heat is molecular motion¹⁸. With regard to the possible approaches

thalamocortical systems (Edelman 1989), 40-hertz rhythmic activity in thalamocortical systems (Llinas et al 1994), extended reticular-thalamic activation system (Newman and Baars 1993), neural assemblies bound by NMDA (Flohr 1995), certain neurochemical levels of activation (Hobson 1997), certain neurons in inferior temporal cortex (Sheinberg and Logothetis 1997), neurons in extrastriate visual cortex projecting to prefrontal areas (Crick and Koch 1995), visual processing within the ventral stream (Milner and Goodale)» (Chalmers, 2000, p. 17), etc. (for review articles, see Chalmers, 1998, 2000; Metzinger, 2000; Seth et al., 2006). Some of them has been rejected, others instead have been integrated with new empirical discoveries and incorporated within more recent theories of consciousness. See, for example the Global Neuronal Workspace Theory (GNWT) (Dehaene et al., 2006, 2011; Dehaene & Naccache, 2001), the Integrated Information Theory (IIT) (Tononi, 2004, 2008, 2017b), the Temporo-spatial Theory of Consciousness (TTC) (Northoff, 2018b; Northoff & Huang, 2017), the Local Recurrency Theory (Lamme, 2006, 2010).

¹⁷ «At this point a common fallacy – *post hoc, ergo propter hoc* – is tempting to commit. This fallacy is committed when one infers that φ caused ϕ simply because φ is correlated with ϕ . [...] Correlation does not entail causation. One needs further rationale to infer that a correlation is best explained by (or suggests) a causal relation. Likewise, a correlation by itself does not entail dependency, identity, or that one correlate is reducible to the other. Given that φ is correlated with ϕ , we need more information to justifiably conclude that φ depends on ϕ , is identical to ϕ , or is reducible to ϕ » (Owen, 2019).

¹⁸ «Though the various metaphysical theories appear neutral with respect to how one approaches the NCC search, it seems that full discovery of the NCC could have consequences for our view of the metaphysics of consciousness. For consider what we should say if highly plausible, strongly systematic NCCs were found for most types of conscious states in most humans (and perhaps monkeys). Those of a reductionist bent would probably be tempted to say that consciousness is then identical with those brain states – that is, we would be tempted to accept the identity theory of consciousness (Smart; Hill). This is not an unreasonable inference – it seems not too different from what happened when we discovered the identity of water and H₂O, or of Lawrence of Arabia and T. E. Shaw. For the identity

exposed above, the brain-based – “how do the mind and its mental features relate to the brain?” – and the brain-reductive ones – “how to reduce the mind to the brain as the final goal” – are the most used within the neuroscientific context.

3.1.2.1 Criticisms and Refinements of NCCs

Yet, there are also critics regarding the meaning, the use and the goals of NCC, especially when it is taken as the ultimate solution to the issues of consciousness. As said, the NCC is supposed to be the minimal neural activations that are sufficient for the emergence of consciousness, in such a way as to enable us to describe and predict the conscious state of a person by looking into his/her brain. However, from a methodological perspective, isolating the neuronal activation that relates exclusively to the conscious state and removing any other neural activity contingently correlated, e.g. stimulus encoding process, higher order cognition, attention, memory, etc. is far from straightforward, and the NCC proper can be confounded with prerequisites for and consequences of conscious processing of the particular content (Aru et al., 2012, p. 744)¹⁹. In other words, the search for the NCC may even confound the spatio-temporal neural dynamics needed for the emergence of a conscious state with the final result (the spatio-temporally punctiform neural activation). Alternatively, relying on subjective reports leads to other issues, e.g. confounding the NCC with neural mechanisms for cognitive functions since that reports presuppose cognitive access, while, on the other hand, no-report paradigm may confound the NCC with neural mechanisms underlying unconscious processes (Schlicht, 2018).

From an ontological perspective, this often depends on whether we understand consciousness more like a process or an event. Also, the NCC is supposed to be extremely useful in detecting residual consciousness in patients with a brain injury, however the reorganization of the neural structure in the injured brain may make hard the comparison between the NCC of a healthy brain and the neural activity of UWS

theorists, identities cannot be explained (Papineau), so for them, the systematic NCC could be the only empirical step required in solving the problem of consciousness» (Hohwy, 2007).

¹⁹ For example, «[d]uring imaging experiments, prerequisites (e.g. stimulus processing, attention) and consequences (e.g. performance, attention, working memory, motor preparation, verbal report, etc.) of consciousness can be easily confounded with the actual NCC» (Morales & Lau, 2019, p. 6).

patients (Chalmers, 2000; Overgaard, 2017), therefore results from brain imaging must be carefully considered²⁰. Indeed, the reverse inference used in this kind of cases (if mental state x, then neural state y; neural state y is true, therefore mental state x is true) does not permit a direct implication of consciousness from the neural activity, nevertheless it provides undeniable and sometimes indispensable help in the clinical and pragmatic assessment of states and levels of consciousness in behaviourally unresponsive patients (Owen, 2013)²¹. Furthermore, by taking up again some of the issues discussed in the first chapter, the localization of the NCC alone does not provide any explanation about the qualitative aspects of that conscious state, nor how it can emerge. In other words, the search for neural correlates of consciousness is still an easy problem.

In view of all this, a refining process of the NCC research has recently begun, in order to offer a more fine-grained correlation between neural activity and experience. For example, Anil Seth argues that a classical NCC can only show relations between specific brain activity and conscious states, while the “explanatory correlate of consciousness” must be based on and eventually account for some of the fundamental properties of our experience, i.e. complexity of the conscious scene we experience (integration and differentiation at the same time, see also (Tononi et al., 2016), perspectivalness (first-person perspective), presence of affective conscious contents, intentionality and sense of agency (Seth, 2009).

²⁰ In particular, Chalmers exposed several methodological warnings, which can be found in detail in Chalmers 2000, e.g. “lesion studies are methodologically dangerous” (as we seen just now), “there may be many NCCs”, “correlation across a few situations is limited evidence”, “we need good criteria for the ascription of consciousness”. For other critics, see Noë & Thompson, 2004; Fink, 2016.

²¹ Despite implying a logical fallacy, i.e. affirming the consequent (if p then q, q is true therefore p is true), the reverse inference is widely used as a good probabilistic tool in cognitive sciences, because of its predictive power (Hutzler, 2014; Nathan & Del Pinal, 2017; Poldrack, 2006, 2011), and it surely demonstrate the insufficiency of the behavioural scales for a reliable detection of consciousness. However, the reverse inference must be always carefully considered, since that having a neuronal-cognitive correlation does not necessarily imply a neuronal-phenomenal inference: «Subjects remaining unconscious may show more or less the same activity pattern during the same kind of tasks. We perform plenty of tasks daily in a rather unconscious mode, meaning that we do not “experience” these tasks. We are thus both responsive and wakeful, but not conscious, with regard to these tasks. This means that responsiveness and wakefulness, including their underlying stimulus-induced or task-related activities, do not imply anything by themselves about the presence or absence of consciousness and its phenomenal features» (Northoff, 2014c, 2016a, p. 112). For this reason, the reverse inference process implied in clinical practices, as in the assessment of consciousness in UWS patients, is often supported by other tests, e.g. multiple repetition of the task, comparison with task-evoked activity in mildly sedated healthy volunteers, and specific tasks related to the autonomous will of the patient, which was explicitly instructed not to follow the command of performing motor imagery (Cruse et al., 2012; Owen, 2013).

Georg Northoff proposes first a preliminary distinction between actual and possible consciousness, where the former is the concrete content of consciousness in a precise point in space and time (e.g. consciousness of the chair), while the latter refers to the condition of possibility for the emergence of any content of consciousness (Northoff, 2014c). Secondly, following the conceptual distillation made by Aru et al., 2012, he argues for the development of more specific neuronal hypotheses related to actual or possible consciousness, i.e. neuronal correlates (content), predispositions (level), prerequisites (form) and consequences of consciousness (cognition) (Northoff, 2014c; Northoff & Heiss, 2015; Northoff & Huang, 2017)²². In this sense, the NCC only refers to actual consciousness, highlighting the sufficient brain activation for that particular content of consciousness in that specific spatio-temporal context. I say “in that specific spatio-temporal context” because arguing that the neuronal state x is sufficient for the actual conscious content y does not entail also the necessity of the correlation, indeed, in another spatio-temporal context with different environmental and bodily conditions, there may be another neuronal state x_1 to elicit the actual conscious content y ²³. Therefore, it would be better to define NCC as non-necessary sufficient conditions of actual consciousness (Northoff, 2014b, 2014c). Next to the NCC, Northoff proposes the identification of neuronal prerequisites of actual consciousness (preNCC), which correspond with the necessary non-sufficient conditions for the occurrence of an actual consciousness, i.e. its “enabling conditions” (Koch, 2004) for setting the “neural substrates” (NCC)²⁴. Regarding the condition of

²² «The various NCC candidate mechanism seem to presuppose consciousness as homogeneous unitary phenomenon. [...] Instead of focusing on one unitary neuronal mechanism only, we need to develop more specific hypotheses that link the various phenomenal features of consciousness to distinct neuronal mechanism in the brain. [...] Besides the phenomenal heterogeneity, we also need to account for the subjective nature of consciousness» (Northoff, 2014c, pp. xxiv–xxv). «Various suggestions for a neuronal mechanism underlying consciousness were discussed; these included synchronization, global workspace and re-entrant connections. However, all these suggestions for a neural correlate of consciousness (NCC) suffered from unspecificity in neuronal, phenomenal and experimental terms. Because of this unspecificity, the current suggestions for NCC remain unable to account for the various problems raised by the philosophical arguments against reducing consciousness to the brain. Hence, we may need to search for neuronal features in the brain other than the current NCC to get a better understanding of the neuronal mechanism underlying consciousness» (Northoff, 2014b, p. 400).

²³ Being in Rome is a sufficient condition for being in Italy, but it is not necessary though, since I do not need to be necessarily there for being in Italy, for example, I could be also in Milan.

²⁴ The enabling conditions « [...] may, for instance, concern the involvement of subcortical structures like the brainstem that may remain insufficient by themselves to yield consciousness (see Chapter 31 for a discussion of subcortical regions and consciousness; see Koch 2004). Psychologically arousal or vigilance may be regarded an “enabling” and thus necessary nonsufficient condition of consciousness» (Northoff, 2014c, p. lvi).

possibility of actual consciousness, i.e. possible consciousness, Northoff proposes the neural predisposition of consciousness (NPC), that is, those unconscious processes that makes possible the emergence of consciousness, i.e. necessary non-sufficient conditions for possible consciousness²⁵. In addition, regarding the difficult distinction between phenomenal consciousness and cognition, he identifies the neural consequences of consciousness (NCCcon) related to the ability to cognitively access and report conscious contents²⁶.

3.1.3 Drastic Scientific Approaches towards Consciousness

So far, we have seen that consciousness has become one of most important targets of neuroscientific research, culminating in the search of the neural correlates of consciousness. As said, this scientific field has remarkable potentialities as well as conceptual and empirical criticisms. Interestingly enough, the NCC per se and its related concept of “correlation” can be interpreted in several ways, some of them leading to particular hypotheses that are very far from the mainstream neuroscience. As said, the fact that there are observable correlations between neural states and conscious states does not automatically imply reducibility, dependence, identity or causation, and also it does not imply any specific direction in the relation (e.g. from brain to consciousness or vice versa). Indeed, the notion of “correlation” is theoretically neutral (Owen, 2019) because it «doesn’t by itself commit one to any particular metaphysical view about the relation between (neural) matter and consciousness» (Hohwy, 2007, p. 461). In this sense, ontologically speaking, it is possible to draw several interpretations about the relation between brain activity and consciousness: causation (consciousness is caused by the NCC), reverse causation (the NCC is caused by consciousness), third cause (consciousness and the NCC are both caused by some third element), constitution (consciousness is constituted by the NCC),

²⁵ «The neural predisposition of consciousness (NPC) concern those neural conditions that predispose the brain to constitute possible consciousness, e.g. actual consciousness and unconsciousness, rather than nonconsciousness» (Northoff, 2014b, p. 410).

²⁶ Of course, Northoff suggests also a specific kind of neuronal activity for each conceptual components of his theory (NPC, preNCC, NCC, NCCcon), but, for the moment, the aim here is only to illustrate a more fine-grained definition of the neural correlates of consciousness. For details, see Northoff & Huang, 2017.

reverse constitution (the NCC is constituted by consciousness), third constitutor (consciousness and the NCC are both constituted by some third element). However, the two models mainly used within the neuroscientific context are the causation and the constitution interpretations. Moreover, from an epistemological point of view, the relation between first- and third- person perspectives can be interpreted in many ways: elimination of FPP as cognitive illusion or epiphenomenon (e.g. Churchland, Crick), reduction of FPP within the TPP methods (e.g. Dennett and generally the neuroscientific approach), FPP as irreducible to the physiological dimension (e.g. Searle), FPP as a central qualitative dimension (e.g. Chalmers), FPP to be investigated by both phenomenological and empirical methods (e.g. neurophenomenology) (Buongiorno, 2014). Again, the most used approaches are the reduction or the elimination within the scientific context.

Thus, consciousness is generally supposed to be caused or constituted by the NCC while the FPP eliminated or reduced by the TPP, and this entails epistemological and ontological consequences on the concept of consciousness within the whole scientific domain. Besides, if consciousness was really a mere effect of the NCC or if it could be constituted by the NCC as H₂O constitutes water, then the neuroscientific findings about the NCC *per se* would be sufficient for a complete and scientific description of consciousness. In order to make things clear, I shall now expose three examples of drastic approaches in neuroscience and cognitive science towards consciousness, which entails – respectively – its reduction, transmutation and scientific elimination.

3.1.3.1 Reduction: Daniel Dennett's Heterophenomenology

Daniel Dennett has developed over the years one of the most influential analyses of consciousness through his own method called “heterophenomenology” (Dennett, 1991, 2005, 2007). According to him, consciousness is utterly private, inaccessible to outsiders, somehow partly incommunicable and uninvestigable by science, rather it must be investigated with the same principle of every other empirical science (Dennett, 1991), hence by third-person perspective and on the empirical methods offered by

contemporary cognitive sciences. However, heterophenomenology does not remove automatically the first-person perspective and the qualitative aspects of experience accessed through introspective methods, but it filters and anchors these dimensions through the third-person methodology²⁷, i.e. all the reports of the rich experiential life of every subject are described in a neutral and objective narration, so that even a Martian without our phenomenality could understand it²⁸. Indeed, the process begins with the extraction and purification of a raw text from the speaking subject and his/her “heterophenomenological world” made of «all the images, events, sounds, smells that the subject (apparently) sincerely believes to exist in his or her (or its) stream of consciousness. Maximally extended, it is a neutral portrayal of exactly *what it is like to be* that subject—in the subject’s own terms, given the best interpretation we can muster» (Dennett, 1991, p. 98). In other words, this fictional world is the description of what subjects naively believe to be going on in their conscious lives, and that instead may be completely inaccurate after scientific investigation²⁹.

In this sense, heterophenomenology is opposed to “lone-wolf autophenomenology” – i.e. the traditional phenomenological approach started by Husserl – «in which the subject and experimenter are one and the same person, is a foul, not because you can’t do it, but because it isn’t science until you turn your self-administered pilot studies into heterophenomenological experiments» (Dennett, 2003, p. 23). According to Dennett, autophenomenology sustains the existence of a privileged access towards our mental states, taken not as empirical data but as

²⁷ «The third-person approach is not antithetical to, or eager to ignore, the subjective nuances of experience; it simply insists on anchoring those subjective nuances to *something*—anything, really—that can be detected and confirmed in replicable experiments» (Dennett, 2005, p. 149).

²⁸ «The idea, variously expressed or just tacitly presupposed, is that the Martians can’t play these games that we can play. They can’t engage in a first-person science of consciousness because they aren’t themselves the right kind of first-persons. They can study Martian consciousness from the first-person point of view, if there is any, but not our consciousness. Or it might be said that they can’t engage in a second-person science of consciousness because, being an alien life form, they can’t form the I–thou bond of empathy such a method presupposes. My question is this: Is there any good reason to believe any of this? And my answer will be No, there is nothing about our consciousness that is inaccessible to the maybe metal-headed methods of the Martians. The third-person methods of the natural sciences suffice to investigate consciousness as completely as any phenomenon in nature can be investigated, without *significant* residue» (Dennett, 2005, p. 29).

²⁹ «We tend to overlook the fact that much of what “we”—you and I and our friends and neighbours—believe about consciousness comes from our huge supply of shared, public, objective representations of the streams of consciousness of other folks, real or fictional» (Dennett, 2005, p. 26).

“undeniable truths” (Dennett, 2005, p. 113)³⁰. This is also the reason why so much effort has been spent over these years for the hard problem and the issue of qualia, which are merely errors made due to cognitive biases and misconceptions about our intuitions and first-person data (Dennett, 2018)³¹. Either the FPP is taken into account through the TPP scientific method or a true science of consciousness will be impossible: «if you leave the subject in your theory, you have not yet begun! A good theory of consciousness should make a conscious mind look like an abandoned factory (recall Leibniz’s mill), full of humming machinery and nobody home to supervise it, or enjoy it, or witness it» (Dennett, 2005, p. 70), as Dennett’s Multiple Drafts theory does (Dennett, 1991). In this regard, Dennett seems to hold a genuine eliminativist position, rather than a reductionist one.

³⁰ «Alluding to this division between acute observation and theoretical explanation, the philosophical school or movement known as Phenomenology (with a capital P) grew up early in the twentieth century around the work of Edmund Husserl. Its aim was to find a new foundation for all philosophy (indeed, for all knowledge) based on a special technique of introspection, in which the outer world and all its implications and presuppositions were supposed to be "bracketed" in a particular act of mind known as the epoché. The net result was an investigative state of mind in which the Phenomenologist was supposed to become acquainted with the pure objects of conscious experience, called noemata, untainted by the usual distortions and amendments of theory and practice. Like other attempts to strip away interpretation and reveal the basic facts of consciousness to rigorous observation, such as the Impressionist movement in the arts and the Introspectionist psychologies of Wundt, Titchener, and others, Phenomenology has failed to find a single, settled method that everyone could agree upon» (Dennett, 1991, p. 44). This extremely brief paragraph in which Dennett declared the failure of the entire phenomenological movement has been criticized, of course, by many (e.g. Zahavi, 2007); in particular, see Varela 1996: «Dennett mixes apples and oranges by putting impressionism and introspectionism in the same bag; he confuses introspectionism with phenomenology which it is most definitely not (vide infra); and he finally draws his conclusion from the absence of some idyllic universal agreement that would validate the whole. Well, we do not demand that everyone could agree upon, say, Darwinism, to make it a remarkably useful research programme. And certainly some people do agree on the established possibility of disciplined examination of human experience. In a book that is in many other respects so savant and insightful, this display of ignorance concerning phenomenology is a symptom that says a lot about what’s amiss in this field» (F. Varela, 1996b, p. 334).

³¹ For example, about the hard problem: «Is there *really* a Hard Problem? Or is what appears to be the Hard Problem simply the large bag of tricks that constitute what Chalmers calls the Easy Problems of Consciousness? I cannot prove that there is no Hard Problem, and Chalmers cannot prove that there is. He can appeal to your intuitions, but this is not a sound basis on which to found a science of consciousness» (Dennett, 2005, pp. 74–75). About qualia: «And these subjective experiences, tradition tells us, have “intrinsic qualities”—qualia, in the jargon of philosophers—that I not only do have access to, but that are inaccessible to objective investigation. This idea has persisted for centuries, in spite of its incoherence, but perhaps its days are finally numbered. [...] Are there qualia? If so, just what are they? “Qualia” is the plural of “quale,” a Latin word for quality, but philosophers have endowed the term with a variety of ill-considered associations and special powers. Since the term has no home use other than in the philosophy of mind, we really have no choice other than to let philosophers define it as they will, if only they would! But there is no agreed-on definition. This does not bother philosophers as much as it might be expected to. [...] Yes, it is indeed difficult to deny that there are qualia. I’ve been working on the task for years, with scant progress! The reason it is difficult is mainly that this “standard, broad sense of the term” is a conspiracy of unexamined presuppositions and circularly defined elaborations» (Dennett, 2005, pp. 78–79).

3.1.3.2 Transmutation: Churchlands' Eliminative Materialism

If Dennett claims that consciousness exists but it is not what people think it is, on the other hand Patricia and Paul Churchland have a possibly even stronger position in this regard. Since the publication of the *Neurophilosophy. Towards a Unified Science of the Mind/Brain* (P. S. Churchland, 1986) and *Eliminative Materialism and the Propositional Attitudes* (P. M. Churchland, 1981) the aim of the Churchlands has been that to proceed towards a progressive connection of different levels of explanation through a process of intertheoretic reduction (E. Nagel, 1961), where, given two theories, one can be reduced within the more basic one, leading to the substitution of concepts and phenomena with new ones, more appropriate scientifically speaking³². Thus, as scientific discoveries often cause shifts not only in the scientific field but also in the cultural and common-sense level, i.e. in the way we see and interpret the world³³, also «[...] it is probable that our commonly accepted ideas about knowledge, reasoning, free will, the self, consciousness, and perception, have no more integrity than prescientific ideas about substance, fire, motion, life, space and time» (P. S. Churchland, 1990, p. 341). The indirect consequence of this approach is “ontological simplification”: «Ontology pertains to what entities and properties exist, and in the event of intertheoretic reduction it may turn out that where we had thought there existed two different kinds of phenomena characterized by the laws of two different theories, there is in fact but one kind of phenomenon that is described by both theories» (P. S. Churchland, 1986, p. 280).

³² For example, the claim that light has been reduced to electromagnetic radiation means (a) that the theory of optics has been reduced to the theory of electromagnetic radiation and (b) that the theory of optics is reduced in such a way that it is appropriate to identify light with electromagnetic radiation (P. S. Churchland, 1986, p. 279).

³³ «Scientific discoveries frequently cause a profound shift in how we think about our universe, and consequently they may reconfigure the very questions we ask. Thus, after Copernicus, it was no longer worth asking what made the crystal spheres turn daily; after Harvey it was no longer worth asking how the heart concocted animal spirits; once Newton framed the laws of motion, no one cared about an object's 'natural place' nor about the properties of 'impetus'; after Lavoisier, the problem of the negative weight of phlogiston could safely be ignored as entirely misbegotten» (P. S. Churchland, 1990, p. 341).

In this sense, also the concept of consciousness is supposed to undergo a process of “transmutation” (P. S. Churchland, 1983) towards a new and more mature theory. There are not two phenomena (the cerebral one and the mental one) taken into account by two theories, but it actually exists only one fact (the cerebral one) and one theory (as a result of intertheoretic reduction). This position, that is, the claim that ordinary understanding about our mental life is deeply mistaken and that common-sense mental states (e.g. beliefs, desires, emotions, etc.) do not really exist, is called eliminative materialism (P. Churchland, 2007; Williams, 2019). So, differently from the heterophenomenology, the first-person data (e.g. what we sincerely believe about our mental states) is not considered, collected and cleaned through the TPP, rather it is rejected as useless for any possible explanation of the mind and substituted by neuroscientific theories³⁴. In this sense, there is nothing like the phenomenal character of experience, consciousness exists but as the activity in the brain’s hardware that can be completely investigated by neuroscience (P. S. Churchland, 2002)³⁵.

Furthermore, according to the Churchlands, the neuroscientific revolution should also lead us to a reconsideration of the way of talking about our mental life, i.e. folk psychology that «will eventually be displaced rather than smoothly reduced by completed neuroscience» (P. M. Churchland, 1989, p. 1)³⁶. Consequently, also our

³⁴ «Eliminative materialism is the thesis that our commonsense conception of psychological phenomena constitutes a radically false theory, a theory so fundamentally defective that both the principles and the ontology of that theory will eventually be displaced, rather than smoothly reduced, by completed neuroscience. Our mutual understanding and even our introspection may then be reconstituted within the conceptual framework of completed neuroscience, a theory we may expect to be more powerful by far than the commonsense psychology it displaces, and more substantially integrated within physical science generally» (P. M. Churchland, 1989, p. 1).

³⁵ « I think Dennett is wrong to see human consciousness as the result of a unique form of “software” that began running on the existing hardware of human brains some ten, or fifty, or a hundred thousand years ago. He is importantly wrong about the character of that background software process in the first place, and he is wrong again to see consciousness itself as the isolated result of its “installation” in the human brain. Instead, as I shall argue, the phenomenon of consciousness is the result of the brain’s basic hardware structures, structures that are widely shared throughout the animal kingdom, structures that produce consciousness in meme-free and von Neumann–innocent animals just as surely and just as vividly as they produce consciousness in us» (P. Churchland, 2007, p. 2). «Consciousness, almost certainly, is not a semimagical glow emanating from the soul or permeating spooky stuff. It is, very probably, a coordinated pattern of neuronal activity serving various biological functions. This does not mean that consciousness is not real. Rather, it means that its reality is rooted in its neurobiology. That a brain can come to know such things as these, and in particular, that it can do the science of itself, is one of the truly stunning capacities of the human brain» (P. S. Churchland, 2002, p. 2).

³⁶ «By “eliminative materialism” I mean the view that holds 1. that folk psychology is a theory; 2. that it is a theory whose inadequacies entail that it must eventually be substantially revised or replaced outright (hence “eliminative”); and 3. that what will ultimately replace folk psychology will be the conceptual framework of a matured neuroscience (hence “materialism”)» (P. S. Churchland, 1986, p. 396).

“mentalistic” language will change in accordance to the new neuroscientific paradigm, aiming to a sort of “neuroscientific experience of the world” with a new neuroscientific vocabulary (P. M. Churchland, 1979, 1989). For example, instead of saying “I don’t believe in beliefs” that would be a pragmatic contradiction for an eliminative materialist, we could say “I gronkify beliefs”, where “gronkification” is a neuropsychological state defined within the mature new theory (P. S. Churchland, 1986, p. 397)³⁷.

3.1.3.3 Dismissal: Elisabeth Irvine’s Scientific Eliminativism

After the proposal, made by the Churchlands, of a radical translation of the folk psychological language into the neuroscientific one, I shall briefly expose another kind of eliminativism, to some extent, completely different from the eliminative materialism. This latter, indeed, moved along the main trend of the previous decades, when the incredible development of cognitive sciences suggested the chance of achieving a complete knowledge of our whole mental life. Not that it is very different now, in fact there are several theories that claim to have the key for explaining with neuroscience every aspects of our conscious experience, but in the recent years some doubts arouse. In this regard, I refer to Elisabeth Irvine’s scientific eliminativism. According to her, there are intrinsic and chronic problems for scientific research about consciousness. There are issues with the definition and the empirical determination of the phenomenon, since, as already seen above, it goes beyond the many specific ways it can be operationalised. In her book *Consciousness as a scientific concept. A philosophy of Science Perspective* and some other articles Irvine sustains that there is not any appropriated method for consciousness (Irvine, 2013b), because there is a lack of convergences between different measurements, indeed it is not clear whether the results of several methods (e.g. phenomenology, introspection, qualitative differences

³⁷ The following is just an anecdote, but still useful for understanding the Churchlands argumentation: «The neurophilosopher Paul Churchland “carries in his wallet a colour picture of his wife. Nothing surprising in it,” remarks the sociologist Bruno Latour, “except it is the colour scan of his wife’s brain! Not only that,” he continues, “but Paul insists adamantly that in a few years we will all be recognizing the inner shapes of the brain structure with a more loving gaze than noses, skins and eyes!” (Latour, 2004, p. 224)» (F. Vidal & Ortega, 2017, pp. 13–14).

and dissociation methods, neurophysiological and behavioural methods, etc.) refer to a single common phenomenon (consciousness as such) or to a variety of conscious phenomena (reportability, confidence ratings, pre-attentional processing, etc.)³⁸: «The methodological problems associated with failures to identify anything like a good candidate mechanism for consciousness again suggest that ‘consciousness’ is not a viable target for scientific research» (Irvine, 2013a, p. 91). Due to underspecification, i.e. failing to describe the multifaceted character and conditions of the phenomenon, taxonomic errors, i.e. unappropriated categorization of a set of phenomena, and misidentification as a scientific kind, Irvine sustains that there is no room for a scientific clarification of the phenomenon, rather she advocates the elimination of the concept from the scientific vocabulary (clarification through elimination)³⁹. This “scientific eliminativism” is different from eliminative materialism and all the other eliminativist positions (see the Deniers and also Dennett) that argues for the non-existence of phenomenal consciousness, i.e. entity eliminativists (Sprevak & Irvine, forthcoming), rather it falls under the category of discourse eliminativism, which denies «the utility of talking about phenomenal (and perhaps access) consciousness in science» (Sprevak & Irvine, forthcoming)⁴⁰.

³⁸ Moreover, «[i]n consciousness science there are also significant differences across production, measurement and detection procedures in experimental protocols for investigating consciousness. It is clear from the above that local clusters of convergence between sets of measures can be observed. However, the convergence that exists within consciousness science can be identified as the product of investigating similar sets of sensory, cognitive or neural phenomena across groups of experimental paradigms. These distinct phenomena are already the subjects of investigation in cognitive and neuroscience. Further, there seem to be no reasons for viewing these phenomena as different types of a common phenomenon» (Irvine, 2013a, p. 90).

³⁹ «However, the claim is not that consciousness does not exist (whatever that might mean), but that it does not constitute a viable scientific concept. An examination of the range of operationalizations, measures and mechanisms of consciousness show that there is nothing in common between them all, and that instead they refer to a wide range of different phenomena and mechanisms already described in cognitive science. While it is possible to label these different phenomena as subtypes of consciousness, the practical and explanatory benefits of having two sets of labels of the same phenomena, one comprised on the vocabulary of perceptual and cognitive abilities, and one with the word ‘consciousness’ added, is of little practical or explanatory benefit. [...] while eliminativism is often seen as a purely negative stance, scientific eliminativism is both necessary to the progress of science, and must also offer a positive contribution in terms of providing conceptual clarification, explanations of current results, and future research questions» (Irvine, 2013a, pp. 154, 172–173).

⁴⁰ «[...] a claim for the elimination of consciousness from science would not necessarily cause any problems for the folk use of the term, as the folk use runs roughly in track with scientific accounts of cognitive/mental phenomena. Interestingly, this may put more pressure on philosophical concepts of consciousness as they can be attacked on two sides; from the cognitive sciences and from experimental philosophy» (Irvine, 2013a, p. 168).

3.2 The Roots of Neuro-centrism

So far, I have exposed how the concept of consciousness has been taken into account by neuroscience, through several operationalizations and empirical quantifications, particularly the (not yet completed) search for the neural correlates of consciousness. Then, I described three examples of drastic scientific approaches towards consciousness, which respectively claim for the reduction of the FFP into the TPP, the transmutation of the concept of consciousness towards the neuroscientific paradigm, or the elimination of its usage in the neuroscientific vocabulary. It is not my intention to deal with these theses in detail, even though it is possible to find some possible counterargument in the two previous chapters (e.g. about the impossibility of complete denials of the concept of consciousness) and in the next chapter. My aim here is to argue that these positions, although they are extreme and minority compared to the neuroscientific trend, can be considered the pinnacle of a way of thinking deeply rooted in the whole neuroscientific field.

It is no secret that cognitive science and neuroscience are based, ontologically, epistemologically and methodologically⁴¹, upon the naturalistic paradigm – in particular the reductive version – derived from the American philosophical debates in the first half of the 20th century, and that can be described – very generally – by these principles: ontological primacy of materialism, universal validity of the scientific method, unicity of knowledge, physical causal closure and exclusion of any supernatural entity (Laudisa, 2014; Papineau, 2016)⁴². And the very same concept of

⁴¹ «Ontological naturalism is the view that ontology (i.e., the complete inventory of reality) is exhausted by the entities and properties invoked by scientific theories. Epistemological naturalism is the view that epistemological questions (e.g., what can we know or what is the nature of knowledge or justification) are to be investigated empirically by, for example, scientific accounts of our cognitive faculties. Methodological naturalism is the view that the methods of philosophy should be restricted to the methods of science. A corollary of methodological naturalism is explanatory naturalism, the view that all genuine explanations are scientific explanations. Any and all of these kinds of naturalism fall under the rubric “scientific naturalism”» (Baker, 2013, p. 5).

⁴² «The term “naturalism” is applied to several different doctrines. In its broadest sense, it may amount to no more than the premise that philosophical theories should not appeal to anything irrevocably mysterious. However, it has also acquired a more specific meaning, which incorporates both epistemological and metaphysical theses: 1 Epistemological: The standpoint and method/s of empirical science comprise the best way to acquire knowledge of every aspect of the world, including ourselves. 2 Metaphysical: The world is comprised solely of the kinds of objects, properties, and causal relations

consciousness frequently used in the neuroscientific field depends on the naturalistic assumptions. Contemporary naturalism is supposed to be the exact opposite of dualism, regarding current studies in cognitive science, in particular between philosophy and neuroscience. Indeed, the key principles showed here above automatically imply the rejection of any kind of dualism, in favour of a monistic view of the world and man. However, I shall argue that it is still possible to find conceptual traces of dualism (in particular from Descartes' doctrine) in the alleged anti-Cartesian naturalism. This analysis will be useful for a better understanding of the meaning of the concept of consciousness.

Thus, before defining any new possible approach – philosophical or neuroscientific – towards consciousness and consequently trying to answer the “what/how/why” questions above mentioned – I believe we need to address first the issue about the epistemological foundations upon which consciousness is generally approached in contemporary philosophy and neuroscience. After having offered a starting definition of consciousness (second chapter) but before giving any new empirical or phenomenological attempt of explanation (fourth and fifth chapters), some clarification about the origin of the scientific concept of consciousness are needed. I already offered a provisional description of the term “consciousness” we use in the previous chapter, fixed in its logical and phenomenological meaning, but still deficient in terms of empirical and phenomenal support. Now, despite exposing here an exhaustive history of the concept of consciousness is almost impossible, I shall highlight some historical and cultural passages that can be useful in order to understand how we have come to interpret it these days, and, if possible and where applicable, I shall try to identify some common genealogical features among different approaches, e.g. common sense, phenomenology, neuroscience and philosophy of mind. As Max Velmans said, «once a given reference for the term ‘consciousness’ is fixed in its phenomenology, the investigation of its nature can begin, and this may in time transmute the meaning (or sense) of the term» (Velmans, 2009, p. 152). And – I would add – in order to give a well-formed phenomenology in which we can fix the

posited by scientific theories. Acceptance of these two theses fuels the project of exhaustive “naturalization”» (Ratcliffe, 2006, pp. 330-331).

reference for “consciousness”, it would be useful to make some clarification about the concepts used to determine the phenomenon and the act of conscious experience.

3.2.1 The Brainhood condition

We already seen how the concept of consciousness nowadays seems intertwined with the concept of brain as the centre of our mental activity. Apart from any technical consideration about the relation between consciousness and the brain, like the debate about the possibility of an identity relation versus other softer alternatives (correlations, contingent identity, etc.), it is already possible to notice how the brain-consciousness relation (whatever it might be) is rooted in our common way of thinking. For example, if I have to undergo a heart surgery, my first concern would be about the healing and recovery process, instead, if I have to undergo a brain surgery, beyond the former concerns I would probably hope that the operation will not cause changes in my personality or cognitive abilities⁴³. Also, even in particular cases in which damages to the body, e.g. the amputation of an arm, provoke also cognitive impairments, e.g. the phantom limb pain, it seems possible to identify the wrong reconfiguration of the brain as the cause of the syndrome (Kaur & Guan, 2018).

The general belief about the essentiality and centrality of our head in the determination of our mental life, subjective characteristics and personhood can be defined as “brainhood”. Brainhood names the ontological condition of “being” (and not simply having) a brain (Fernando Vidal, 2009), as well as personhood with the condition of being an individual person, and it refers to an anthropological figure of modernity named by Fernando Vidal the “cerebral subject” (F. Vidal & Ortega, 2017; Fernando Vidal, 2009). Indeed, it seems self-evident – at least apparently – that we are our brain and if the brain loses its normal features and activity, we may lose consequently our personhood and individuality. The intuitive power of thought experiments like the brain in a vat hypothesis (Putnam, 1981) or the brain transplant

⁴³ Moreover, as already said, brain trauma can cause prolonged losses and disorders of consciousness, neurodegenerative diseases like Alzheimer’s provoke changes in personalities (Robins Wahlin & Byrne, 2011), as well as brain tumours can cause also psychiatric symptoms (Madhusoodanan et al., 2015).

(Shoemaker, 1963) confirms that general belief⁴⁴; without going into details about the logical coherence of the brain in a vat scenario (a modern reinterpretation of the Cartesian evil genius) and entering the debate about scepticism and realism, it is interesting that the brain is taken as the only necessary case of our whole existence, so much that not only the environment but also the rest of body is left out. Thus, in this respect, it is completely normal that nowadays we are talking about the possibility in the very next future to perform a human head transplant (cephalosomatic anastomosis), i.e. the surgical transference of a healthy head on a surgically beheaded body under deep hypothermic conditions (Canavero, 2015). It is generally said that a transplant is a medical operation in which a new organ – a part – is transferred into someone's body – the whole system – that needs that new organ, in this case, instead, the head – the part – is said to be transplanted, while actually it is the rest of the body that is removed first (of the old/ill one) and then substituted with the new one. Basically, the operation consists of preservation of about the 20% of the body (the head) against the replacement of the remaining 80% of it. If we look at this operation as the substitution of body parts, as it is usually understood, we may talk about body transplant, however in this case, rather unusually, the most important thing is the preservation of a little part of the body, i.e. the head, and therefore it is not the body that is transplanted but the head, shifted from a “host body” to another⁴⁵. Thus, the brainhood condition is strongly rooted into the contemporary scientific context but also into our folk psychological and conceptual structures. In this sense, concepts like mind, consciousness, but also personhood and individuality have been inevitably related to that of the brain, i.e. the mind-brain relation/problem. Now, the important thing to realize how this strong relationship has emerge in both cultural and scientific contexts.

⁴⁴ «Consider the hypothesis that you are a disembodied brain floating in a vat of nutrient fluids. This brain is connected to a supercomputer whose program produces electrical impulses that stimulate the brain in just the way that normal brains are stimulated as a result of perceiving external objects in the normal way. (The movie ‘The Matrix’ depicts *embodied* brains which are so stimulated, while their bodies float in a vats.) If you are a brain in a vat, then you have experiences that are qualitatively indistinguishable from those of a normal perceiver. If you come to believe, on the basis of your computer-induced experiences, that you are looking at a tree, then you are sadly mistaken» (McKinsey, 2018).

⁴⁵ «Reacting to the thought-experiment version of the above-quoted formula – ‘If the brain of A could be transplanted into the body of B, then it is not B who would receive a new brain, but A who would gain a new body’ – leading neuroscientist Michael Gazzaniga (2005: 31) commented: ‘This simple fact makes it clear that you are your brain.’ Neuroscientists’ writings and interviews for general audiences, media discussions of neuroscientific research, and the vast neurocultural constellation we shall deal with below offer countless variations of such a claim» (Fernando Vidal, 2009, p. 6; Gazzaniga, 2005).

3.2.2 Pre-modern concepts of Brain and Consciousness

As said, it is not possible to discuss here a fine-grained exposition of the history of consciousness, nevertheless I shall describe some points of interest which I believe are useful for understanding how nowadays we connect the mental domain (self, consciousness, mind, emotions, memories, etc.) to the brain in an automatic way. The idea that the brain has an essential role in our cognitive and mental life in general was already present in ancient Greece, in fact a V century BC text attributed to Hippocrates refers to the brain as the one and only origin point of our own thoughts and emotions⁴⁶, so that Hippocrates can be considered not only the father of medicine but also the forefather of neurology (Breitenfeld, Jurasic, & Breitenfeld, 2014)⁴⁷. Thus, the idea that brain and mind are reciprocally related is not recent, indeed also Plato thought that the brain, together with heart and liver, was one of the seats of activities of the soul (Plato, *Timaeus*)⁴⁸. Notwithstanding this, it would be inaccurate to claim that people from ancient Greece had the same concept of mental life and consciousness we have now. Indeed, if on the one hand the brain has been related very soon to mental activity, on

⁴⁶ «Men ought to know that from nothing else but the brain come joys, delights, laughter and sports, and sorrows, griefs, despondency, and lamentations. And by this, in an especial manner, we acquire wisdom and knowledge, and see and hear, and know what are foul and what are fair, what are bad and what are good, what are sweet, and what unsavory; some we discriminate by habit, and some we perceive by their utility. By this we distinguish objects of relish and disrelish, according to the seasons; and the same things do not always please us. And by the same organ we become mad and delirious, and fears and terrors assail us, some by night, and some by day, and dreams and untimely wanderings, and cares that are not suitable, and ignorance of present circumstances, desuetude, and unskilfulness. All these things we endure from the brain, when it is not healthy, but is more hot, more cold, more moist, or more dry than natural, or when it suffers any other preternatural and unusual affection. And we become mad from its humidity. For when it is more moist than natural, it is necessarily put into motion, and the affection being moved, neither the sight nor hearing can be at rest, and the tongue speaks in accordance with the sight and hearing. [...] In these ways I am of the opinion that the brain exercises the greatest power in the man. This is the interpreter to us of those things which emanate from the air, when the brain happens to be in a sound state. But the air supplies sense to it. And the eyes, the ears, the tongue and the feet, administer such things as the brain cogitates. For in as much as it is supplied with air, does it impart sense to the body. It is the brain which is the messenger to the understanding. For when the man draws the breath into himself, it passes first to the brain, and thus the air is distributed to the rest of the body, leaving in the brain its acme, and whatever has sense and understanding. For if it passed first to the body and last to the brain, then having left in the flesh and veins the judgment, when it reached the brain it would be hot, and not at all pure, but mixed with the humidity from flesh and blood, so as to be no longer pure» (Walshe III & Walshe, 2016).

⁴⁷ However, it must be said that the first historical document that describes the brain in relation with cranial injuries is the Edwin Smith Papyrus, dated around 2500 BC (Finger, 2005a).

⁴⁸ Differently, Aristotle claimed that the brain was a sort of cooling mechanism for the blood, because of its form, while that the heart was the centre of sensation and movement (Aristotle, *De Anima*).

the other hand it cannot be taken for granted that concepts like “mind”, “consciousness” and “self” have remained unchanged for centuries or millennia.

To give an example, according to Julian Jaynes, the analysis of the psychological vocabulary used in the Homeric poem *Iliad* shows an image of the human mental life quite far from the contemporary one (Jaynes, 1976). Indeed, Jaynes argues that in ancient Greece any psychological term was related to physical features where the human seems to be moved not by an ego or a conscious will but by body parts (e.g. “the heart is eager to fight”), while the inner speech was interpreted as the voice of the gods, very far from our concept of inner voice as the expression of our own autonomous subjectivity⁴⁹. Regarding this, Jaynes develops the theory of the bicameral mind, according to which, the mind in the ancient history was split in two parts: explicit speech and the guiding voices. Progressively, the mind developed an autonomous thought through the dismissal of the heterogeneous voices (the breakdown of the bicameral mind) and finally giving rise to consciousness, which, therefore, according to Jaynes is not a biological product of evolution present in animal species others than the human one, rather it is a learned process that strongly depends on linguistic and social constructions, so that its development is not yet finished.

⁴⁹ «There is in general no consciousness in the *Iliad*. [...]. And in general therefore, no words for consciousness or mental acts. The words in the *Iliad* that in a later age come to mean mental things have different meanings, all of them more concrete. The word *psyche*, which later means soul or conscious mind, is in most instances life-substances, such as blood or breath: a dying warrior bleeds out his *psyche* onto the ground or breathes it out in his last gasp. The *thymos*, which later comes to mean something like emotional soul, is simply motion or agitation. When a man stops moving, the *thymos* leaves his limbs. But it is also somehow like an organ itself, for when Glaucus prays to Apollo to alleviate his pain and to give him strength to help his friend Sarpedon, Apollo hears his prayer and “casts strength in his *thymos*” (*Iliad*, 16:529). The *thymos* can tell a man to eat, drink, or fight. Diomedes says in one place that Achilles will fight “when the *thymos* in his chest tells him to and a god rouses him” (9:702f). But it is not really an organ and not always localized; a raging ocean has *thymos*. [...] The characters of the *Iliad* do not sit down and think out what to do. They have no conscious minds such as we say we have, and certainly no introspections. It is impossible for us with our subjectivity to appreciate what it was like. When Agamemnon, king of men, robs Achilles of his mistress, it is a god that grasps Achilles by his yellow hair and warns him not to strike Agamemnon (I:197ff.). [...] There is also no concept of will or word for it, the concept developing curiously late in Greek thought. Thus, *Iliadic* men have no will of their own and certainly no notion of free will. [...] It is one god who makes Achilles promise not to go into battle, another who urges him to go, and another who then clothes him in a golden fire reaching up to heaven and screams through his throat across the bloodied trench at the Trojans, rousing in them ungovernable panic. In fact, the gods take the place of consciousness. The beginnings of action are not in conscious plans, reasons, and motives; they are in the actions and speeches of gods. To another, a man seems to be the cause of his own behavior. But not to the man himself» (Jaynes, 1976, pp. 69–70, 72; for other details, Di Francesco, Marraffa, & Tomasetta, 2017, pp. 39–43).

This kind of historical and anthropological analysis could be useful in order to highlight that the term “consciousness” as we know, use and ascribe to persons is not an original and primitive concept, but depends also on conceptual, epistemological and – more generally – cultural mutations in the history of philosophy and science. Nevertheless, the risk here is to infer from a cultural, linguistic, sociological argument – i.e. the fact that ancient Greeks did not have concepts about mental and conscious life – to an ontological argument – i.e. the fact that ancient Greeks did not have consciousness or a mental life as we have now (Wilkes, 1988). Neither Plato nor Aristotle nor the Greek tragedians used concepts like psyche (*Ψυχή*) as a perfect synonyms of the modern term “consciousness”, however this does not mean that they did not possess their own self-awareness or that they did not experience the world in a similar way as we do (Perconti, 2011)⁵⁰. Therefore, I would consider this analysis on the origin of consciousness from a conceptual perspective, rather than from an ontological one, looking for phases and shifts in the cultural and scientific dominion that could be useful to understand the historical evolution of (the concept of) consciousness and its relation to the brain.

So far, I have tried to show how the concept of mind, with all the other correlated concepts, immediately established a relationship with that of the brain, and this can be seen in both medical (Hippocrates) and philosophical (e.g. Plato) contexts. After that, the attempt to understand the functional structure of the brain gave birth to several theories, among which the theory of ventricular localization of the mental faculties⁵¹ attributed to Galen or to some of his scholars (Green, 2003), that remained one of the most influential until at least the XVI century. It is impossible to enter into the detail of the evolution of the science of the brain (for a historical instruction through the exposition of the forefathers, pioneers and more recent founders of neuroscience, see Finger, 2005b), and despite Ancient and the Medieval history are inevitably relevant for the understanding of our conceptual and theoretical roots, I shall claim that the very turning point to consider for the analysis of the concept of consciousness

⁵⁰ Moreover, it seems that dualism is not merely a product of Western philosophy, rather it seems to be rooted also in people and cultures that do not make explicit use of dualistic language and beliefs (e.g. subject-object, spirit-matter, man-nature, etc.) (Astuti, 2001). Thus, someone may speak of a sort of “natural”, “intuitive” or “implicit” dualism (Bloom, 2004; Forstmann & Burgmer, 2015; Perconti, 2011).

⁵¹ The theory sustains that sensation, cognition, memory, etc. are located in specific parts of the three ventricles in the brain.

is the Modern era, especially through two key figures, Galileo Galilei and René Descartes.

3.2.3 Modernity as the Cradle of the contemporary concept of Consciousness

The Modern era is not simply relevant for understanding the conceptual transition and mutation of consciousness, the change we are now going to take in consideration are relative to the whole structure of scientific research and philosophy as we know it today. In the XVII century, Galileo Galilei and René Descartes respectively laid the foundations for the scientific method and modern philosophy. This caused the reconceptualization – or even more the emergence – of the relation between subjectivity and objectivity, producing a conceptual distinction that turned out useful both for scientific progress and philosophical research. Galilei gave birth to the modern scientific method, consequently producing a reformulation of the relation between subject and object. According to him, a theory is scientific if founded on “sense experiences and necessary demonstrations” (Galilei, 1615). The sense experience refers to the empirical perception of the world, against any aprioristic dogmatism on natural facts, while, the necessary demonstration is the elaboration of the experiential observations within a mathematical framework, in order to avoid any uncontrolled and non-rigorous empirical collection. So, the all-new science (*Novissima scienza*) became an interdependent integration between experimental data and mathematical models, and this integration is theoretically unlimited, towards an increasingly perfectible approximation (Zanin, 2011). In other words, with the entry of the mathematical framework within the empirical collection of data, the natural world becomes progressively measurable, reproducible and thus predictable.

In order to do so, Galilei overturned the classic Aristotelean opposition between what today we would call subjective and objective properties, eliminating the former from the scientific domain. Focusing on perceptiveness, Aristotle distinguishes between the “proper” or “exclusive” sensible, and the “common” sensible⁵²: the

⁵² «Perceptible objects are spoken of in three ways: in two cases we say perceptible objects are perceived in their own right, and in one co-incidentally. Of the first two, one is exclusive to an individual sense and the other common to them all. By exclusive I mean what cannot be perceived by another sense and about what one cannot be deceived. For example, sight is of colour, hearing is of sound, and taste is of

former is the set of objective properties really possessed by the perceived thing, which can be directly sensed – e.g. colour, sound, taste, etc.; the latter is that kind of properties that can be indirectly sensed by more than one sense together – e.g. size, motion, weight, unity, number, figure – and these, due to their indirect derivation, they are capable of subjective error⁵³ (Aristotle 2001, B6, 418a 16-19). Galilei overturns this opposition and what he calls “primary” or “real” “qualities” or “accidents” – Aristotle’s common sensibles – become objective properties of the world, while what he calls “secondary qualities” or “affections” – Aristotle’s exclusive sensibles – become products of the “sensitive body” (from the Italian “*nel corpo sensitivo*” (Galilei, 1623, V; Galilei, 2017, p. 197) and not “in consciousness” as in most English translations; see Buyse, 2015)⁵⁴. Hence, without the (bodily) perceiving and subjective act of knowledge, secondary qualities do not truly exist, and, in this way, Galileo

flavour, whereas touch has a number of different objects. In any case, each sense discerns these and is not deceived that there is colour or that there is sound-as opposed to what or where the coloured or sounding thing is. Accordingly, these sorts of objects are said to be exclusive to each sense. Common objects include motion, rest, number, shape, and magnitude, since these sorts of objects are exclusive to no one sense but are, rather, common to them all. For in some cases movement is an object of perception for both touch and sight. [...] Moreover, there cannot be some special sensory organ for the common objects, which we perceive co-incidentally by each sense for example, motion, rest, shape, magnitude, number, and unity. For all of these we perceive by motion: for example, we perceive magnitude by motion (and consequently, shape, since shape is a sort of magnitude); something at rest in virtue of its not being moved; number in virtue of the lack of continuity as well as by the peculiar objects of perception (since each sense perceives one thing). Consequently, it is clear that there cannot be any special sense for these, for instance, for motion. For in this way things will be just as they are now when we perceive the sweet by sight: this occurs because we as a matter of fact have a perception of them both, by which we come to know them at the same time when they occur together. Otherwise, we would never perceive such things other than co-incidentally» (Aristotle, *De Anima*, 418a 8-19, 425a 13-25).

⁵³ «[P]erception of exclusive objects is true, or is subject to falsity in the smallest degree. Second, perception is of something's being an attribute of something; and already here it is possible to be mistaken. For one is not mistaken that there is white; but if one says that this or that other thing is what is white, one is mistaken. Third, perception is of the common objects which follow upon the attributes to which the exclusive objects belong (I mean, for instance, motion and magnitude, which are attributes of the objects of perception), concerning which there is already, most of all, deception in the realm of perception» (Aristotle, *De Anima*, 428b 18-25).

⁵⁴ «Now I say that whenever I conceive any material or corporeal substance, I immediately feel the need to think of it as bounded, and as having this or that shape; as being large or small in relation to other things, and in some specific place at any given time; as being in motion or at rest; as touching or not touching some other body; and as being one in number, or few, or many. From these conditions I cannot separate such a substance by any stretch of my imagination. But that it must be white or red, bitter or sweet, noisy or silent, and of sweet or foul odor, my mind does not feel compelled to bring in as a necessary accompaniment. Without the senses as our guides, reason or imagination unaided would probably never arrive at qualities like these. Hence I think that tastes, odors, colors, and so on are no more than mere names so far as the object in which we place them is concerned, and that they reside only in the sensible body. Hence if the living creature were removed, all these qualities would be wiped away and annihilated. But since we have imposed upon them special names, distinct from those of the other and real qualities mentioned previously, we wish to believe that they really exist as actually different from those» (modified translation) (Galilei, 1623/1957, p. 274).

voluntarily limited the domain of scientific knowledge, excluding from the natural world secondary qualities (Galilei 1623/1995). Or better yet, according to the Galileian paradigm, the only way for the pursuit of knowledge to be “scientific” – i.e. rigorous and objective observation towards measurable and reproducible phenomena – was to exclude all the subjective experiences from its research program, consequently leading to the omission of the secondary qualities from the quantitative knowledge. As Hume claimed, this distinction between primary and secondary qualities is the “fundamental principle” of modern philosophy against the Scholastic-Aristotelean tradition⁵⁵, and it can be traced, in different versions, in many authors of the same time period, like Boyle, Newton and Locke (Thompson, 1994)⁵⁶; even more, for Hobbes subjective appearances are nothing but in the brain⁵⁷.

Regarding the “necessary demonstrations”, Galilei connected mathematics to physics, setting the boundaries of the knowable within what is susceptible to mathematical correction, in order to reach a progressively perfectible mind-world

⁵⁵ «The fundamental principle of that philosophy is the opinion concerning colours, sounds, tastes, smells, heat and cold; which it asserts to be nothing but impressions in the mind, deriv'd from the operation of external objects, and without any resemblance to the qualities of the objects. Upon examination, I find only one of the reasons commonly produc'd for this opinion to be satisfactory, *viz.* that deriv'd from the variations of those impressions, even while the external object, to all appearance, continues the same. [...] For upon the removal of sounds, colours, heat, cold, and other sensible qualities, from the rank of continu'd independent existences we are reduc'd merely to what are called primary qualities as the only real ones, of which we have any adequate notion» (Hume, 1738/2007, pp. 149–150).

⁵⁶ «[I]f we should conceive all the rest of the universe to be annihilated, save one such body - suppose a metal or a stone - it were hard to show that there is physically anything more in it than matter and the accidents we have already named [figure, shape, motion, and texture]» (Boyle, 1666/1991, p. 30). «These I call original or primary qualities of body, which I think we may observe to produce simple ideas in us, *viz.* solidity, extension, figure, motion or rest, and number. Secondly, such qualities which, in truth, are nothing in the objects themselves, but powers to produce various sensations in us by their primary qualities, i.e. by the bulk, figure, texture, and motion of their insensible parts, as colours, sounds, tastes, &c., these I call secondary qualities» (Locke, 1689/1838, p. 75). «For the rays, to speak properly, are not coloured. In them there is nothing else than a certain power and disposition to stir up a sensation of this or that Colour [...] Colours in the Objects are nothing but disposition to reflect this or that sort of rays more copiously than the rest. in the rays they are nothing but disposition to propagate this or that Motion into the Sensorium, and in the Sensorium they are Sensations of those Motions under the forms of Colours» (Newton, 1704/1730, pp. 108–109).

⁵⁷ «[W]e are to consider, first that upon every great agitation or concussion of the brain (as it happeneth from a stroke, especially if the stroke be upon the eye) whereby the optic nerve suffereth any great violence, there appeareth before the eyes a certain light, which light is nothing without, but an apparition only, all that is real being the concussion or motion of the parts of that nerve; from which experience we may conclude, that apparition of light is really nothing but motion within. [...] namely, that image and colour is but an apparition to us of that motion, agitation, or alteration which the object worketh in the brain or spirits, or some internal substance in the head» (Hobbes, 1650/1811, pp. 5–6).

correspondence⁵⁸. And for “scientific” he meant a theory that has solid method and criteria, which permits to reach a God-like intensive knowledge – not God-like extensive though – about mathematical truths (Machamer, 1998). In mathematical correspondence model, the reason makes questions to the nature in mathematical language and the nature answers correctly if the questions correspond to the reality⁵⁹. This is possible only with an idealization of the nature, that is the approximation of the real event into a non-subjective model-event. In this sense, as Husserl said, the objectivity of reality is not reached but presupposed. And, as we will see later, after Galileo even the subjective properties will return as indirect object of mathematization, by means of the reduction of psychological data in composition of common observations⁶⁰. Thus, Galilei can be considered one of the key figures of the modern turn in philosophy and science, with the mathematization of the natural world and the separation between primary and secondary qualities (Husserl, 1970b; Israel, 1996)⁶¹.

The second key figure I shall briefly expose is Descartes, who distinguished, from both the epistemological and ontological domains, between a physicalist objectivism and a rational subjectivism: the former entails a mechanistic view of the world – not only the external world but also all human functions (sensorial and motor processes, memory, digestive system) – except for the conscious thinking (Descartes, 1637/2003). Indeed, the latter is not part of the material world, nevertheless regards

⁵⁸ «For Platonism, the real had a more or less perfect methexis in the ideal. This afforded ancient geometry possibilities of a primitive application to reality. [But] through Galileo’s mathematization of nature, nature itself is idealized under the guidance of the new mathematics; nature itself becomes — to express it in a modern way — a mathematical manifold» (Husserl, 1936/1970, p. 23).

⁵⁹ Interestingly enough, despite Galilei was a convinced empiricist, the Platonic influence has been decisive in him for the mathematization project of the world and the subdivision into primary and secondary qualities (De Caro, 2017).

⁶⁰ «If we adhere strictly to Galileo’s motivation, considering the way in which it in fact laid the foundation for the new idea of physics, [...] we must ask, accordingly, how he could hit upon this conception, namely, that everything which manifests itself as real through the specific sense-qualities must have its mathematical index in events belonging to the sphere of shapes—which is, of course, already thought of as idealized—and that there must arise from this the possibility of an indirect mathematization, in the fullest sense, i.e., it must be possible (though indirectly and through a particular inductive method) to construct *ex datis*, and thus to determine objectively, all events in the sphere of the plena» (Husserl, 1936/1970, p. 37).

⁶¹ «Whereas the primary characteristics are intersubjective as long as harmonious identification can take place at all, and have only [the above-mentioned] relativity of perfection and approximation (and also the relativity of interreal causality), the secondary characteristics are relative in another way, namely, relative to the normality and abnormality of experience and thus “accidentally” to the subjects, and can change with the subjects. In addition, the perfection-limit of the secondary qualities is not measurable; it is only “intuitable.” But it is intersubjectively determined and determinable through relation to the mathematical limits of the primary characteristics» (Husserl, 1970b, pp. 309–310).

the rigorous and deductive thought process that forms the foundation of the scientific knowledge⁶². As such, the study of mind was excluded from the horizon of the physical world. This entails the famous separation between *res extensa* and *res cogitans*, which in turn leads to the mind-body problem, i.e. the issue regarding the relationship between mental properties and physical properties⁶³. Even Descartes proposed a separation similar to that between primary and secondary qualities, where the former refers to “modes” of the *res extensa*, e.g. size, shape, motion, position, etc., while the latter are derived from the interaction between *res extensa* and *res cogitans*, e.g. colours, smells, sounds, hardness, pleasure, emotions, etc., which are not in the natural things⁶⁴ and nevertheless are not illusory or superfluous, because they are the natural product of the mind-body union (Ortín Nadal, 2019). So, the primary/secondary divide comes from the split of reality into “cogitative things” and “material things” (Descartes, 1641/2003), however secondary qualities are not purely immaterial mind-dependent entities, but they «originate from the close and profound union of our mind with the body»⁶⁵. Besides, the clear-cut separation between body and mind is called into question by Descartes himself⁶⁶.

⁶² This is one of the reasons why Descartes’ ideal of a *mathesis universalis* as queen of sciences is different from Galileo’s view of mathematics as experimental tool for the quantitative understanding of nature (Israel, 1996).

⁶³ Differently from Thomas Aquinas, who defined the separation between spirit and nature in human beings as the intellect-sense divide, Descartes redrew the main separation as a division between the body-machine and the mind-*cogitatio*, and this latter including also sense-perception, volition and emotion.

⁶⁴ «And certainly from the fact that I am sensible of different sorts of colours, sounds, scents, tastes, beat., hardness, etc., I very easily conclude that there are in the bodies from which all these diverse sense-perceptions proceed certain variations which answer to them, although possibly these are not really at all similar to them» (Descartes, 1641/2003, p. 114, VI meditation).

⁶⁵ «I do not recognize more than two principal kinds of things: one is intellectual or cogitative things, that is, things pertaining to the mind or to thinking substance; and the other, material things, or things pertaining to extended substance or body. Perception, volition, and all modes of perceiving and willing pertain to thinking substance; while size (or extension in length, width and depth), figure, motion, situation, divisibility of its parts, and such, pertain to extended substance. However, we also experience in ourselves certain other things which should be attributed neither solely to the mind nor solely to the body, and which, as I shall show later in the proper place, originate from the close and profound union of our mind with the body: specifically, the appetites of hunger, thirst, etc.; and similarly the emotions or passions of the soul (which do not consist solely in thought), for example, the emotions of anger, merriment, sadness, love, etc. ; and finally all sensations, such as pain, pleasure, light and color, sounds, odors, tastes, heat, hardness, and the other tactile qualities» (Descartes, 1644/1982, pp. 21–22).

⁶⁶ «Nature likewise teaches me, through these very feelings of pain, hunger, thirst, and so forth, that I am not present in my body only as a pilot is present in a ship, but that I am very closely conjoined to it and, so to speak, fused with it, so as to form a single entity with it. For otherwise, when the body is injured, I, who am nothing other than a thinking thing, would not feel pain as a result, but would perceive the injury purely intellectually, as the pilot perceives by sight any damage occurring to his ship; and when the body lacks food or drink, I would understand this explicitly, instead of having confused

With Descartes, the natural world is conceived as a mechanic machine, ruled by rational principles understandable by means of rigorous law of physics. Descartes developed an ontological separation between the world-object and the mind-subject, however what concerns our experience emerges from the interaction of these two categories. For this reason, despite they do not entirely fall within the mechanistic framework because they depend also on our mind, secondary qualities can be analysed through a mechanistic approach⁶⁷. Indeed, according to him the mind – or the soul – is not simply juxtaposed to the body, rather it is in constant contact and communication with the body⁶⁸, in particular with the brain⁶⁹. In the brain – and this is perhaps the most famous fact in the philosophy of Descartes – there is the pineal gland (now called *epiphysis cerebri*) that in the eyes of the French philosopher seemed the better candidate for being the strongest point of confluence between immaterial mind and extended matter (Descartes, 1649/2015, § 32). This point is quite interesting since Descartes and his ontological dualism are often considered as the exact opposite of the empirical method of modern science. However, the Cartesian attempt to find the interaction between mental features and body/brain functions can be considered a prototype of the contemporary mind-based approach, perfectly in conformity with the scientific paradigm (Zilio, 2018a). What remains out of the mechanical laws of the material world is the rational and conscious thought, while the rest of the “mechanical

feelings of hunger and thirst. For certainly, these feelings of thirst, hunger, pain, and so forth are nothing other than certain confused modes of thinking, arising from the union and, so to speak, fusion of the mind with the body» (Descartes, 1641/2003, p. 114, VI meditation).

⁶⁷ «Descartes’ plan is to reduce the class of metaphysically suspect properties, such as heat, weight, taste, to the empirically quantifiable attributes of size, shape, and motion. In other words, Descartes intends to replace the “mentally” influenced depiction of physical qualities in Scholastic natural philosophy with a theory that requires only the properties of extension to describe the manifest order of the natural world» (Slowik, 2017).

⁶⁸ For an interpretation of the Cartesian man as and holistic mind-body unity, see Marion, 2018. Also, Merleau-Ponty interpreted the soul-body relation made by Descartes not as a mere juxtaposition but as an (failed) attempt to unify the form (soul) and the matter (body) of the human being (Merleau-Ponty, 1995, chap. 2).

⁶⁹ «We also need to know that, although the soul is joined to the whole body, there is nonetheless a certain part of the body in which it exercises its functions more particularly than in all the rest. This is commonly believed to be the brain, or perhaps the heart; the brain, because the sense-organs are connected with it; and the heart, because that is where we seem to feel our passions. But, having examined the matter with care, it seems to me that the part of the body in which the soul immediately exercises its functions is by no means the heart, or the brain as a whole, but only the most inward part of it, which is a certain tiny gland situated in the middle of the substance of the brain, and suspended over the channel through which the spirits of the anterior cavities communicate with those of the posterior cavity in such a way that the slightest movements that take place within it can do much to change the flow of these spirits; while, conversely, the slightest changes I the flow of the spirits can have major effect on the movements of the gland» (Descartes, 1649/2015, p. 208).

body” can be investigated by physical and geometrical knowledge, so that, from a functional perspective, the whole processes of the human body – except for the rational thought – cannot be distinguished from those of a human-like robot or from a corpse (Descartes, 1641/2003, II meditation). This, of course, went against the Aristotelian and Scholastic hylomorphism that accounts for an indissoluble relationship between the bodily matter and the soul. In this regard, for Aristotle the corpse is different from the body, because the latter is shaped and structured by the form/soul.

Taken together, the attempt to develop a new scientific approach towards the natural world on the basis of a new ontological and epistemological paradigm has also contributed in the establishment of the parallel dimension of the subject, characterized by non-material thinking, resistance to the empirical and quantitative methods, mind-dependence, introspection and first-person perspective, self-ness, individuality, etc. It is easier to understand now how certain contemporary philosophical positions that go against the complete reduction of consciousness come from⁷⁰. The repercussions are visible in the recent debate about the relationship between neuroscience and philosophy, focusing, in particular, on some epistemological and ontological issues, such as the Explanatory Gap Problem, the Knowledge Argument and the Hard Problem of consciousness (Chalmers, 1995; Jackson, 1982; Levine, 1983). In very general terms, all three problems claim that any physical explanation will never be able to account for the emergence of conscious experience. In other words, the target of the argument is the alleged possibility to completely explain consciousness within a pure physicalist perspective and by means of a reductive approach to science. The main difference among the three is the degree of ontological commitment that any single thesis sustains against reductive physicalism: for example, while the explanatory gap tries to prove the epistemic incommensurability between the function explained and the experience felt, the hard problem of consciousness suggests the idea

⁷⁰ «[...] consciousness is a first-person affair, whereas the objective methodology of science is a third-person affair. Opinions, subjectivity, and the like cannot be admitted in scientific experiments. An equally large factor affecting the scientific approach to consciousness can be attributed to the influential thoughts of René Descartes. Sometime after Galileo, Descartes essentially removed the mind from nature. He did this by thought alone, concluding that there were two substances: *res extensa*, extended things that were susceptible to physics, and *res cogitans*, thinking things that were not extended in space and that were unavailable to physics. Descartes’s dualism and its various subsequent derivations have had a profound effect on the approach to consciousness as a valid scientific target» (G.M. Edelman, 2006, pp. 7–8).

that conscious experience involves non-physical properties⁷¹. Hence, theories of qualia, property and substantial dualisms, which claim that there is something that is impossible to catch through the scientific methodology, do not step outside the modern paradigm, because they conceive conscious experience as something extra-physical, extra-natural, therefore referring to the primary-secondary divide, or, more generally, the subject-object division.

3.2.4 From Modern Philosophy to Contemporary Cognitive Sciences

At the same time, the very same scientific and mechanistic paradigm that initially excluded secondary qualities or conscious thought from the scientific approach will be later the basis of following materialist and reductionist approaches towards consciousness. As I just tried to explain, Galilei and Descartes respectively developed the scientific method through mathematization of the world and the mechanistic view of the nature, while both have contributed to the exclusion of subjectivity from the objective domain of the natural phenomena. This offered the epistemological starting point of contemporary naturalism. As Husserl explained in *The Crisis of European Sciences*, Galilei and Descartes developed in different ways the foundation of the naturalistic paradigm that now guides contemporary cognitive sciences (Husserl, 1936/1970). In fact, ontological primacy of materialism, universal validity of the scientific method, unicity of knowledge, physical causal closure, are all naturalistic principles derived from the dualistic division between primary and secondary qualities and exclusion of the subject from the objective rationalism. However, even though there was no room for a quantifying explanation of the mental dimension within the epistemological categories offered by Galilei and Descartes, the development of scientific research, in particular the psychological and neuropsychological ones, has

⁷¹ As Daniel Stoljar argues (Stoljar, 2006, p. 145), most of Chalmers' works are generally based on the "structure and dynamics argument": «First: Physical descriptions of the world characterize the world in terms of [spatiotemporal and causal] structure and dynamics [i.e. changes of dispositions and relations over time]. Second: from truths about structure and dynamics, one can only deduce further truths about structure and dynamics. And third: Truths about consciousness are not truths about structure and dynamics» (Chalmers, 2002, p. 258). Apart from its validity or inconsistency, this argument is perfectly in line with the primary vs secondary divide from modern philosophy.

progressively led to the attempt of approaching the mind and all its features within the scientific paradigm, i.e. through quantification and reduction to physical laws.

The XIX century saw the development of new methods and disciplines for the study of the brain and human behaviour, e.g. psychophysics with Fechner, von Helmholtz, neurophysiology with Ramon y Cajal, Bois-Reymond, Broadmann, Broca, and experimental psychology with Wundt and James, behaviourism with Watson and Skinner (C. Frith & Rees, 2007). It is not my aim to expose here the entire history of the development of psychology and of the birth of neuroscience, rather I shall highlight that what was initially excluded from science, i.e. the mind-dependent dimension, later it has been reintroduced into the scientific dominion through multiple attempts of objectification of the mental features. Quoting Edelman:

«The mind remains well removed from nature. But as Whitehead duly noted [*Science and the Modern World*], the mind was put back into nature with the rise of physiology and physiological psychology in the latter part of the nineteenth century. We have had an embarrassing time knowing what to do with it ever since. Just as there is something special about relativity and quantum mechanics, there is something special about the problems raised by these physiological developments. Are observers themselves “things,” like the rest of the objects in their world? How do we account for the curious ability of observers (indeed, their compelled need) to carve up their world into categories of things-to refer to things of the world when things themselves can never so refer? When we ourselves observe observers, this property of intentionality is unavoidable» (Gerald M. Edelman, 1992, p. 11)⁷².

⁷² This is the passage from Whitehead which Edelman refers to: « The study of mind divides into psychology, or the study of mental functionings as - considered in themselves and in their mutual relations, and into epistemology, or the theory of the knowledge of a common objective world. [...] This is a very uneasy division, giving rise to a host of perplexities whose consideration has occupied the intervening centuries. As long as men thought in terms of physical notions for the objective world and of mentality for the subjective world, the setting out of the problem, as achieved by Descartes, sufficed as a starting point. But the balance has been upset by the rise of physiology. In the seventeenth century men passed from the study of physics to the study of philosophy. Towards the end of the nineteenth century, notably in Germany, men passed from the study of physiology to the study of psychology. The change in tone has been decisive. [...] The effect of physiology was to put mind back into nature. The neurologist traces first the effect of, stimuli along the bodily nerves, then integration at nerve centres, and finally the rise of a projective reference beyond the body with a resulting motor efficacy in renewed nervous excitement» (Whitehead, 1925/1970, pp. 146–147, 148).

Thus, the mind starts to come back gradually in the XIX and – in particular – XX century, in the form of physiological processes, behavioural responses and neuronal activity. With the progress of psychology and neurophysiology the idea of an empirical, i.e. quantifiable and mathematizable, science of mind begins to be no longer a mirage. After the behaviouristic rejection of the mental dimension from science (Watson, 1925; Ryle, 1949)⁷³, during the cognitive revolution (1950s) scientists began to “subdivided” the mind into different features like memory, perception, cognition, attention, emotions, high-order thought, introspection, etc., while nevertheless the word “consciousness” was not frequently used (but still latent in their works) by the first cognitive psychologists in the middle of the XX century (C. Frith & Rees, 2007), so that, recently, someone even spoke of the taboo of consciousness or subjectivity (Revonsuo, 2010; Wallace, 2000)⁷⁴. Owen Flanagan pointed out some reason why consciousness has not been immediately welcome within the paradigm of cognitive sciences (Flanagan, 1991, p. 312): “positivistic reserve”, i.e. the fact that consciousness still refers to a private and ineffable dimension seemed a problem for an objective science of consciousness, “piecemeal/example-driven approach”; i.e. the fact that it seemed better to talk about single parts of consciousness, like memory and attention, instead of the whole concept, it was too embarrassing for scientists; “conscious inessentialism”, i.e. the fact that it seemed possible to have a machine functionally identical to a human being but without consciousness; “epiphenomenalist suspicion”, i.e. the fact that even if consciousness was a part of human mental life, it possibly would have no causal influence⁷⁵.

3.2.5 The Brainhood condition and the Body-Machine model

As I already exposed in the first part of this chapter, despite the doubts raised about the possibility of a science of consciousness, in the last three decades the technological

⁷³ «Behaviorism claims that “consciousness” is neither a definable nor a usable concept; that it is merely another word for the “soul” of more ancient times. The old psychology is thus dominated by a kind of subtle religious philosophy» (Watson, 1925, p. 3). Compare with Irvine’s scientific eliminativism.

⁷⁴ However, as Galen Strawson noted, consciousness never really left the academic scene in the XX century, in particular in the philosophical context, as I also exposed in the first chapter (Strawson, 2017).

⁷⁵ These doubts about the possibility of a reliable science of consciousness resemble the proposal of scientific eliminativism offered by Irvine (see above).

advancement has permitted almost conclusively to talk about the construction of a (neuro)scientific theory of consciousness. The various brain-imaging techniques can increasingly identify instances of relationship between brain activity and consciousness, which does not seem to be found likewise between consciousness and other body parts. Moreover, these techniques give the feeling to go beyond the limits of the observable behaviour, allowing to see what happens “inside the machine”, offering potential counterarguments to behaviourism.

Thus, as showed before, the mind has returned in nature as neural patterns and processes, so that brain and consciousness seem intrinsically related, but this time the relation is even stronger. There is no room for secondary qualities out of the objective dimension of the natural world, rather conscious states like perceptions, thoughts, will, emotions, pleasures and pains are products of the mind, and the mind is a genuine natural phenomenon produced in and by the brain. Everything is brought back to the brain, within the ontological and epistemological paradigm of contemporary naturalism and the empirical methodology of neuroscience. Accordingly, since we could define ourselves as individuals or persons because we possess particular features like self-consciousness, rational thought, free will, emotions, memories, etc., and since those characteristics depend on our neural activities, therefore there seems to be no reason to deny the fact that we are our brains (Caruso & Flanagan, 2018; P. S. Churchland, 2014; Thagard, 2010), i.e. cerebral subjects with the property of being a brain. Theoretically, the mind can be transmuted (Churchland), or reduced (Dennett) into the brain, so that it would be possible to simply change the word “soul” with “brain” in order to make the Lockean scenario of the prince and the cobbler a contemporary topic of discussion in neuroethics:

«The body, as well as the [brain], goes to the making of a man. And thus may we be able, without any difficulty, to conceive the same person at the resurrection, though in a body not exactly in make or parts the same which he had here,—the same consciousness going along with the [brain] that inhabits it. But yet the [brain] alone, in the change of bodies, would scarce to any one but to him that makes the [brain] the man, be enough to make the same man. For should the [brain] of a prince, carrying with it the consciousness of the prince’s past life, enter and inform the body of a cobbler, as soon as deserted by his own [brain], every one sees he would be the same person with the prince, accountable only for

the prince's actions: but who would say it was the same man [human being]? The body too goes to the making the man, and would, I guess, to everybody determine the man in this case, wherein the [brain], with all its princely thoughts about it, would not make another man: but he would be the same cobbler to every one besides himself» (Locke, 1690/1860, p. 213, modified).

Similarly, we can substitute the word “little finger” with “brain” in another Lockean scenario:

Thus every one finds that, whilst comprehended under that consciousness, the [brain] is as much a part of himself as what is most so. Upon separation of this [brain], should this consciousness go along with the [brain], and leave the rest of the body, it is evident the [brain] would be the person, the same person; and self then would have nothing to do with the rest of the body (Locke, 1690/1860, p. 214, modified).

Both the “prince's brain in the cobbler's body” and the “consciousness in the brain” scenarios derived here are based on the aforementioned brainhood condition, in particular they are perfectly in line with the brain-in-a-vat hypothesis or the instance of the head transplant I discussed above. Now, without going into detail about the distinction that Locke made between person and man/human being, what is noteworthy is that, despite the reification of the mind/soul and the substitution of those metaphysical concepts with a more empirically grounded notions (e.g. neural activity, brain, etc.)⁷⁶, the modern paradigm has not changed at all. The contemporary naturalistic paradigm still reveals underlying dualistic structures between consciousness and body, or – to put it otherwise – between brain and body, which implicitly refer to the (alleged) Cartesian but also Lockean mind-body relation⁷⁷. Armstrong's “central materialism theory” is a clear example of a naturalistic paradigm in which, after having rejected any non-physical property, entity or activity, the central

⁷⁶ «Nineteenth-century brain research fulfilled both d'Alembert's and Haller's desiderata: on the one hand, it abandoned the concept of soul and the search for its organ or seat; on the other, it evolved towards increasing technical, descriptive and argumentative sophistication and precision. The bond of brain to self and personhood was thereby confirmed and refined» (Fernando Vidal, 2009, p. 15)

⁷⁷ Fernando Vidal, after the analysis of the two Lockean thought experiments, notes that: «Bodies become things we have, not things we are; personal identity becomes purely psychological, and distinct from bodily identity. From the standpoint of ontological theory, and in comparison with the earlier insistence on the essential corporality of the self, the Lockean theory implied a loss of body» (Fernando Vidal, 2009, p. 14).

nervous system (CNS) has become the central processor unit, where mental states are mechanical functions (D. M. Armstrong, 1968). Thus, the “ghost in the machine” (Ryle, 1949) has become the “brain in the body-machine” (we *are* brains and we *have* bodies) (Zilio, 2018a), therefore, this means that some conceptual features of dualism lie unaltered at the basis of contemporary science⁷⁸, but also of common sense and medicine (Arnaudo, 2017; Joubert, 2014; Ventriglio & Bhugra, 2015).

After the analysis of some of the cultural, medical, historical and philosophical factors that mostly influenced our contemporary presuppositions about the relationship between the mental and the physical, it is easier to understand the neurocentric background which underpinned the rush to the localization of consciousness, starting from phrenology to neurophysiology, and nowadays from cognitive neuroscience up to the search of the neural correlates of consciousness. The next step is to analyse flaws and shortcomings of the neurocentric view of consciousness, and then to propose some epistemological and ontological alternatives to it.

⁷⁸ «Dualism has persisted in various forms to the present day. For example, while apparently monistic, behaviorism is simply dualism reduced by denial of the mind as a scientific object, and therefore left with one end hanging. Behaviorists solve the dilemma by examining behavior and ignoring intentionality. They do not attempt to put the mind back into nature; they simply deny its validity as a scientific object. And many nonbehavioristic psychologists, while asserting that they are materialists and not substance dualists, are nonetheless property dualists. While conceding that the mind and the brain arose from a single substance they insist that psychological properties must be dealt with exclusively in their own terms, which necessarily differ from those used for the physical objects or bodies giving rise to these properties» (Gerald M. Edelman, 1992, pp. 11–12)

4. The Appearance-Reality Fallacy as the Root of contemporary Studies of Consciousness

In this chapter, I proceed with the analysis of the cultural, epistemological and ontological basis on which contemporary researches on consciousness are grounded, including not only empirical studies but also analytical studies from the philosophy of mind. The thesis here is that Cartesianism is diffusely present in contemporary cognitive sciences and also that the modern dichotomy between primary and secondary qualities, instantiated first by Galilei, continues to exercise a considerable influence even today. This represents the background within which what I call the Appearance-Reality Fallacy takes shape, which is the epistemological and ontological error to separate the world and also consciousness in two realms, the apparent/superficial one and the real/deep one, as if the laptop in front of me right now was just the appearance of a truer reality beyond it, like a thing-in-itself. This fallacy is present both in neuroscience and philosophy of mind and I will present a series of problems concerning neuro-centrism (myth of localization, mereological fallacy, fallacy of misplaced concreteness) as well as mento-centrism (Sartre's critique of mentalism, i.e. hypostatization of psychic objects), in order to avoid the immediate objectification of experience both in neuronal and mental sense. In this regard, on the basis of the analysis prepared in the first chapter, I will distinguish between actual and objectual consciousness, identifying the subject of the subsequent phenomenological analysis in the intrinsic and unavoidable perspectivalness of the actual consciousness, while instead most of contemporary studies on consciousness are focused on how consciousness is objectified by means of neuronal correlation or philosophical concepts (representations, qualia, self, etc.). This will lead me to propose ways of escape from reductionist naturalisms or irreconcilable dualisms, towards an alternative approach based on ontological pluralism and interdisciplinary epistemology.

4.1 The Cartesianism in contemporary Cognitive Sciences

We have already seen the conceptual foundations of the mainstream neuroscience of consciousness, which lie on the dualistic paradigm of modern science and philosophy. Even the drastic approaches like those I exposed before, despite promoting a monistic and empirical method and rejecting any dualistic alternative, seem to be still an attempt to deal with the heavy consequences of modernity. The three approaches – reductionism, eliminative materialism and scientific eliminativism – are rooted in the same naturalistic paradigm of the modern science, and, in this regard, Irvine's scientific eliminativism is the position that better respect the conditions of possibility of the objective science posed by Galilei and the mechanistic ontology posed by Descartes. Consciousness seems something hard to describe in a pure objective way and it seems something different from any other natural phenomenon – as we have seen in the first chapter – for these reasons Irvine does the same move of Galilei and Descartes, expelling the consciousness domain from the empirical analysis of the natural world, in favour of a better mathematization and idealization of this latter.

On the contrary reductionism and eliminative materialism try to analyse from the perspective of the objective science what seems, at least in part, intrinsically outside the very same domain of objective science, i.e. conscious experience. It is kind of like if the nozzle of the vacuum cleaner became the criterion of dustiness inside a room, in the corner of the room that the nozzle cannot reach there is no dust. Similarly, since the epistemological structure of contemporary sciences is taken also as the ontological criterion of natural phenomena, where scientific methodologies does not detect anything, there is nothing relevant at all, hence we can proceed reducing and eliminating (Bouwsma, 1948; cf. Nunziante, 2012, p. 133). This is what I would call an instance of methodological-ontological fallacy (Northoff, 2016d), that is, the determination of what exists (or does not exist) in the ontological sense through the determination of what we can know from a particular methodology. Furthermore, and more precisely related to neuroscience, the attempts to “put back the mind into nature” (Gerald M. Edelman, 1992) through the naturalization of consciousness, i.e. the reduction of its ontological dimension to the neural activations in the brain, can be

interpreted as examples of the empirical-ontological fallacy (Northoff, 2004a), that is, inferring existence and ontological characteristics of something from the empirical results about it¹ (e.g. the project of localization of consciousness inside the brain or the definition of personal identity as a brain's property; for a critical discussion see Wagner & Northoff, 2017). On the other hand, as we will see in more details later, someone might argue for the existence (or inexistence) of something inferred from what we can know from a particular perspective or epistemic stance, e.g. first- or third-person perspective; in this case, something may exist because it can be reliably included into the range of a subjective and qualitative point of view (FPP), like mental properties, while conversely something else may exist only if it can be reliably observed from an objective and external point of view (TPP), like the quantification of the mass of an object. Inferring from an epistemic stance an ontological commitment may imply an epistemic-ontological fallacy (Northoff, 2004a).

I will come back to these issues in detail later, now I would like to once again stress the point that all these theoretical models like the brain-body/machine, the brain-consciousness and the mind-body relationships come from the ontological and epistemological turn of modern science and philosophy. Interestingly enough, Michael Wheeler specifically exposed the “principles of Cartesianism” that still permeate the methodological and ontological structure of contemporary cognitive sciences (Wheeler, 2005):

- 1) «The subject–object dichotomy is a primary characteristic of the cognizer's ordinary epistemic situation. [p. 23]
- 2) Mind, cognition, and intelligence are to be explained in terms of representational states and the ways in which such states are manipulated and transformed. [p. 24]
- 3) The bulk of intelligent human action is the outcome of general purpose reasoning processes that work by (i) retrieving just those mental representations that are relevant to the present behavioral context, and then (ii) manipulating and transforming those representations in appropriate ways so as to determine what to do. [p. 38]
- 4) Human perception is essentially inferential in nature. [p. 42]

¹ This empirical–ontological fallacy «historically can be traced to Kant [Kant, 1787/1991] and his characterization of Locke as a “physiologist of reason”» (Northoff, 2018b, p. 249).

- 5) Perceptually guided intelligent action takes the form of a series of sense-represent-plan-move cycles. [p. 43]
- 6) In typical cases of perceptually guided intelligent action, the environment is no more than (i) a furnisher of problems for the agent to solve, (ii) a source of informational inputs to the mind (via sensing), and, most distinctively, (iii) a kind of stage on which sequences of preplanned actions (outputs of the faculty of reason) are simply executed. [p. 45]
- 7) Although the informational contents carried by bodily sensations and certain primitive perceptual states may have to be specified in terms that appeal to particular bodily states or mechanisms, the cognitive scientific understanding of the operating principles by which the agent's mind, given that information, then proceeds to generate reliable and flexible intelligent action remains conceptually and theoretically independent of the scientific understanding of the agent's physical embodiment. [p. 51]
- 8) Psychological explanation is temporally austere, in that it is neither necessary nor characteristically typical for good scientific explanations of mental phenomena to appeal to richly temporal processes [p. 53]».

After the analysis given in the third chapter, it should not be difficult to recognize how these principles are common points of both the Cartesian – more generally, modern – psychology and the contemporary attempts to offer a scientific explanation of consciousness. I shall focus mainly on the principles 1, 2, 5, 6, 7 and 8, as they can be easily observed within the neurocentric paradigm. Indeed, we can summarize the neurocentric argument about brain and consciousness in this way: the brain is the most relevant organ of human body for cognition and consciousness (7), and it can be defined as the organ of the mind, the seat of consciousness or the source of our subjectivity towards the outside world (1); it can be understood as a processing unit that constantly receives inputs from the environment and the body (e.g. photons on the retina) (6), encoding them in terms of information in and by the neural activity that transform these amounts of information in representational states of the outside world (2); once these information are integrated and selected, or globally broadcasted in the cortex (e.g. according to Integrated Information Theory or Global Neuronal Workspace Theory), they give rise to a conscious representation of the outside world. This neural representation may cause a behavioural reaction (e.g. the representation of something threatening for the subject) and a consequent action (e.g. moving away from

danger) (5); the project of a contemporary science of consciousness should deal with these neural computations, giving of an explanation of the causal structure of the entire information processing, as a series of discrete neural states put in a time sequence (8).

4.1.1 The persistent Primary-Secondary Qualities Dichotomy

Northoff interprets the contemporary neurocentric paradigm as an epistemological and ontological position where the vantage point is taken from within the brain, where there is no more need for something like a mind, since the neural activity is the real substrate that gives rise to mental features that are nothing but a type of process among those described above. And more generally, the neuro-centric view replace the mind with our brain as the ontic centre of the world, which is supposed to “move around” the brain (Northoff, 2018b). In this sense, it is not surprising that neuroscientists but also philosophers of mind and non-expert people speak of the brain as responsible of the production of the reality, by a world-simulation in the head², and even of some details that are not “really out there”, like colours, tastes, sounds, emotions, etc., i.e. secondary qualities³. Thus, it is possible to find articles and books that try to explain how the brain “creates”, “produces”, “makes”, “constructs” the mind and our neurorepresentations⁴ of the self and world (Dehaene, 2014; Eagleman, 2015; Feinberg, 2001; C. D. Frith, 2007; D. Hoffman, 2010; Hood, 2012; Metzinger, 2009), according to what Thomas Fuchs calls neuroconstructivist epistemology, i.e. phenomenal reality is to be understood as an internal mirroring or a reconstruction of the outer world by means of neuronal processes (Fuchs, 2017). Consequently, the world only seems to be colourful, textured, rich of sounds, scents and shapes, and for arguing this you do not have to be an illusionist necessarily, you can sustain a realist

² «The assumption that everything that people experience is, in reality, a construction or even an illusion created by their brains is one of the common convictions of neuroscientists and neurophilosophers. In particular, neuroimaging results, due to their seemingly simple and suggestive presentation, have ignited the enthusiasm of researchers, lay people, and the media alike. Assuming that we can literally watch the brain thinking, perceiving, or feeling, there is hardly a phenomenon, from pain or anger to colors or music and even to love or faith, which is not accommodated somewhere in the brain. The almost taken-for-granted view that reality can be found in the head turns perception, so to speak, into a physiological illusion» (Fuchs, 2017, p. 3)

³ [T]he world around you, with its rich colours, textures, sounds, and scents is an illusion, a show put on for you by your brain [...] If you could perceive reality as it really is, you would be shocked by its colorless, odourless, tasteless silence (Eagleman, 2015, p. 37).

⁴ Neurorepresentationalism is «a variant of representationalism that confines substrates for representation to the brain, and explicitly searches for neural substrates and mechanisms for representation within the brain» (Pennartz, 2015, p. 138, 2018).

position about the existence of qualities, but they still remain within the brain or the mind. In this sense, the brain does not process information that already are there in the world, but it seems that the brain itself creates relevant information for our conscious representation of the world, so that we do not actually perceive the world as such, rather we live within a product of the brain⁵. In the most “eliminative” case, our mental states and mental life in general are nothing but brain processes (Place, 1956; Smart, 1959).

As I already explained in the previous chapters, also some of those philosophers of mind that argue for the irreducibility of qualia to the neurophysiological analysis come from the same ontological roots of materialists. Indeed, when I speak of modern roots of cognitive sciences, I am not referring merely to the neuroscience and its neuro-centric paradigm, but also to some of the theories of consciousness from the Anglo-American philosophy of mind, that I would call “mentalistic” or “mento-centric”. Thus, the defence of some “extra ingredient” still needed for a complete explanation of conscious experience (Chalmers, 1995), of a sort of “insideness” of the mind that is different from the way that your brain is inside your head (T. Nagel, 1987), of the non-physicality of the redness of the tomato that Mary the neuroscientist cannot grasp with all her physical knowledge about vision (Jackson, 1986), all these cases come from the persisting modern subdivision between primary and secondary qualities, subject and object, mind and world, so that their solution – or dis-solution – depends on a complete revision of the underlying conceptual paradigm⁶;

⁵ Neuroscientists like Anil Seth and philosophers like Andy Clark, in attempting to explain how the brain works to non-expert people, go so far as to say that the conscious reality is like an hallucination. Since there is nothing in the outside world like the qualitative experiences we have every day, Seth considers the active processes the brain instantiates as a production of a sort of hallucinations, in the sense that the only difference from a “real” hallucination (when you perceive something that is not actually there) is that we collectively agree about our experiences in general (Rao, 2018; Seth, 2019). Similarly, Andy Clark describes perception as “controlled hallucination”, that is, the world we perceive is the production of brain’s expectations plus sensory information that work as a feedback on these expectations and anticipations of the neural activity – the main part of the perceiving activity (Clark, 2019). Although these are not argumentations coming from scientific articles, they clearly give an idea about the neurocentric paradigm and the persistent separation between primary and secondary qualities.

⁶ «I think it’s strictly true that we can’t, as things stand now, so much as imagine the solution of the hard problem. The revisions of our concepts and theories that imagining a solution will eventually require are likely to be very deep and very unsettling. (That’s assuming what’s by no means obvious: that we are smart enough to solve it at all.) Philosophers used to think (some still do) that a bit of analytical tidying up would make the hard problem go away. But they were wrong to think that. There is hardly anything that we may not have to cut loose from before the hard problem is through with us» (Fodor, 2007, pp. 9–10).

these latter philosophical positions try to argue for the autonomy and reality of the secondary ones, while eliminative materialist argues for their ontological removal; still the roots are the same. The extreme of this paradigm is panpsychism, according to which consciousness is not confined to biological entities, e.g. as an epiphenomenon or a emergent property, but is a fundamental feature of all physical matter; in this sense, this can be seen as an overthrow of the Galilean primary-secondary distinction (Goff, 2019).

Taken together, regardless of which ontological commitment (be it a monistic materialism or a substantial dualism) or methodological stance (be it empirical or conceptual) we take, it seems that the main contemporary lines of research on consciousness are still based on the modern scientific revolution. Despite modernity has been an essential turning point for the scientific progress, I believe that we need to find a new way for understanding experience in such a way as to keep together all the dualisms that have fractured the phenomenon of consciousness (subject vs object, mind vs world, mind vs body, qualitative vs quantitative, inner vs outer, etc.). If we remain stuck within what I shall call the appearance-reality divide, I think the only possible alternatives are a monist position that reduces, eliminates, degrades consciousness to illusion or a dualistic position that risks separating the metaphysical concept of consciousness from the empirical findings on it. One of the aims of this chapter and the following will be to conceive, epistemologically, an interdisciplinary but integrative approach on consciousness, and, ontologically, a plurality within the unity of the phenomenon of consciousness. In order to do this, I shall focus on the divide between appearance and reality understood as the basis of the contemporary projects of a “science of consciousness” (from both the scientific and philosophical perspective).

4.2 Consciousness beyond Surface and Deep Properties

As already discussed in the previous chapters, the debate around consciousness is often characterized by two main positions, the realist one and the eliminative/illusionist one. For the former consciousness is an autonomous and distinct

phenomenon, not reducible to any empirical finding regarding the brain activity, while for the latter consciousness is an old, vague concept that needs to be replaced by the empirically informed and neuronal-based – or more generally physical-based – knowledge. In between we have several moderate positions that argue for a distinction between apparent and real properties of consciousness, in the same way as we would define any other natural phenomenon. For instance, getting back to the direct reference theory, we can refer to something – e.g. “heat” – to its superficial – e.g. the sensation of heat – or deep – e.g. the molecular motion – properties⁷. In line with a kind of scientific essentialism, on the basis of this divide there is the belief that the regional ontology offered by the scientific paradigm is more “real” or “fundamental” than others, like the regional ontology of phenomenal qualities. This scientific essentialism, despite not being explicitly endorsed, is widespread among most of the scientific theories of consciousness, where indeed the main goal is to lead the mental features – superficial properties – to the neuronal activity – deep properties, i.e. how the brain “causes” the mind. Now, I shall challenge this position starting from this distinction between superficial-deep properties, or more generally, between appearance and reality.

In *The Rediscovery of the Mind*, Searle claims that there is an ontological – and not just epistemological – difference between, for instance, the conscious state of pain and the perceptible property of redness. The red colour can be defined in an ostensive way with a series of examples, but if we causally reduce the phenomenon to its physical composition – e.g. referring to photon emissions of 600 nanometres – then we can «carve off and eliminate the subjective experience of color from the “real” color. Real color has undergone a property ontological reduction to light reflectances» (Searle, 1992, p. 115). On the contrary, we have already seen with Kripke that this kind of reduction is not possible with some conscious state – e.g. pain, pleasure, etc. Moreover, Searle argues that this argument is consistent with the phenomenon of consciousness as such:

⁷ In the second chapter I argued that consciousness seems to resist this kind of distinction, since its alleged superficial properties (like the feeling of pleasure or pain) are far from being easily separated from its intrinsic structure.

«But now we come to an apparently shocking asymmetry. When we come to consciousness, we cannot perform the ontological reduction. Consciousness is a causally emergent property of the behavior of neurons, and so consciousness is causally reducible to the brain processes. But – and this is what seems so shocking – a perfect science of the brain would still not lead to an ontological reduction of consciousness in the way that our present science can reduce heat, solidity, color, or sound» (Searle, 1992, p. 116).

Such kind of thesis can be found also one of the most important publications of Gerald Edelman and Giulio Tononi, *A Universe of Consciousness* (Gerald M Edelman & Tononi, 2000, pp. 12–13):

«Scientific explanations can provide the conditions that are necessary and sufficient for a phenomenon to take place, can explain the phenomenon's properties, and can even explain why the phenomenon takes place only under those conditions. But no scientific description or explanation can substitute for the real thing. We all accept this fact when we consider, say, the scientific description of a hurricane: what kind of physical process it is, why it has the properties it has, and under what conditions it may form. But nobody expects that a scientific description of a hurricane will be or cause a hurricane. Why, then, should we not apply exactly the same standards to consciousness? We should provide an adequate description of what kind of physical process it is, why it has the properties it has, and under what conditions it may occur. As we shall see, there is nothing about consciousness that precludes an adequate scientific description of the particular kind of neural process it corresponds to. What, then, is special about consciousness? What is special about consciousness is its relationship to the scientific observer. Unlike any other object of scientific description, the neural process we are attempting to characterize when we study the neural basis of consciousness actually refers to ourselves—it is ourselves—conscious observers. We cannot therefore tacitly remove ourselves as conscious observers as we do when we investigate other scientific domains»⁸.

⁸ See also (Gerald M Edelman, 2003, p. 5521): «To expect that a theoretical explanation of consciousness can itself provide an observer with the experience of “the redness of red” is to ignore just those phenotypic properties and life history that enable an individual animal to know what it is like to be such an animal. A scientific theory cannot presume to replicate the experience that it describes or explains; a theory to account for a hurricane is not a hurricane. A third-person description by a theorist of the qualia associated with wine tasting can, for example, take detailed account of the reported

The theoretical strategy is approximately the same in both Searle and Edelman and Tononi: at the heart of the argument there is a malleable distinction between two types of ontological domains. On one hand, there is the domain of the “surface features” and the “appearance” (Searle, 1992, p. 119) what we may call the observer-related thing⁹. On the other hand, the domain of the underlying reality, the “real thing” as such, the natural world described by physics¹⁰. The first, the appearance domain, may include secondary and some of the primary qualities – e.g. the subjective perception of red and the property of solidity¹¹ – and it may be progressively excluded from the scientific explanation, and substituted by the “real domain”, in which their underlying ontological causes are determined. The problem with consciousness and specific conscious states is that, regarding that kind of explanation, our interest is necessarily focused on the subjective experience as such. In other words, the distinction between appearance and reality does not work with consciousness as the object of the investigation: «Where appearance is concerned we cannot make the appearance-reality distinction because the appearance is the reality» (Searle, 1992, p. 112) and in such cases «no scientific description or explanation can substitute for the real thing» (Gerald M Edelman & Tononi, 2000, p. 12)¹².

personal experiences of that theorist and his human subjects. It cannot, however, directly convey or induce qualia by description; to experience the discriminations of an individual, it is necessary to be that individual».

⁹ «Unlike any other object of scientific description, the neural process we are attempting to characterize when we study the neural basis of consciousness actually refers to ourselves—it is ourselves—conscious observers» (Gerald M Edelman & Tononi, 2000, p. 13).

¹⁰ «If I, as a “naive human being” who is “deceived by the senses” have yielded to my inclinations to develop such reflections, I now recall, as a “scientific” human being the well-known distinction between secondary and primary qualities according to which the specific qualities pertaining to the senses are “merely subjective,” and only the qualities dealt with in geometry and physics are “Objective.” The color, the sound, the odor, and the taste of the physical thing, however much they appear “in person” in the thing, as qualities included in its essence, are not themselves actually as what they appear: to be there, but instead are “signs” of certain primary qualities. [...] Rather the entire essential contents of the perceived physical thing, thus the whole physical thing standing there “in person” and all its qualities, including all those which could ever be perceived, is a “mere appearance” and that the “true physical thing” is the one “determined” by physics» (Husserl, 1913/1982, §40).

¹¹ As well as secondary qualities, «[t]he same pattern works for the primary qualities: Solidity is defined in terms of the vibratory movements of molecules in lattice structures, and objective, observer-independent features, such as impenetrability by other objects, are now seen as surface effects of the underlying reality. Such redefinitions are achieved by way of carving off all of the surface features of the phenomenon, whether subjective or objective, and treating them as effects of the real thing» (Searle, 1992, pp. 119–120).

¹² «[...] [W]ith respect to consciousness, there is no difference between appearance and reality. With consciousness, what there seems to be is what there is. What there seems to be is all there is» (Honderich, 2006, p. 5)

4.2.1 The intrinsic Perspectivalness of Consciousness

Now, this is in line with what we discussed in the first and second chapter, that is, the fact that it is difficult to rigidly separate appearance and reality domains in the case of consciousness, since some of the features we generally would ascribe to the former dimension actually seem to be intrinsic characteristics, which are necessary for the definition of consciousness. This has been argued as a reply to the illusionist position (while the objectual consciousness may be illusory, the actual consciousness is undeniable; Rowlands, 2001) and to the brain-consciousness identity theories (there is no easy separability between deep and superficial properties in consciousness). However, this argument against reductivism, thus developed, may also lead to a misinterpretation of the issue here exposed as the appearance-reality relation. The reductive scientific view as well as the main conceptual paradigm of contemporary philosophy of mind are all rooted on this problematic ground. As both Searle and Edelman sustain, scientific research pursues the progressive determination of the real thing by excluding from the explanation all subjective contents – e.g. the “real” red as photon emissions as well as the real description of solidity as a particular disposition of molecules, and so forth. On the basis of this distinction, it follows that consciousness is a very strange phenomenon, because it seems impossible to remove completely neither the subjective point of view (epistemological version of the problem) nor the subjective content of the experience (ontological version of the problem)¹³. I shall argue that the basic problem here is not that consciousness seems to escape the reduction to pure physical descriptions, as Searle argues, since this is the symptom – and not the cause – of the fundamental appearance-reality issue.

I already described some cases in which it seems that the apparent-reality distinction does not work well, e.g. pain-consciousness – consciousness (of) pain *à la*

¹³ About that, Husserl explained that a purely empirical-objective psychologism is a non-sense, because it seems that the objective science, once chased the subject out through the door, yet it now seeks to bring it back in through a window (Husserl, 1936, §64). This entails a vicious circle, that is the attempt to explain the subjective phenomenon with natural laws that depends on the presupposed world idealization made by the subjective stance itself. In other words, according to Husserl, a truly psychological science conceived as an objective science of subjective structures has become impossible with this epistemological position (§78).

Sartre – and others (pleasures, beliefs, sensations, etc.), however someone might argue that these are extremely rare instances of intrinsic relation between the conscious act and its object. Pain, pleasure, etc. were useful in order to highlight the constitutive presence of the conscious perspective even for their existence, and not only for a complete description of both phenomena, but there seems to be lots of cases in which the phenomenal aspects of consciousness are not necessary for the existence and description of object in question. It is not my claim to support any kind of phenomenalism in which all the objects of the world depends upon the existence of consciousness, however, I shall try to apply the same “perspectival-dependence” on other phenomena that apparently do not require any non-physical, subjective or observer-related content in general for a complete description. Let us take the example illustrated by Edelman and Tononi, the hurricane. They say – correctly – that a scientific explanation can provide the necessary and sufficient conditions for the hurricane to “emerge”¹⁴, but it cannot substitute the real thing nor cause it¹⁵. Ontologically speaking, if we put in a supercomputer all the information, all the contextual contingences and all the variables carved from the analysis of a hurricane, no one would expect to witness a real hurricane arising; epistemologically speaking, not a single scientist would require from his colleagues the description of “what it is like” to be inside a hurricane for giving an exhaustive scientific description of the event – lucky for him. In short, we do not need any subjective features in order to define the nature of the real thing, rather we need to remove ourselves as conscious observer for a scientific explanation as such (Gerald M Edelman & Tononi, 2000). And this is when the problems arise.

Indeed, it is not true that in these cases we can remove the observer-relation from the “normal” objects of scientific research. We cannot remove our perspectivalness, not only regarding consciousness, but also with any other analysis of natural phenomena. We can remove possible qualities from the first-person perspective (e.g. the fact that H₂O is tasteless), but we cannot have a “pure

¹⁴ The concept of emergence is not used innocently, because it is one of the epistemological cornerstones on which most of contemporary scientific theory of consciousness lies. See (Searle, 1992, pp. 111–116).

¹⁵ For example, see Patricia Churchland on the thought experiment of Mary the neuroscientist: «Just as the obstetrician does not become pregnant by knowing all about pregnancy, so Mary does not have the sensation of redness by knowing all about the neurophysiology of perceiving and experiencing red» (P. S. Churchland, 1986, p. 332).

knowledge”¹⁶ as a point of “view from nowhere” (T. Nagel, 1986)¹⁷ nor from everywhere, like a “God’s-eye view” (Putnam, 1981)¹⁸ – where the two alleged non-perspectival positions, despite being literally opposites, lead to the same result (Northoff, 2018b). We just pretend to analyse any process, event or fact with “centreless” point of view, understood as an observer-independent perspective. The third-person perspective (TPP) is often considered as the most “real” view, through which we can consequently approach the real thing – the external world – beyond all the perspectival and subjective appearances. Despite this assumption, we often confuse the fact that things may present qualitative aspects that depends on subjectivity with the most basic fact that we never get rid of our perspectivalness¹⁹. On the contrary, this is the ground zero of every relationship between us and the world, i.e. being in a situation already given to me from the start, which I can approach only partially from time to time. Thus, the third-person perspective is not the detector of an alleged underlying reality made of physical facts, but rather it is a kind of abstract agreement between perspectives that it may become independent of the particular and contingent observer – e.g. the single researcher – but not of the observation as such (Hales, 2014). This is not a critique to the scientific method; it is a claim against the idea of a division between a physical-objective-real domain and a non-physical-subjective-apparent

¹⁶ «The point of view of pure knowledge is contradictory; there is only the point of view of engaged knowledge. This amounts to saying that knowledge and action are only two abstract aspects of an original, concrete relation. The real space of the world is the space which Lewin calls “hodological.” A pure knowledge in fact would be a knowledge without a point of view; therefore a knowledge of the world but on principle located outside the world. But this makes no sense; the knowing being would be only knowledge since he would be defined by his object and since his object would disappear in the total indistinction of reciprocal relations. Thus knowledge can be only an engaged upsurge in a determined point of view which one is. For human reality, to be is to-be-there; that is, “there in that chair,” “there at that table,” “there at the top of that mountain, with these dimensions, this orientation, etc.” It is an ontological necessity. (Sartre, 1956, p. 308)

¹⁷ «It includes all the individuals in the world, of every kind, and all their mental and physical properties. In fact it is the world, conceived from nowhere within it. But if it is supposed to be this world, there seems to be something about it that cannot be included in such a perspectiveless conception-the fact that one of those persons, TN, is the locus of my consciousness, the point of view from which I observe and act on the world» (T. Nagel, 1986, p. 56).

¹⁸ «On this perspective [the perspective of metaphysical realism], the world consists of some fixed totality of mind-independent objects. There is exactly one true and complete description of ‘the way the world is’. Truth involves some sort of correspondence relation between words or thought-signs and external things and sets of things. I shall call this perspective the externalist perspective, because its favorite point of view is a God’s Eye point of view» (Putnam, 1981, p. 49).

¹⁹ «There is no pure third-person perspective, just as there is no view from nowhere. To believe in the existence of such a pure third-person perspective is to succumb to an *objectivist illusion*. This is, of course, not to say that there is no third-person perspective, but merely that such a perspective is exactly a perspective from somewhere. It is a view that we can adopt of the world. It is a perspective founded upon a first-person perspective, or to be more precise, it emerges out of the encounter between at least two first-person perspectives; that is, it involves intersubjectivity» (Gallagher & Zahavi, 2012, p. 40).

domain. Surely, we can take a detached view upon the world, as if consciousness does not exist. Figuratively speaking, we can easily evade our subjective limits of inhabitants of planet Earth and imagine an objective point of view from which we can place the planet Earth revolving within the Solar System and the latter within the Milky Way galaxy and so on, then in order to measure all the phases of the stars and planets (we will see later how important this capacity is for an ecological perspective on consciousness). However, it would be a mistake to think that this scientific attitude can replace and erase the natural perspective (Husserl, 1991). Once out of scientific abstraction, we still come back to see the sun “rising” and “going down” in relation to Earth as our horizon/ground-zero (in a Ptolemaic-like way).

4.3 The Appearance-Reality Fallacy

In this sense, consciousness is not merely a contingent point of view – e.g. the redness of the tomato as in the thought experiment of Mary the neuroscientist – nor a non-physical object with some undetectable features – e.g. as for mysterianism or substance dualism – nor only a peculiar way of knowing the world – e.g. as if consciousness were identifiable with FPP. All these accounts entail a kind of opposition (interior-exterior; real-apparent; knowing-known, subject-object, etc.) that derives from the modern scientific revolution, as seen, particularly from the scientific method and the mechanistic view of the world. As Sartre said, and before him Nietzsche and Husserl, trying to find something more true or more real behind the so-called “appearance” is the result of the “illusion of worlds-behind-the-scene”, and this is what Searle, Edelman, Tononi, Jackson and so many other figures of contemporary neuroscience and philosophy of mind do.

The appearances which manifest the existent are neither interior nor exterior; they are all equal, they all refer to other appearances, and none of them is privileged. Force, for example, is not a metaphysical conatus of an unknown kind which hides behind its effects (accelerations, deviations, etc.); it is the totality of these effects. [...] No one of these actions alone is sufficient to reveal it. But no action indicates anything which is behind itself; it indicates only itself and the total series. The obvious conclusion is that the dualism of being and appearance is no

longer entitled to any legal status within philosophy. The appearance refers to the total series of appearances and not to a hidden reality which would drain to itself all the being of the existent. And the appearance for its part is not an inconsistent manifestation of this being. To the extent that men had believed in noumenal realities, they have presented appearance as a pure negative. It was “that which is not being”; it had no other being than that of illusion and error. But even this being was borrowed, it was itself a pretence, and philosophers met with the greatest difficulty in maintaining cohesion and existence in the appearance so that it should not itself be reabsorbed in the depth of non-phenomenal being. But if we once get away from what Nietzsche called “the illusion of worlds-behind-the-scene,” and if we no longer believe in the being-behind-the-appearance, then the appearance becomes full positivity; its essence is an “appearing” which is no longer opposed to being but on the contrary is the measure of it. For the being of an existent is exactly what it appears (Sartre, 1956, pp. xlv–xlvi).

What does this long Sartrean quote have to do with what has been discussed so far with respect to the objectivity of science and the appearance of subjective qualities – in particular with consciousness? Someone might interpret these lines as the anti-realistic disruption of the appearance-reality distinction through the defence of a solipsistic idealism, where everything is appearance and there is nothing “real” behind it, leading inevitably to the loss of consistency of every scientific inference about the world. However, this is not the case, since Sartre explicitly rejects any form of idealisms, phenomenalism or mind-reductive approaches, indeed accepting a constitutive perspectivalness within every inference, act or knowledge on the world does not necessarily imply any ontological primacy of the mind²⁰. But, what does this mean for our present purpose regarding the possibility of a coherent and constitutive account of consciousness? How Sartre’s phenomenological ontology can be useful for

²⁰ «This “transphenomenality of Being” means that the object of consciousness is always outside and transcendent, that there is forever a resistance, a limit offered to consciousness, an external something which must be taken into consideration. Nevertheless we have not substituted a realistic position for the idealistic. For without consciousness, Being does not exist either as a totality (in the sense of “the world,” “the universe”) or with differentiated parts. It is a fullness of existence, a plenitude which can not possibly isolate one part so as to contrast it with another, or posit a whole over against its parts, or conceive a “nothing” in opposition to which it is “everything”. It is simply undifferentiated, meaningless massivity. Without consciousness there would not be a world, mountains, rivers, tables, chairs, etc.; there would be *only* Being. In this sense there is no *thing* without consciousness, but there is not *nothing*. Consciousness causes there to be things because it is itself nothing. Only through consciousness is there differentiation, meaning, and plurality for Being» (Sartre, 1956, p. xx). Against the solipsistic scepticism, see also (Husserl, 2006, §18).

understanding whether the concept of consciousness can be analysed or not under the appearance-reality paradigm, i.e. whether is it possible a purely objective account of consciousness or not? Perhaps, an example made by Sartre himself can be useful for our purpose.

In *Being and Nothingness*, during the exposition of the theme of the body, Sartre discusses the problem of subjectivity and objectivity. He takes the example of volunteering for some psychologic experiment in which you are asked to sit in front of a screen, follow the instructions and then answers the experimenter, a typical attempt of scientific research to isolate the individual from the world in order to focus on specific elements of perception, consciousness, cognition, etc. and analyse them in the more objective way.

«From time to time the experimenter asked me if the screen appeared to me more or less illuminated, if the pressure exerted on my hand seemed to me stronger or weaker, and I replied; that is, I gave objective information concerning things which appeared in the midst of my world. Sometimes an inept experimenter asked me if "my sensation of light was stronger or weaker, more or less intense." Since I was in the midst of objects and in the process of observing these objects, his phrase would have had no meaning for me if I had not long since learned to use the expression "sensation of light" for objective light as it appeared to me in the world at a given instant. I replied therefore that the sensation of light was, for example, less intense, but I meant by this that the screen was in my opinion less illuminated» (Sartre, 1956, p. 311).

Even though Sartre is talking about an example of experiment made seventy-eighty years ago, the epistemological structure is the same of the actual scientific paradigm. With the "inept experimenter" that asks for my sensation of light, Sartre is trying to highlight the typical attitude to separate the subjective sensations we have about the world (the sensation of light) and the stricter objective world as such (the light as such and the experimental measurements):

«Since I actually apprehended the screen as less illuminated, the phrase "in my opinion" corresponded to nothing real except to an attempt not to confuse the objectivity of the world-for-me with a stricter objectivity, which is the result of

experimental measures and of the agreement of minds with each other» (Sartre, 1956, p. 311).

This is exactly what I mean with the appearance-reality distinction/fallacy. The idea that there are two series of objects in the world, the apparent ones (subjective, erroneous, malleable, unreliable) and the real ones (objective, true, unchangeable). We can clearly recognise this kind of distinction during the experiments about vision, as I already expose in the first chapter. The question always is whether we actually see less than we think to see or whether we produce lots of subjective artefacts when we look at the world, hinting that what we perceive is not the real world.

«Why indeed should we use the term “subjectivity” for the ensemble of luminous or heavy or odorous objects such as they appeared to me in this laboratory at Paris on a day in February, etc. And if despite all we are to consider this ensemble as subjective, then why should we recognize objectivity in the system of objects which were revealed simultaneously to the experimenter, in this laboratory, this same day in February? We do not have two weights or two measures here; we do not encounter anywhere anything which is given as purely felt, as experienced for me without objectivation» (Sartre, 1956, p. 312).

I am not saying that our conscious experience is always right and must be taken as completely reliable. This is evident, for example, in the studies of how vision works, indeed the experiment often shows our limits in identifying colours, movements, things in the scene, despite in our daily life we think to see a colourful and detailed world at any moment. The issue here is another, that is, there is no point in degrading the “subjective” qualities of the world in something fictional, unreal and unimportant. For instance, it is generally said that hallucinations correlated to psychotic disorders, e.g. schizophrenia, are mere illusions and it is nothing real, because it does not match with the physical reality, so they probably are only images “in the head” of the patient. Treating these cases in such a way it would be like trying to convince a schizophrenic patient that the voices he hears are in his head, by recording the sounds in the room with a microphone and then making him hear the recording in order to show that there are no voices at all. Besides being completely inappropriate for the patient, such a “physicalist” approach cannot even help to interpret the nature of the hallucinations, because it discredits and degrades to “unreal” or “untrue” from the beginning anything

that does not precisely match with the mathematico-physical ideal of the world. Instead, every mode of being is equally “real”, without either being reducible to the other, or explicable in terms of some absolute mode of being.

Now, I am not equating the perspective of a single subject with the scientific approach of quantitative measurement, since this latter can be surely chosen for a reliable description of some portion of the world, but there is no reason for interpreting it as the detector of a superior reality “behind the scene”:

«[...] for example, I can establish that the warm water appears cold to me when I put my hand in it after having first plunged my hand in hot water. But this establishment which we pompously call “the law of relativity of sensations” has nothing to do with sensations. Actually we are dealing with a quality of the object which is revealed to me: the warm water is cold when I submerge my heated hand in it. A comparison of this objective quality of the water to equally objective information which the thermometer gives me simply reveals to me a contradiction. This contradiction motivates on my part a free choice of true objectivity. I shall give the name subjectivity to the objectivity which I have not chosen. [...] Scientific objectivity consists in considering the structures separately by isolating them from the whole; hence they appear with other characteristics. But in no case do we get out of an existing world» (Sartre, 1956, p. 312)²¹.

While the choice between different series of objective facts (i.e. relative to objects) in some cases is actually simple – e.g. in an empirical study of the characteristics of non-Newtonian fluids I would definitely give priority to the information offered by fluids mechanics – in other the different sets of information (physiological, behavioural, psychological, introspective, etc.) must be taken all together in consideration. *A fortiori* during the examination of a phenomenon on the line between qualitative and quantitative, subjective and objective characteristics, as consciousness.

²¹ «In this passage, Sartre appreciates that experience does not take the form of an internal ‘technicolour phenomenology’ but a rich, intricately structured experiential world. Consciousness is finding oneself amongst a collection of objects in a laboratory at a particular time, as opposed to being a collection of mysterious internal sensations that occur when one is in the laboratory» (Ratcliffe, 2007, p. 486).

4.3.1 The Insufficiency of the Physicalist account of Reality

Regarding consciousness, indeed, there are different series of objects to consider, as we have already seen in the first and second chapter. An exclusive focus on the brain activity as the cause of all the other series of information (psychological, behavioural, cognitive, existential, and so forth) would inevitably entail the ontological reduction of consciousness to neurophysiology, i.e. brainhood and the neurocentrism, and more generally the epistemological reduction to the naturalistic paradigm. As well as the example of the schizophrenic patient, arguing that there is nothing like a real phenomenon called consciousness (or self, etc.), since nothing in the brain suggests its existence (P. S. Churchland, 2002, 2013; Metzinger, 2003, 2009), completely misses the point: of course there is nothing like a self or consciousness in the brain, because you are vacuuming with the wrong nozzle (see above Bouwsma's metaphor). And I believe that these forms of scientism, understood as the idea of a universal applicability of the scientific method with the exclusion of other perspectives, are incompatible with a complete study of consciousness.

At the same time, as I will describe in details later, we must be careful not to fall on the opposite side of the fallacy appearance-reality fallacy, since the identification of certain epistemological and ontological limits by the physicalist paradigm does not automatically imply the existence of other spooky entities in the world. In this sense, some of the attempts to claim the autonomy of consciousness from the natural world, as a distinct element incompatible to any scientific description would lead to the exclusion of any series of information but the mentalistic one, i.e. mento-centrism, substantial dualisms and idealism, rejecting any possible relation between different domains (e.g. brain and mental features). Thus, if we want to go beyond the problematic project of a pure "objective", neuro-maniac science of consciousness, we need to consider first the possibility of alternative ontologies regarding the world and consciousness, i.e. a different way to conceive the relation between different sets of information we can collect from different perspective. Instead, the reduction of these relations into one single ontological sets of entities – be they neuronal, mental, behavioural, and so forth – will always be based on the appearance-reality fallacy.

An alternative way of considering the relationship between various groups of elements could be to conceive them not in a hierarchical way in which the aim is to find what causes what (mere efficient cause, as required by the physicalist model), but as different layers of the same phenomenon, i.e. different appearances that cover only some portion of it and where none of them can claim to be the fundamental one from which all the other descend. These appearances, as Sartre would say, do not refer to any non-apparent reality behind it, it has not the «double relativity of Kant's *Erscheinung*. It does not point over its shoulder to a true being which would be, for it, absolute. What it is, it is absolutely, for it reveals itself as it is. The phenomenon can be studied and described as such, for it is absolutely indicative of itself» (Sartre, 1956, p. xlvii). Every single appearance (be they a qualitative affection or a quantitative measurement) reveals the entity in question and refers to the totality of all its possible appearances, so that the revealing act cannot complete the determinations of that entity once for all, i.e. the transphenomenality of being²². In other words, every manifestation towards consciousness or, if we want, every disclosing activity of consciousness (Rowlands, 2018) can be understood as an act of “decompression of being” (Sartre, 1956). Furthermore, there is no hierarchical difference in the ontological determination of the thing, and any domain reveals a part of it, from different perspectives and

²² «That is why we can equally well reject the dualism of appearance and essence. The appearance does not hide the essence, it reveals it; it is the essence. The essence of an existent is no longer a property sunk in the cavity of this existent; it is the manifest law which presides over the succession of its appearances, it is the principle of the series. [...] But the appearance, reduced to itself and without reference to the series of which it is a part, could be only an intuitive and subjective plenitude, the manner in which the subject is affected. If the phenomenon is to reveal itself as transcendent, it is necessary that the subject himself transcend the appearance toward the total series of which it is a member. [...] In a particular object one can always distinguish qualities like color, odor, etc. And proceeding from these, one can always determine an essence which they imply, as a sign implies its meaning. The totality "object-essence" makes an organized whole. The essence is not in the object; it is the meaning of the object, the principle of the series of appearances which disclose it. But being is neither one of the object's qualities, capable of being apprehended among others, nor a meaning of the object. The object does not refer to being as to a signification; [...] The object does not possess being, and its existence is not a participation in being, nor any other kind of relation. It is. That is the only way to define its manner of being; [...] By not considering being as the condition of revelation but rather being as an appearance which can be determined in concepts, we have understood first of all that knowledge can not by itself give an account of being; that is, the being of the phenomenon can not be reduced to the phenomenon of being. [...] The phenomenon of being requires the transphenomenality of being. That does not mean that being is found hidden behind phenomena (we have seen that the phenomenon can not hide being), nor that the phenomenon is an appearance which refers to a distinct being (the phenomenon exists only qua appearance; that is, it indicates itself on the foundation of being). What is implied by the preceding considerations is that the being of the phenomenon although coextensive with the phenomenon, can not be subject to the phenomenal condition-which is to exist only in so far as it reveals itself-and that consequently it surpasses the knowledge which we have of it and provides the basis for such knowledge» (Sartre, 1956, pp. xlvii, xlviii, xlix, l).

approaches. The fact that the set of physical information about the world seems fundamental and related to the “real thing”, compared to other ontological domains that seems to regard only to the appearance, is an assumption constructed in the course of the last centuries of great progress of the scientific research²³. On the other hand, arguing – as I shall do – that the physicalist and mathematical approach towards the objects is a particular way – among others – of revealing things through a presupposed abstraction and idealization of world, this argument in question is not, in any way, an attempt to discredit the potentiality, usefulness and peculiarity of the scientific method, nor it is my intention to claim that science offers only artefactual notions of the world²⁴. On the contrary, it is my aim to open the empirically driven domain of science to other sets of ontological facts upon the things of the world, and with the possibility of integration and complementarity of different approaches, as I shall demonstrate in the next chapter.

4.3.2 Lifeworld as the Horizon of all the Worlds

So far, I have exposed the appearance-reality dichotomy and its relationship with the mainstream physicalist approach of science. Now, I shall focus on how the divide is related with consciousness. How can this ontological analysis relate to the issue of consciousness? We do not need to posit a realm of various appearances distinct from the physical, nor to deny the reality of phenomenal experience, rather we may try to establish a connection between different modes of being. Moreover, if we

²³ Someone might argue that the distinction between primary and secondary qualities it is far older than Galilei, Locke and the others, since Democritus, the forefather of materialism, seems to have said, as reported by Galen: “There seems to be colour, there seems to be sweetness, there seems to be bitterness. But really there are only atoms and the void”. However, as Galen Strawson noted, « the quoted passage is only half of an imaginary exchange Democritus stages between “The Intellect” and “The Senses”. The Intellect speaks first: “There seems to be colour, there seems to be sweetness, there seems to be bitterness. But really there are only atoms and the void.” But then The Senses reply: “Poor Intellect, do you hope to defeat us while from us you borrow your evidence? Your victory is your defeat”» (Strawson, 2017, p. 95-96; for a critical analysis of the idea of Democritus as the first proponent of the primary-secondary distinction, see also Pasnau, 2007).

²⁴ «The physical thing as determined by physics, however, is nothing foreign to what appears sensuously "in person;" rather it is something which makes itself known originaliter in it and, more particularly, a priori (for indefeasible eidetic reasons) only in it. Accordingly, even the sensuous determination-content of the X which functions as bearer of the determinations ascribed in physics is no clothing foreign to these determinations and hiding them: rather, only because the X is the subject of the sensuous determinations is it the subject also of the determinations ascribed in physics which, for their part, make themselves known in the sensuous determinations» (Husserl, 1913/1982, §52).

acknowledge the impossibility to remove consciousness from experiencing the world, this must not be understood as the defence of a solipsistic position, but as the manifestation of a series of phenomena, which reveal themselves as they appear, with no hidden reality behind but through various modes of appearance (physical, perceptual, psychological, cognitive, physiological, etc.). As we will see later, this series of phenomena given to us already “on hand” (Husserl, 1913/1982) in different ways, meanings and possibilities is the background condition of being-in-the-world (Heidegger, 1927/1996; Husserl, 1913/1982). Thus, by taking this phenomenological approach, the third-person perspective – in particular the physicalist ontology related to it – becomes (only) one of the possible views that describe the world through abstractions of the “lifeworld”²⁵. As Husserl said,

«When physics determines the physical thing given exclusively by such concepts as atoms, ions, energies, and so forth, and as, in any case, space-filling processes for which the only characterizations are mathematical expressions, it means them as something transcendent to the whole physical-thing content standing there "in person." As a consequence, it cannot mean the physical thing as something located in the natural space pertaining to the senses. In other words, the space of physics cannot be the space belonging to the world given “in person” in perception» (Husserl, 1913/1982, p. §40)²⁶.

Thus, first and foremost every act of consciousness (“actual consciousness”; Rowlands, 2001) is based on and immersed into what Husserl called the “natural attitude”, the “natural concept of the world” (Husserl, 1913/1982, 2006) or – even better – “lifeworld” (*Lebenswelt*; Husserl, 1936/1970), in which we all live surrounded by a pre-given spatiotemporal structure, and from which we start when we bring about the philosophical and scientific transformation of our viewpoint²⁷. This primordial

²⁵ «[...] [W]e must note something of the highest importance that occurred even as early as Galileo: the surreptitious substitution of the mathematically substructured world of idealities for the only real world, the one that is actually given through perception, that is ever experienced and experienceable – our everyday life-world. This substitution was promptly passed on to his successors, the physicists of all the succeeding centuries» (Husserl 1970, 48-49).

²⁶ See also: «It is this world [the pregiven world] that we find to be the world of all known and unknown realities. To it, the world of actually experiencing intuition, belongs the shape of space-time together with all bodily shapes incorporated in it; it is in the world that we ourselves live, in accord with our body (*leiblich*), personal way of being. But here we find nothing of geometrical idealities, no geometrical space or mathematical time with all their shapes» (Husserl, 1970, §9).

²⁷ «[...] the life-world does have, in all its relative features, a general structure. This general structure, to which everything that exists relatively is bound, is not itself relative. We can attend to it in its

horizon of world experience is at the same time a foundational and a founded ontology, a burden but also a springboard for consciousness (Merleau-Ponty, 1995), subjective-relative and, conversely, given prior to all ends (Husserl, 1970b; Zelić, 2008). On the one hand, this pre-given reality is what we need to disrupt and investigate in order to reveal its unthought presuppositions, through phenomenological analysis, but also with philosophical conceptualization and scientific research. On the other hand, this “actual” lifeworld “in flesh and blood” (*Leibhaft*) is the pre-reflective and intuitive condition of possibility of every further theoretical idealization or metaphysical construction of the world – therefore it is not the “real” world behind the apparent worlds²⁸. In other words, it is the «horizon within which all of the other worlds, the worlds of scientific objectivity, of perceptual objectivity, and of pre-scientific interests, are constituted and given. However, the lifeworld is not one world beside all of the others. Rather, the lifeworld is, to put it in a formula of transcendental philosophy, the condition of the possibility of all worlds qua worlds of natural conception» (Zelić, 2008, p. 423)²⁹.

Within the horizon of modes of being – percepts, meanings, objects, tools, values, and so forth – every experience reveals some part of the infinite series of the world³⁰ from a particular perspective and, only after, from a specific approach

generality and, with sufficient care, fix it once and for all in a way equally accessible to all [...] Prescientifically, the world is already a spatiotemporal world; to be sure, in regard to this spatiotemporality there is no question of ideal mathematical points, of “pure” straight lines or planes, no question at all of mathematically infinitesimal continuity or of the “exactness” belonging to the sense of the geometrical a priori. The bodies familiar to us in the life-world are actual bodies, but not bodies in the sense of physics» (Husserl, 1970b, p. §36).

²⁸ Husserl thematized the concept of life-world in different ways, e.g. the forgotten meaning-fundament of natural science, the unexplored presupposition for Kant’s philosophy, the pre-given world, the correlate of the natural attitude, the theme of historians who try to reconstruct the life-worlds of peoples, the theme of a non-transcendental ontology, the theme of a transcendental science, the unthematized horizon which never has been brought to explicit attention., an intuitive, real and concrete world, compared with the world of science as objective, ideal and abstract (Moran, 2012).

²⁹ «It should be noted, therefore, that talk of regions of Being receives its sense only in the light of this essentially necessary correlation with the modes of consciousness in which an entity of any region is given, and that there could be no other way to justify making such distinctions of regions of Being, if these distinctions are to be anything other than generalizations of concepts already acquired empirically. It is in this sense that Husserl develops the inquiry into the regions of Being as a transcendental philosophical inquiry, as an inquiry into the necessary correlation of Being and consciousness, and this means into the constitution of Being for consciousness. So understood, this inquiry is prior to any metaphysical decision that would interpret this relationship in an idealistic or realistic sense» (Landgrebe, 1981, p. 136).

³⁰ «The thing is given in experiences, and yet, it is not given; that is to say, the experience of it is givenness through presentations, through “appearings.” Each particular experience and similarly each connected, eventually closed sequence of experiences gives the experienced object in an essentially incomplete appearing, which is one-sided, many-sided, yet not all-sided, in accordance with everything

(philosophical, psychological, physical, mathematical, phenomenal, etc.)³¹. This position is opposed to the claim that the physical is a sort of super-class of facts in which other sub-sets (e.g., in ascending order, chemical, organic, biological, psychological, sociological) are inserted and dependent on it, and where, consequently, consciousness is considered a physical phenomenon (Jylkkä & Railo, 2019). And, interestingly, this does not imply that consciousness must be a non-physical or extra-natural entity. Indeed, following the rejection of the appearance-reality fallacy, I would prefer to say that consciousness is “also” a physical phenomenon (as objectual consciousness; Rowlands, 2001), according to the regional ontology of the mathematico-physical world; but this is not the point yet. The problem here is that someone may believe that the appearance-reality fallacy can be solved by throwing all “realities” and “appearances” into the cauldron of the mathematico-physical “noumenon”³². However, the absolutization of physicalist monism and mathematical method as – respectively – the ontological basis and the epistemological structure of the whole reality is a mere assumption made by the contemporary scientific paradigm. Instead, despite Husserl’s fierce – and maybe somewhat exaggerated³³ – criticism, Galileo did not divide the world in the real-independent one and the apparent-

that the thing “is.” Complete experience is something infinite. To require a complete experience of an object through an eventually closed act or, what amounts to the same thing, an eventually closed sequence of perceptions, which would intend the thing in a complete, definitive, and conclusive way is an absurdity; it is to require something which the essence of experience excludes» (Husserl, 2006, §14).

³¹ «There is more to experience than what actually appears from one’s current vantage point; there is also a sense of what would be revealed from another vantage point or by manipulating the object in various ways. Experience is structured by what Husserl and Merleau-Ponty call horizons, possible ways of perceiving and manipulating a thing that shape how the thing is actually experienced» (Ratcliffe, 2007, p. 488).

³² [...] [W]e defend a form of physicalistic monism, which holds that all concrete phenomena are physical—this is typically the starting point in modern science. Complex phenomena, such as those studied by chemistry, biology, or neuroscience, belong to the same ontological category as fundamental physical processes (those studied by fields such as particle physics), and are based on them. We could say that “physical” denotes the set that contains all concrete phenomena, and terms like “biological” or “neural” denote subsets of that all-encompassing set. By the term “physical” we mean the ontological category of phenomena studied by physics — thus, on our reading “physical” is a placeholder for whatever is the fundamental nature of phenomena studied by physics. [...] If we endorse the physicalistic premise that all concrete natural phenomena are physical, it follows that our experiences are physical phenomena. This might appear to be a contradictory claim, given that the term “physical” is often defined in opposition to “mental”. For instance, according to the Oxford Dictionary, “physical” is something “relating to the body as opposed to the mind”, or as “relating to things perceived through the senses as opposed to the mind” (emphasis added) (Physical, n.d.). Here we do not adhere to this standard usage of “physical”, because it a priori commits us to dualism. On our reading, terms like “experiential” or “conscious” denote certain natural phenomena which belong to the class of physical phenomena.

³³ «Certainly, Husserl’s Galileo is “idealized”, presented, by explicit recognition of the founder of phenomenology, as a symbol of the entire formation process of modern physics, without claiming to faithfully outline the image of the “historical Galileo”» (Bucci, 2006, p. 84, my translation).

dependent one, rather he tried to analyse some portions of the world, by means of the idealization of the nature – i.e. the approximation of the real event into a model-event with the lack of subjective properties and non-measurable contingencies, and it has nothing to do with the idea of a complete knowledge and ontological determination of the world, achieved by the exclusion of “useless” features related to the subject, as we see in some eliminative and reductionist projects (see chapter 3), where “to be is to be measurable”³⁴. From the fundamental and extraordinary attempt to rigorously measure natural phenomena thanks to the mathematical abstraction from the lifeworld, it is passed to the (alleged) gradual process of substitution of the lifeworld (Zelić, 2008).

4.3.3 From Ontological Pluralism to an Interdisciplinary and Relational Approach to Consciousness

In this sense, any attempt to interpret consciousness only as a product of physical and/or evolutionary processes remains within the framework of the appearance-reality dichotomy and those theories that entirely reduce consciousness to this kind of description slip into the appearance-reality fallacy. For example, the cognitive scientist Donald Hoffman in his last book *The Case Against Reality. How Evolution Hid the Truth from Our Eyes* argues that animals that perceive the environment “as it really is” are less likely to survive in the evolutionary competition rather than those whose consciousness constructs a user friendly but false world in order to survive (D. D. Hoffman, 2019)³⁵. Indeed, we see what we need to survive, so

³⁴ «To overcome the vagueness and relativity of ordinary experience, science performs a set of abstractions and interpretations upon the world as it originally presents itself. First it focuses upon the shape-aspect of the world [...], then it interprets these shapes as pure geometrical shapes in order to deal with them in geometrical terms. But it forgets that this first move is an abstraction from something and its second an interpretation of something [...]. Having forgotten the abstractive and idealizing role of scientific thought, the philosophical interpretation comes up with an ontological claim: to be is to be measurable [...]» (Carr, 1970, p. 240 in Zelić, 2008, p. 417).

³⁵ «Think of it like a video game. You have to hunt for points as fast as you can, and if you get enough, you get to advance to the next level. If you don't, you die. Evolution is the same. We're hunting for “fitness payoffs.” And if you get them, your offspring advances to the next level. Fitness payoffs are really what it's all about. Most of my scientific colleagues think that if you're looking for fitness payoffs, you're looking for the truth. My team and I found that to be wrong. Fitness payoffs themselves don't have info about the truth – they are what you have come to understand are important for survival and reproduction. If you're chasing fitness payoffs, you slowly lose info about the true structure of the world in favor of what is important for helping you pass on your genes. Inevitably, then, you can't see the true structure of the world. Organisms that see reality as it is are never more fit than those just tuned to see fitness payoffs» (D. D. Hoffman & Ashbach, interview, 2019).

that objects, properties and qualities in the world are like the icons in the desktop of our laptop, i.e. user illusions and simplistic representations of a deep and complex reality made by numbers, bits, etc. incomprehensible for the majority of us (Dennett, 2017). Folders and icons are not “really there” but they are crucial for a good use of the laptop (instead, for example, of directly writing codes without any graphic interface)³⁶. The problem here is not that consciousness – or the brain, from a more physicalist position – seems to be responsible for the construction of a world made by objects and properties that are useful for the organism to survive. In fact, as we also see later, we constantly experience things to be grasped, obstacles to be avoided, attractive shapes to be reached, etc., in a word, affordances (Gibson, 1979). All these invites, messages, meanings are not found inside the physical structure of the world but come from the interaction between the human organism and the environment. However, the claim that the world deprived of all the conscious constructs and illusions is the actual reality is a mere assumption which stems from the appearance-reality fallacy.

Now, instead of reducing consciousness to the mathematico-physical abstraction of the world, what I aim to is the construction of new relationships among different sets or layers of information on consciousness, instead of the definition of the most fundamental one according to the mainstream scientific paradigm (what causes what). Consequently, the kind of relations involved in this alternative ontological and epistemological paradigm may not be only the efficient causality, but it may be included the formal and final causations, and other interactions between ontological domains, like iteration, reciprocity, similarity, interpenetration, diffraction, where different sets of information are different “modes of appearance” or “adumbrations” (Husserl, 1913/1982) through which an object is given³⁷. Moreover, beyond new kinds

³⁶ «Another way to think about this concept is looking at the world as a user interface – a mask. Three-dimensional space is really just like a desktop on a computer. The stuff we see is just there as an icon. An icon of a book that’s blue doesn’t mean the book and its contents are actually blue – it’s a simple symbol representing something much bigger. Not seeing the truth actually helps you – if we had to know all the truth about the computer and how to operate every circuit to make each action needed, we wouldn’t use it. Evolution literally hides the truth from us on purpose. Perception lets us control the truth while we’re utterly clueless about what actually is» (D. D. Hoffman & Ashbach, interview, 2019).

³⁷ «The physical thing which he observes, with which he experiments, which he continually sees takes in his hand puts on the scale or in the melting furnace: that physical thing, and no other, becomes the subject of the predicates ascribed in physics, such as weight, temperature, electrical resistance, and so forth. Likewise, it is the perceived processes and concatenations themselves which become determined by means of concepts such as force, acceleration, energy, atom, ion, etc. The sensuously appearing

of relationship, there may be also new kind of questions along with the empirical-scientific ones, i.e. what consciousness is (description), how it comes to existence (explanation), which its functions are (functionality) (Van Gulick, 2018). Thus, there may be questions about the meaning of being conscious (existential conditions), the conditions of possibility of our lived experience (transcendental analysis). Moreover, on the epistemological ground, it would be possible to connect, following Burge's terminology, empirical explanations (how that phenomenon arises) and constitutive explanations (the general and conceptual conditions under which something can be what it is) (Burge, 2010).

This epistemological project towards a multidimensional phenomenon like consciousness is based on a relational structure and a pluralist approach rather than on a hierarchical one, besides it does not give priority to the empirical domain nor to the conceptual one, nor to the folk/naïve psychology from common-sense. Quoting Husserl:

«In the skeptical mood necessarily created by critical epistemological reflection (here I mean the initial reflection that occurs within the natural mode of thought and thus precedes any scientific critique of knowledge), every positive science and every positive scientific method ceases to count as an available possession [...] and here exact knowledge is no less enigmatic than non-exact knowledge, scientific no less than prescientific. What becomes questionable is the possibility of knowledge, more precisely, the possibility of knowledge making contact with an objectivity that is, after all, what it is in itself. [...] The most rigorous forms of mathematics and mathematical natural science here have not the slightest advantage over any actual or alleged knowledge belonging to common experience» (Husserl, 1999, pp. 20–21).

Also, as already seen, this pluralist approach recognizes the impossibility of a pure starting point devoid of any epistemic stance, approach, or perspective. Thus, within this relational epistemology and regarding our purposes here, the main aim is not to build a science of consciousness upon the empirical findings in order to determine the concepts, nor to define a priori the theories in order to force data into them. There is

thing, which has the sensuous shapes, colors, odor- and taste-properties, is thus anything but a sign for some other thing; rather it is, so to speak, a sign for itself» (Husserl, 1913/1982, p. §52).

no absolutization of any particular method or approach, rather (as we will see soon) a critique of the attempt to reduce all the different layers of knowledge to one, i.e. absolutizing it. In this sense, there is no room for a strong foundationalist account of knowledge³⁸, be it physicalist or phenomenological. For instance, perhaps *contra* Husserl³⁹, the point here is not to determine whether the phenomenological approach, despite its ability to reveal the background of meanings and objects upon which any scientific or pre-scientific account lies, may or may not provide the “critical” (Husserl, 1913/1982, §62) and “absolute grounding” (Husserl, 1931/1977, §3) of human knowledge (Husserl, 1902-1903/1999).

Therefore, in an interdisciplinary approach like this one, partially inspired by coherentist accounts of epistemic justification⁴⁰, the best explanation would not be the exposition of the most fundamental layer of causation (e.g. at the neuronal level or even deeper, like the quantum vibrations of the microtubules in the brain; Hameroff & Penrose, 2014), because this would mean that we are still stuck within the ontological reductionism-model. Instead, a relational epistemology must consider a particular system of knowledge, e.g. the one concerning philosophy, phenomenology and neuroscience, as a web made by different findings, conceptual schemes, models, analyses, in which every single “piece” maintains its independence, for example, an empirical discovery *per se* does not need a to be phenomenologically justified and *vice versa*, however the different sets of knowledge (e.g. about neural activity, phenomenological description and philosophical conceptualization regarding the dreaming state) get more explanatory power once connected together and interpreted

³⁸ «Foundationalism is a view about the structure of justification or knowledge. The foundationalist’s thesis in short is that all knowledge or justified belief rest ultimately on a foundation of noninferential knowledge or justified belief» (Hasan & Fumerton, 2018).

³⁹ Actually, there is still a debate about whether Husserl was really a convinced foundationalist: «Husserl was a foundationalist to the extent that he wanted phenomenology to be a pure presuppositionless science that provided an ultimate foundation for all other sciences. However, he was not a foundationalist to the extent that he accepted neither the rationalist nor the empiricist suppositions about absolute foundation. Furthermore, Husserl did not assume in advance that any particular science (e.g. mathematics) offered a

model or normative ideal for what science as systematic knowledge should be (see Cartesian Meditations § 3). In some respects Husserl was a coherentist in that he maintained that science itself consisted in a web of interrelated beliefs. Husserl’s main claim was that all knowledge worthy of the name had to be justified by evidence» (Moran & Cohen, 2012, pp. 128–129). In this respect, see also (Berghofer, 2018; Drummond, 1990; Hopp, 2008).

⁴⁰ «According to the coherence theory of justification, also known as coherentism, a belief or set of beliefs is justified, or justifiably held, just in case the belief coheres with a set of beliefs, the set forms a coherent system or some variation on these themes» (Olsson, 2017).

from an overall view of the phenomenon (e.g. hypothesis of an isomorphism between phenomenology and neurophysiology of dreams; Hobson, Pace-Schott, & Stickgold, 2000). Thus, the explanation of the phenomenon is as adequate and complete as it is able to hold together different perspectives and produce links through different levels of description. One can certainly present cases of contradiction between two levels of description, but this must not immediately imply the a priori elimination of one domain in favour of the other, first there should be an attempt to account for such a discrepancy. I may say that the key question is no longer about how the domain x can explain and incorporate the domain y, but how the x and y domains can be linked together.

4.4 Problems with Neuro- and Mento-centrism

So far, I have exposed the appearance-reality distinction from which some physicalist but also mentalist positions may derive. This fallacy produces the idea that there is a fundamental, deep, true reality “behind the scene” which may be either the physical or the phenomenal world. In the first case, the only true ontology is the one that can be explored by means of mathematics, for example through empirical measurements, leading to the conclusion that “to be is to be measurable”, while all the apparent characteristics (e.g. secondary qualities) are nothing but illusions⁴¹. In the second case, instead, consciousness plays an important ontological role, being a distinct entity from the physical realm, and more or less autonomous depending on the type of dualism (Chalmers, 2002). Against these paradigms, I showed how physicalism pretends to consider the world as devoid of any potential conscious perspective, while mentalism addresses consciousness as an object, reifying its features in a property-like way. Then, I argued for a rejection of the appearance-reality fallacy, proposing an approach that considers all the different ontological domains not as a hierarchy of realities – from the most fundamental to the illusory ones – but as different modes of manifestation of being within the very same pre-given and

⁴¹ «Physics, supposedly, is the supreme basic science and all other sciences are secondary and derivative from, reducible to, physics. We increasingly accept the concretely incredible: that things which are, ultimately, unable to be explained as causal functions of atoms in the void are, ultimately, illusory, unreal figments of “folk psychology”» (Thompson III, 1997, p. 224).

lifeworld, which can be understood as the ground of every intentional act and condition of possibility of all the other ontological dimensions (physical, phenomenal, psychological, social, etc.). The lifeworld is neither a noumenal world, the thing-in-itself which lies behind appearances, nor a phenomenalist production of a solipsistic subject, but the steppingstone from which – knowingly or unknowingly – we start any possible, action, intention, perception, investigation in and of the world. In other words, the world is not the totality of all the things (in-itself), but the condition of manifestation of being in all its appearances (emotions, thoughts, things, institutions, people, theories, etc.).

In this sense, all the different paradigms of beliefs, values, techniques (e.g. the neuroscientific or the philosophical ones) that constitute any ontological dimension can be seen not as elements of the causal chain (e.g. “the mind emerges from the brain” or “the mind creates the brain”), nor as an interaction between different domains (“the mental activity interacts with the brain activity”), but as “diffractions” (Costa & Cesana, 2019, p. 91) of the original unity of the very same phenomenon. Both the domains come from the same existence (e.g. a particular individual in a specific environmental context) and they cannot be conceived separately, as they were first independent and then they acted on each other. Now, focusing on these two ontological domains – the neural and the mental – when one of these paradigms is absolutized as the fundamental ontological base of consciousness or as the main source of knowledge about consciousness, this may lead – respectively – to neuro-centrism or mento-centrism⁴². Then, I shall argue demonstrate where and how these paradigms fail to give a consistent and complete account of consciousness.

4.4.1 Neuro-centrism

As seen in the previous chapter, current research programmes in neuroscience they are generally based on a physicalist and monist ontological commitment and characterized by a reductionist – if not even eliminativist – attitude towards consciousness, looking

⁴² I already exposed widely the genesis of neuro-centrism in the third chapter, while I dedicated less space to mento-centrism, however I consider the former a consequent product of the latter (see the genealogy of neuro-centrism from the apparent rejection of mind-body dualism in chapter 3).

for an explanation of how conscious experience emerges from the brain, by means of necessary and sufficient neural correlates (NCC). However, I already highlighted that this paradigm, on closer examination, seems to reveal an underlying dualism coming from the dichotomy between primary and secondary qualities and from the mechanistic approach developed by modern philosophers and scientists like Descartes and Galilei. Indeed, it may be said that contemporary neuroscience and cognitive sciences, instead of getting through the mind-body dualism, has reintroduced it in a new way – the dualism between the brain and the body-machine (Zilio, 2018a) or “crypto-Cartesianism”⁴³ – so much that most of the principles of Cartesianism (Wheeler, 2005) are still present within the core ideas of contemporary theories of consciousness and cognition (see above). After this exposition of the most general problematic issues related to the neuroscience of consciousness, I shall now go through some more specific details.

4.4.1.1 The Myth of Localization

First, there is what I would call the myth of localization, i.e. the idea that thanks to the incredible new technologies of brain imaging we are about to pinpoint the neural source/seat/cause/area/mechanism of consciousness, for instance, comparing healthy people and patients with various brain damages, and examining which lesions are relevant for the maintenance of the level of consciousness (e.g. the Ascending Reticular Activating System, ARAS; Fischer et al., 2016) and which ones instead elicit different cognitive disabilities. Similarly, the content of consciousness may be traced in the brain by measuring the complexity of integrated information in a physical system, in this case the brain, according to the Integrated Information Theory of consciousness (IIT; Tononi, 2017). The system – the brain – is a set of possible states within an internal structure with causal-effect relations, which determine the system itself. In a system there can be a complex – a set of elements with causal power – that generates a local maximum of integrated information ($\Phi^{\max}$). A set of these elements

⁴³ «This is a crypto-Cartesian view of the relation between our psychological attributes and the workings of the brain. It is rooted in the misconception, which we assailed in our book [...], that the essence of consciousness is its association with qualia and hence that these ineffable qualitative characteristics of experience might be found to be caused by cardinal or pontifical cells» (Bennett et al., 2007, p. 167).

shapes a “concept” of integrated information, that is impossible to subdivide due to the irreducibility of its cause-effect relations. In turn, an irreducible and indivisible structure of these concepts is called “conceptual structure” that composes the experience as such: the configuration of the structure defines the quality, while the level is defined by the quantity of integration (Oizumi, Albantakis, & Tononi, 2014). Apart from this specific theory, there is a general problem with this kind of approach, which – I believe – is shared by almost all neuroscientific theories of consciousness. This problem is the idea that a good account of consciousness must “localize” where conscious experience is or arises (e.g. Lloorits, 2018), in so far as localization means explanation and identification⁴⁴.

Epistemologically speaking, it is unclear how localization means explanation, i.e. the fact that we can detect some specific brain area that show correlation – or even assuming causation – with a specific conscious state is not an explanation of what consciousness is and how it works. A lesion to the ascending reticular activating system may lead to loss of consciousness, like in coma due to stroke. Metaphorically speaking, arguing that activation of ARAS explains in some way – even partially – consciousness it is like saying that the electric current delivered by spark plugs can explain what an engine is and how it works. Ontologically speaking, even assuming that some particular brain area may cause consciousness, this does not mean that that area is identical with the act of consciousness. Even adopting a consciousness-brain identity, it is not clear how and why the localization of relevant brain activity regarding consciousness necessarily lead to the identity of the two phenomena, since that experience may arise from the physical, but it is not necessarily entailed by it (Chalmers, 1995)⁴⁵. Metaphorically speaking, arguing that that particular brain

⁴⁴ When I speak of localization, I am not referring to the methodological debate about localizationist and holistic views about the brain, that is, the old question from neuroscience and neurology about localizing mental states and their contents in particular regions or networks of the brain, or alternatively to holistically consider the whole brain as participating in processing mental contents. (Northoff, 2014a). The issue here discussed instead is mainly epistemological and ontological.

⁴⁵ Regarding IIT in particular, «IIT seems to make the argument that if experience arises from the structure and dynamics of integrated information states, then it is entailed by those integrated information states, and so posits an identity to explain that entailment. Yet, this move should give one pause — just because consciousness might arise from integrated information does not mean that it is identical to integrated information» (Mindt, 2017, p. 142). «[...] IIT could entail a concept of identity not as equation, but as the asymmetric relation of composition (Place, 1956). From here, the relation between conscious experience and integrated information would be similar to the sentence “lightning is (composed by) electrical discharge, but not vice versa; clouds are (composed by) water molecules, but not vice versa”. Similarly, we can say that “conscious experience is composed by integrated information

activation/information network is identical with conscious experience, it is like saying that the electric current delivered by spark plugs is identical with the functioning of the car's engine.

Moreover, the fact that conscious experience might be elicited by brain activity does not mean that the process of “being conscious of something” is restricted to neural states (Rowlands, 2010a), indeed every intentional act, as we will see in details in the next chapter, includes many different sub-processes involving the neuronal firing, bodily actions (sensory-motor coupling) and, of course, the intentional object emerging from the environmental background. Even assuming a representational account of consciousness, it remains unclear how, where, and why a representation of “me intentionally approaching a chair in front of my body” can be located in the head, precisely in some specific neuronal pattern. From an epistemological perspective, the fact that we nowadays we can find evident changes in the brain, while it seems more difficult to understand and describe bodily-environmental interactions and loops correlated with conscious activity, does not imply that consciousness is in the brain or is caused only by the brain.

Taken together, the claims regarding alleged localizations of mental states in the head/brain result from a process of abstraction and subsequent reification of reality, which Whitehead called the fallacy of simple location⁴⁶, i.e. the error of assuming that anything that is real must have a simple spatial location in absolute space and time within which there are independent unchanging bits of things and objects as fixed, static, non-temporal, and non-relational (Northoff, 2016d; Weber & Weekes, 2009).

in a complex”, but at the same time there is no sense in saying that “integrated information in a complex is composed by conscious experience”» (Zilio, 2019, p. 113).

⁴⁶ «What I mean by matter, or material, is anything which has this property of simple location. By simple location I mean one major characteristic which refers equally both to space and time, and other minor characteristics which are diverse as between space and time. The characteristic common to both space and time is that material can be said to be here in space and here in time, or here in space–time, in a perfectly definite sense which does not require for its explanation any reference to other regions of space–time. Curiously enough this character of simple location holds whether we look on a region of space–time as determined absolutely or relatively. For if a region is merely a way of indicating a certain set of relation to other entities, then this characteristic, which I call simple location, is that material can be said to have just these relations of position to the other entities without requiring for its explanation any reference to other regions constituted by analogous relations of position to the same entities. In fact, as soon as you have settled, however you do settle, what you mean by a definite place in spacetime, you can adequately state the relation of a particular material body to space-time by saying that it is just there, in that place; and, so far as simple location is concerned, there is nothing more to be said on the subject» (Whitehead, 1925/1970, p. 50).

This fallacy can be found in several disciplines and the attempts made by some neuroscientific researchers to localize the precise point/instant in the brain situated into an outer spatio-temporal dimension from which consciousness should rise, is exactly an example of simple location⁴⁷. The error is not the methodology *per se*, since neuroscience needs to pinpoint mathematically the brain activity, in order to operationalize the mental feature in question, the problem is when the abstraction made according to the scientific paradigm becomes an ontological claim, i.e. methodological-ontological fallacy⁴⁸.

4.4.1.2 The Mereological Fallacy

One of the main consequences of this localizationist perspective towards consciousness and the brain is the so-called mereological fallacy (Bennett & Hacker, 2003) (or homunculus fallacy; Kenny, 1971), i.e. the attribution of specific characters and properties that do not belong to the part but to the whole system, in this case the human being⁴⁹. In this context, because the brain seems to be a causal condition for mental phenomena, someone might ascribe psychological predicates to parts of the human being that apply only to the whole, e.g. the brain is conscious, sees, perceives, interprets information, feels, etc.⁵⁰. However, as I explained earlier, even if the brain played an essential role for consciousness concluding that the experiencing subject is the brain is a fallacy. It is not the brain that is the subject of experience, rather the brain

⁴⁷ «Applied to the brain [the simple location] this means that the brain cannot be conceived as thing or object at one particular point or location in time and space within the world. Accordingly, a property-based ontology of brain may be empirically implausible even of logically plausible» (Northoff, 2016d, p. 3).

⁴⁸ «The concept of the “methodological–ontological fallacy” concerns the fallacious inference from features (like discrete points in time and space) based on the application or use of a particular methodology (like the scientific method) in investigating things and objects to their underlying ontological existence and reality in the world» (Northoff, 2016d, pp. 6–7).

⁴⁹ The fallacy is not related merely to psychological and neuroscientific sciences, rather it has been developed long before these contemporary issues, for example by Aristotle, Wittgenstein, Lewes, and others (Smit & Hacker, 2014).

⁵⁰ «For example alterations at the level of cells and neural circuits are a prerequisite for retaining information, as mutations affecting the processes of long term potentiation and depression show. But although these alterations are a prerequisite for the retention of information, they do not show that cells, neural circuits, or brain structures retain information. For it makes sense to say that the animal as a whole has a memory, is able to retain information, and possesses knowledge, but it makes no sense to say that for example the hippocampus or cortex possesses knowledge. It is the animal as a whole that retains information and exercises the ability to retain information, not its parts» (Smit & Hacker, 2014).

is a part of the experiential structure⁵¹. And this “mereological mode” of conceiving the human psychology in relation to the brain reveals again the crypto-Cartesianism rooted in the current neuroscientific research, indeed correlating mental properties solely to the brain produces, as a reflex reaction, an undesired dualism, i.e. the isolation of the brain and the transformation of the body as a passive machine without psychological properties, even more aporetic than the original mind-body dualism, since the former insists on a difference between part and whole of the very same entity.

4.4.1.3 Fallacy of Misplaced Concreteness

Another problem for neuro-centrism related to the modern scientific paradigm is the interpretation of the brain as the “creator” or – softer version – “modulator” of the qualitative aspects of the world. As already seen, the modern dichotomy between primary and secondary qualities has been transposed into the conceptual scheme of current neuroscience, so that the “external” world is sometimes conceived as devoid of any quality, not only colours and sounds but also shapes, tastes, etc. When these “neutral” stimuli coming from the environment meet our body, they are processed by the brain that selects and filters them, producing all the qualities we meet during perception, a little like wearing coloured lenses in a black and white world or, as Seth, Clark and Hoffman probably would say, being constantly hallucinated by the brain (see above)⁵². This can be understood as what Whitehead called fallacy of misplaced concreteness, i.e. «the accidental error of mistaking the abstract for the concrete»

⁵¹ «The subject of experience is not a bit of your body. You are not your brain. The brain, rather, is part of what you are» (Noë, 2009, p. 7).

⁵² This is what Whitehead called the bifurcation of nature: «The primary qualities are the essential qualities of substances whose spatio-temporal relationships constitute nature. [...] But the mind in apprehending also experiences sensations which, properly speaking, are qualities of the mind alone. These sensations are projected by the mind so as to clothe appropriate bodies in external nature. Thus the bodies are perceived as with qualities which in reality do not belong to them, qualities which in fact are purely the offspring of the mind. Thus nature gets credit which should in truth be reserved for ourselves: the rose for its scent: the nightingale for his song: and the sun for his radiance. The poets are entirely mistaken. They should address their lyrics to themselves, and should turn them into odes of self-congratulation on the excellency of the human mind. Nature is a dull affair, soundless, scentless, colourless; merely the hurrying of material, endlessly, meaninglessly» (Whitehead, 1925/1970, pp. 55–56).

(Whitehead, 1925/1970, p. 59)⁵³, of which the simple location is an example⁵⁴. Many theories of consciousness mistake internal correctness of the model for external adequacy to the reality it was devised to explain⁵⁵. But, again, this leads to the transition from a methodological issue to an ontological one, in this case from the identification of relevant correlation between brain activity and phenomenal experience (e.g. when I see red, the brain area x activates) to the attribution of mental properties to the brain (e.g. when I see red, the brain area x is processing neutral stimuli, adding in the experience the perception of red). Just because we conceive the brain as a stimulus processor for our scientific purposes, this does not mean that the brain is like that⁵⁶. As I will argue in the next chapter, the brain does not create the world nor assembles isolated impulses into an unified representation of it, rather it is situated and embedded and dependent from the start within an environment (and a body) that already has spatio-temporal and meaningful structures.

4.4.2 Mento-centrism

The fact that neuro-centrism seems to be controversial regarding both its epistemology and ontology does not mean that a mento-centric approach upon consciousness is necessarily better. Many philosophers consider consciousness as a mental entity or activity that can be conceived and analysed – partially or completely – independently from the brain. As we have already seen in the first chapter, there can be substantial and non-substantial dualisms, idealisms and solipsisms, etc. An extreme version of this defensive approach towards the independence of consciousness may be called mento-

⁵³ «Whitehead wants to say that our concrete experience of real particular situations is not respected by scientific analyses that narrowly abstract from that situation and build pictures of reality based upon only certain restricted fields of experience» (Thompson III, 1997, p. 229).

⁵⁴ Whitehead is not the first and the only one that explicitly exposed this error of confusing the abstraction for the concrete. Other ways of describing this genre of fallacies is reification, hypostatization, “the map is not the territory”.

⁵⁵ In this sense, going at the root of the problem, even the primary-secondary dichotomy (the bifurcation of nature; Whitehead, 1920) and the ontological priority given to the former group of qualities are instances of mistaking an abstraction with the concrete reality, i.e. fallacy of misplaced concreteness.

⁵⁶ «In either case of ascribing physical or mental properties to the brain, the methodological domain as introduced by the scientific observation of the brain at particular points in outer time and space, is not clearly detached from the ontological domain. [...] The characterization of the brain by either physical (or mental) properties rendering the brain a physical (or mental) object or thing may then be traced to a methodological–ontological fallacy or, as Whitehead would say, a “fallacy of misplaced concreteness”» (Northoff, 2016d, p. 10).

centrism, where the mind-based epistemic stance upon the world is mistaken for the ontological foundation of the world⁵⁷. In order to understand better this point, we may take again the figurative example I made above, regarding Earth and Sun. We can conceive Earth revolving around the Sun, even if our personal standpoint as terrestrial inhabitants show us the Sun moving around Earth (but still, as shown, we cannot replace that way we see the Sun from Earth). Metaphors aside, we can evade our subjective epistemic stance upon consciousness in order to look at it in a more detached and objective way, e.g. investigating the correlation with brain activity, but this cannot replace our starting approach upon consciousness, i.e. as conscious perceivers. Now, if the absolutization of the objective approach may lead to neuro-centrism, i.e. consciousness is a product of the brain, on the other hand the absolutization of the subjective stance may lead to mento-centrism, i.e. consciousness is analysable only through our mental introspection.

4.4.2.1 Hypostatization of Psychic Objects

As well as neuro-centrism, mento-centrism commits a sort of fallacy of misplaced concreteness, mistaking an abstraction (e.g. categorization of mental features) for the concrete reality. There are different examples of this kind of error, and I shall expose them through the analysis Sartre made about the mistakes we may make when we reflect introspectively upon conscious experience (Eshleman, 2016). As seen in the second chapter, during our daily life we are generally immersed in our activities within the environment, without any specific and overt awareness of ourselves, rather we are pre-reflectively conscious of what we are conscious of, like an implicit and tacit sense of self-awareness that may become explicit in some situations. These situations can be cases of environmental anomalies, new physical activities to learn (e.g. driving or dancing), when we focus on what we are doing, when we feel like someone is watching

⁵⁷ «Moreover, such mento-centric view of the world is, at the same time, egocentric or anthropocentric rather than allocentric or eco-centric (see below for the latter). Since we suppose that the mind as ontic center of the world also provides our own ontic location within that very same world, the mento-centric view of the world cannot be but egocentric or anthropocentric: we conceive the world in terms of our own mental features and thus our mind. The identification of mind as both ontic center and ontic location excludes any possible allocentric or eco-centric view of the world—the latter views are simply not included as epistemic options in the logical space of knowledge as entailed by a vantage point from within mind» (Northoff, 2018b, p. 383).

us⁵⁸, during major life decision in which we need to retrospectively look at our personal history and predicted future, when, of course, we theorize about consciousness and so forth. In this kind of cases, reflective consciousness – like King Midas who transformed everything he touched into gold – explicitly posits everything (and also us) as an object of consciousness, e.g. our mental states, our thinking upon something, our physical activity, our own conscious activity, etc.

A clear example of this passage from a pre-reflective to a reflective consciousness is made by Sartre in *Being and Nothingness* about pain. Pain is difficult to describe, so that we often use analogies, metaphor and allegories for making it explicit. Generally, we usually say something like “my eyes hurt” like if “I feel pain in the eye”. According to Sartre, pain as psychic object, i.e. a particular mental or psychophysical state, is the product of impure reflection that thematizes pain like an object we have in the body or in our consciousness (“I have a headache...now the headache is gone...now is coming back”). Before thematizing pain as an illness (*mal*) in the body⁵⁹, like “I have a headache”, pain manifest itself through a pre-reflective consciousness (of) pain (*doleur*), which exists only as it appears to consciousness, like consciousness (of) something painfully (as a mode of existence or perceiving; Olivier, 2003). Sartre made the example of vision-as-pain: we are reading a book, while the letters start trembling, we periodically stop and resume the reading and we need to read three times a blurry line in order to understand it. In that moment we are conscious of what we are reading in a painful way, i.e. the pain is not – yet – “in” the eyes, but

⁵⁸ See the concept of the look in Sartre: «Let us imagine that moved by jealousy, curiosity, or vice I have just glued my ear to the door and looked through a keyhole. I am alone and on the level of a non-thetic self-consciousness. This means first of all that there is no self to inhabit my consciousness, nothing therefore to which I can refer my acts in order to qualify them. They are in no way known; I am my acts and hence they carry in themselves their whole justification. I am a pure consciousness of things, and things, caught up in the circuit of my selfness, offer to me their potentialities as the proof of my non-thetic consciousness (of) my own possibilities. This means that behind that door a spectacle is presented as "to be seen," a conversation as "to be heard." [...] But all of a sudden I hear footsteps in the hall. Someone is looking at me. What does this mean? It means that I am suddenly affected in my being and that essential modifications appear in my structure - modifications which I can apprehend and fix conceptually by means of the reflective cogito. [...] First of all, I now exist as myself for my unreflective consciousness. It is this irruption of the self which has been most often described: I see myself because somebody sees me - as it is usually expressed» (Sartre, 1956, p. 317).

⁵⁹ For example, see Toombs' four-stage model of illness, inspired by Sartre's ontology of the body, whereby the body is gradually objectified: it ceases to be something lived and transcended by me and instead becomes the biological organism researched by science (Toombs, 1992; see also Svenaeus, 2006).

the pain is “the” eyes⁶⁰, our consciousness posit the contents of the book in a painful way, while pain is in no way an intentional object of consciousness (Olivier, 2003). Only then we may “find out” that we have a slight headache, in other words we may tend to reflect upon our own pre-reflective conscious act lived painfully and transform it in a psychic object, like a mental psychophysical state in the body-as-object⁶¹, for a better re-presentation of it⁶². As Sartre said, the construction of a psychic object by reflection, i.e., a-spatial and a-temporal, cognitive, representative, conceptual, is a degraded form compared to the pre-reflective consciousness (of) pain, i.e. immediate, non-cognitive, non-objectual.

⁶⁰ «It must be understood that if pain is given as pain “in the eyes,” there is no mysterious “local sign” there nor any knowledge either. Pain is precisely the eyes in so far as consciousness “exists them.” As such it is distinguished from other pain by its very existence, not by a criterion nor by anything added on. To be sure, the expression pain in the eyes supposes a whole constitutive work which we shall have to describe. But at this stage in the argument, there is not as yet any reason to consider this, for it is not made. Pain is not considered from a reflective point of view; it is not referred back to a body-for-others. It is the-eyes-as-pain or vision-as-pain; it is not distinguished from my way of apprehending transcendent words. We ourselves have called it pain in the eyes for the sake of clarity; but it is not named in consciousness, for it is not known. Pain in the eyes is distinguished from other possible pains inexpressibly and by its very being. This pain however does not exist anywhere among the actual objects of the universe. It is not to the right or to the left of the book nor among the truths which are revealed through the book nor in my body-as-object (the body which the other sees and which I can always partially touch and partially see), nor in my body-as-a-point-of-view as the latter is implicitly indicated by the world. Neither must we say that the pain is an “overprint” or that it is like a harmony “superimposed” on the things which I see. Those are images which have no meaning. Pain then is not in space. But neither does it belong to objective time; it temporalizes itself, and it is in and through this temporalization that the time of the world can appear. What then is this pain? Simply the translucent matter of consciousness, its being-there, its attachment to the world, in short the peculiar contingency of the act of reading. The pain exists beyond all attention and all knowledge since it slips into each act of attention and of knowledge, since it is this very act in so far as the act is without being the foundation of its being» (Sartre, 1956, pp. 332–333).

⁶¹ Another example: «We can find a helpful hint in Husserl’s discussion of a toothache: ‘the perceived object is not the pain as it is experienced [lived through], but the pain in a transcendent reference as connected with the tooth’ (Husserl, 1970a, p. 866). Pain as it is experienced or lived through by the bodily subject is a characteristic of the body-as-subject; and this is different from the pain perceived as a characteristic of the body-as-object, as when I point out the tooth that hurts to the dentist» (Gallagher & Aguda, 2015, p. 98).

⁶² «The first movement of reflection is therefore to transcend the pure quality of consciousness in pain toward a pain-as-object. Thus if we restrict ourselves to what we have called an accessory reflection, reflection tends to make of pain something psychic. The psychic object apprehended through pain is illness. This object has all the characteristics of pain, but it is transcendent and passive. It is a reality which has its own time, not the time of the external universe nor that of consciousness, but psychic time. The psychic object can then support evaluations and various determinations. As such it is distinct even from consciousness and appears through it; it remains permanent while consciousness develops, and it is this very permanence which is the condition of the opacity and the passivity of illness. But on the other hand, this illness in so far as it is apprehended through consciousness has all the characteristics of unity, interiority, and spontaneity which consciousness possesses-but in degraded form. This degradation confers psychic individuality upon it» (Sartre, 1956, p. 335).

Using Sartre's description of pre-reflective, non-positional pain-consciousness and reflective, positional consciousness of illness, we can understand that the error that someone can make is to confuse the abstractions and objectifications produced by reflection for the pre-reflective lived experience, and this kind of misplaced concreteness may happen in different ways. Matthew C. Eshleman listed six errors, some of them are explicitly named by Sartre, other are indirectly derived from many of his works (Eshleman, 2016). The projection error, i.e. projecting aspects of reflective consciousness (e.g. thinker-thought, subject-object, substance-attribute distinctions) on to pre-reflective consciousness (Sartre, 2004a, 2004b)⁶³; the reification error (also related to bad faith; J.-P. Sartre, 1956), i.e. application of qualities suitable for objects of the world to our conscious experience; the illusion of immanence, i.e. the attribution of spatial qualities to consciousness, understanding it as a spatial container of mental states, representations, ideas, psychic objects in general⁶⁴ – this error can be attributed also to the neuro-centric attitude to find out where consciousness is inside the brain; the warping of thought, i.e. when we transform representational and ideal schemes and analogies into real qualities of consciousness⁶⁵; the illusion of the primacy of knowledge, i.e. the Cartesian error of confusing the knower-known dyad produced by positional and reflective consciousness for an essential structure of pre-reflective consciousness, transforming the latter in the

⁶³ «The "Psychic" is given solely to a special category of cognitive acts, the acts of the reflective For-itself. On the unreflective plane, in fact, the For-itself is its own possibilities in the non-thetic mode; and since its possibilities are possible presences to the world beyond the given state of the world, what is revealed thetically but non-thematically across these possibilities is a state of the world synthetically bound with the given state. [...] We use the term Psyche for the organized totality of these virtual and transcendent existents which form a permanent cortege for impure reflection and which are the natural object of psychological research» (Sartre, 1956, p. 163).

⁶⁴ «At the first reflective glance, we see that we have so far committed a double error. We thought, without justifying it to ourselves, that the image was in consciousness and that the object of the image was in the image. We depicted consciousness as a place peopled with small imitations and these imitations were the images. Without any doubt, the origin of this illusion must be sought in our habit of thinking in space and in terms of space. I will call it: the illusion of immanence» (Sartre, 2004a, p. 5) «[...] the image cannot in any way be reconciled with the necessities of synthesis if it remains an inert psychic content. It can only enter into the stream of consciousness if it is itself a synthesis and not an element. There are no and there couldn't be any images in consciousness. But the image is a certain type of consciousness. The image is an act and not a thing. The image is consciousness of something» (Sartre, 2012, p. 144).

⁶⁵ «One can see the danger; a slight preference is enough, it is enough to momentarily consider the spatial relations of the schema for themselves and to let them be affirmed or modified in accordance with the laws proper to spatiality: the thought is irremediably warped, we no longer follow the idea directly, we think by analogy. It appears to me that this insensible degradation of thought is one of the most frequent causes of error, particularly in philosophy and psychology» (Sartre, 2004a, p. 117).

knowledge activity⁶⁶; the standpoint error, i.e. the attempt to take a completely detached point of view upon our own point of view, as if the reflective activity could make different from itself (see the Sun-Earth metaphor above)⁶⁷. To sum up, these psychological and philosophical errors about reflection and introspection can be gathered under the category of hypostatization of the abstractions made by our reflection and report upon our lived experience, i.e. confusing the representation of our conscious life with the conscious life as such.

4.4.2.2 Confusing Actual and Objectual Consciousness

This is related with the distinction I made in the first chapter regarding actual and objectual consciousness (Rowlands, 2001). Objectual consciousness refers to consciousness as a collection of items and properties and this can regard both monistic and dualistic approaches. What is important here is not to confuse objectual consciousness with actual consciousness, understood as the very same disclosing activity of experience. As Sartre would say, consciousness is not related to its experiences as knower to known, subject towards an object or as a set of properties (qualitative or quantitative). Actual consciousness is not related to its experiences, rather it is the experience. Addressing consciousness only as made of phenomenal properties, material causes, features like representations, ideas, sensations, etc. means to have in mind merely the objectual consciousness, which is still important as we need it in order to have an analogical description of what it is like to experience. However, the error of mento-centrism is to replace objectual consciousness with actual consciousness, i.e. with the experiential act as such. Even “consciousness” when hypostatized as a property or an entity falls into the fallacy of misplaced concreteness.

⁶⁶ «Thus reflection has no kind of primacy over the consciousness reflected-on. It is not reflection which reveals the consciousness reflected-on to itself. Quite the contrary, it is the non-reflective consciousness which renders the reflection possible; there is a pre-reflective cogito which is the condition of the Cartesian cogito. [...] as I showed earlier, we must avoid defining pleasure by the consciousness which I have of it. This would be to fall into an idealism of consciousness which would bring us by indirect means to the primacy of knowledge. Pleasure must not disappear behind its own self-consciousness; it is not a representation, it is a concrete event, full and absolute» (Sartre, 1956, pp. liii–liv).

⁶⁷ «Reflection is a knowledge; of that there is no doubt. It is provided with a positional character; it affirms the consciousness reflected-on. But every affirmation, as we shall soon see, is conditioned by a negation: to affirm this object is simultaneously to deny that I am this object. To know is to make oneself other. Now the reflective can not make itself wholly other than the reflected-on since it is-in-order-to-be the reflected-on. Its affirmation is stopped halfway because its negation is not entirely realized. It does not then detach itself completely from the reflected-on, and it can not grasp the reflected-on “from a point of view.”» (Sartre, 1956, p. 155).

More generally, the error is the confusion of the dimension of the phenomenal (mental features, qualitative properties, psychological states, etc.) with the lifeworld, which instead is the revealing condition of possibility of all the other dimensions of appearances.

That said, those theories that appeal to the irriducibility of peculiar qualitative states and properties (like qualia) or entities (like mind) risk to reify a theoretical object coming from abstraction and mistake it for the concrete experience, which is the final and real subject of a theory of consciousness: «Descriptions of consciousness in terms of subjective sensations, raw feels and the like are selective abstractions from consciousness that bear little relation to the structure of experience» (Ratcliffe, 2007, p. 484). As Rowlands said, from the fact that when we have an experience of a red tomato we feel like there is something to be conscious of a red tomato, this does not mean that there are entities in the world like the redness of the tomato⁶⁸. And from the fact that we can posit the redness of the tomato as an intentional object of consciousness, this does not mean that we can ontologically posit an entity called “redness”. If we metaphysically overload the concept of qualia we fall into one of the errors here above exposed and we slip into the fallacy of misplaced concreteness in the same way as the neuro-centric position, looking for an ontological chimera. Take, for example, the notions of representation and sensation, which are paradigmatic examples of the appearance-reality dichotomy and, according to Sartre, respectively invented by philosophers and psychologists. They can be useful as analogies and concepts for the determination of a precise experiential state, but the constant risk is to hypostatize them like things we have “in the mind”, i.e. to confuse the map for the

⁶⁸ «That an experience seems or feels a certain way to the subject that has it does not entail that the subject

is aware of certain qualia or phenomenal properties of that experience. Indeed, it is false that the subject is aware of these items. Qualia, or phenomenal properties, it is commonly supposed, are the phenomenal constituents of the what it is like of experience. The term ‘qualia’ has, perhaps, so far been conspicuously absent in a book that purports to be concerned with the what it is like of experience. I have, of course, effectively discussed qualia at some length in the previous chapter, it is just that I chose to refer to them as phenomenal properties. This is no accident. Talk of qualia tends to bring with it a lot of theoretical baggage, most of it distinctly unhelpful. That is, qualia, understood as the apprehended phenomenal properties of experience, are often loaded down with various other properties, of varying degrees of implausibility, and then attacked from this standpoint. Thus, we are often told that qualia are intrinsic properties of experiences, that our access to them is privileged, that we can know incorrigibly of their presence, perhaps because we are directly aware of them, and so on. But to attack qualia because of their possession of a series of improbable properties that one has oneself foisted on them smacks of cheating» (Rowlands, 2001, p. 181).

territory. If the experimenter asks me how, “in my opinion”, “my sensation” of the light in the screen is, he believes to look something somewhere inside my mind or my brain, while actually the fact refers to my relation with the light, which is something out there in the world⁶⁹. In other words, «[e]xperience of the world is completely overlooked by such descriptions; a sense of being part of a world is replaced by reference to a few free-floating sensations. The exercise is analogous to that of attempting to describe the structure of a cathedral by focusing only on a few choice tiles that one has noticed hanging off its roof» (Ratcliffe, 2007, p. 485). In this sense, besides hypostatizing psychic objects in place of the experience as such, those theories of consciousness focused exclusively on the nature of qualia, raw feels or sense-data commit the mereological fallacy, since they often refer to a sort of “internal” subject accompanied by these mental, private and ineffable states that has nothing to do with the body and the “external” world. As mentioned in the first chapter, they confuse phenomenality (internal mental life, qualia, subjective states, what-it-is-likeness) with phenomenology, which instead is interested in «the very dimension of givenness or appearance [...] beyond any divide between physical interiority and physical exteriority, since it is an investigation of the dimension in which any object – be it external or internal – manifests itself» (Zahavi, 2007, p. 31).

Taken together, both neuro-centrism and mento-centrism present several issues – methodological, epistemological and ontological – but more importantly they both are products of the appearance-reality fallacy, which precludes an adequate analysis of consciousness as experience. Indeed, both approaches constantly risk to reify consciousness into brain or mental states, founding their studies upon an abstraction or – even more – a “caricature” (Ratcliffe, 2007) of experience. In the next chapter,

⁶⁹ «In this sense everything is given, everything is present to me without distance and in its complete reality. Nothing of what I see comes from me; there is nothing outside what I see or what I could see. Being is everywhere around me; it seems that I can touch it, grasp it; representation, as a psychic event, is a pure invention of philosophers. [...] Such is the notion of sensation. We can see its absurdity. First of all, it is pure fiction. It does not correspond to anything which I experience in myself or with regard to the Other. We have apprehended only the objective universe; all our personal determinations suppose the world and arise as relations to the world. Sensation supposes that man is already in the world since he is provided with sense organs, and it appears in him as the pure cessation of his relations with the world. [...] Sensation, a hybrid notion between the subjective and the objective, conceived from the standpoint of the object and applied subsequently to the subject, a bastard existence concerning which we cannot say whether it exists in fact or in theory-sensation is a pure daydream of the psychologist. It must be deliberately rejected by any serious theory concerning the relations between consciousness and the world» (Sartre, 1956, pp. 217, 314, 315).

we need to find out how to deal with conscious experience without reducing it into one single dimension (psychological, phenomenal, physical, etc.), that is, identifying a methodological approach which may pursue interdisciplinarity and the integration of different modes of appearances. Moreover, I shall address again the structures of conscious experience delineated in the second chapter – phenomenality, intentionality, self-awareness – and place them into a broader phenomenological context, introducing the concept of being-in-the-world as the fundamental condition of consciousness. Lastly, this concept will be investigated in both neuroscientific and phenomenological ways.

5. Pluralist and Interdisciplinary Approaches to Consciousness

In this chapter, I propose a perspective which is mainly grounded on Sartre's phenomenological ontology, as an alternative to reductionist, mento-centric and neuro-centric approaches. First, I will present a method of analysis that considers on the same ontological level various dimensions of our experience, physical, phenomenal, psychological, etc., which consequently needs the contribution of many disciplines; in particular I will argue for a fruitful interconnection between phenomenology, philosophy and neuroscience, exposing different types of non-reductive approaches (neurophenomenology, non-reductive neurophilosophy, neuropragmatism, 4E cognition). Secondly, regarding the concept of consciousness as such, I will focus on the lived experience as the starting point and at the same time the presupposition of every other experience (introspection, empirical analysis, conceptualization, etc.). From here, the analysis will focus mainly on the necessary condition of consciousness as being-in-the-world, in a concrete sense (we exist as immersed in an objectual and pre-given context) as well as the activity of sense-disclosing (consciousness is at the same time an activity of meaning-apprehending and meaning-making in the world). In this sense, the provisional features I have introduced in the second chapter, will here be inserted within the context of consciousness as being-in-the-world, which works as a pre-condition for intentionality, phenomenality and self-awareness. The attempt will be to propose a non-mentalistic and non-reductionist account of experience, starting with a philosophical-phenomenological analysis that will then develop it in chapter six through a phenomenological-neuroscientific approach. In the last part of the chapter, I will also build a bridge between Sartre's phenomenology and some contemporary philosophical position about consciousness, like intentionalism, representationalism, disjunctivism and adverbialism, to ensure greater conceptual clarity.

5.1 Looking for an Alternative Approach

As we have seen in the second chapter, so far I have only provide a provisional structure of consciousness, from a logical and a phenomenological analyses, made by phenomenality, i.e. the fact that there is something like to be conscious/to undergo an experience; intentionality, i.e. the fact that to be conscious is always to be conscious of something; self-awareness, i.e. the fact that any conscious state of something intrinsically involves a form of awareness of itself as being conscious (not necessarily a reflective, cognitivist, high-order self-consciousness). However, these aspects still lack a concrete characterization in terms of ourselves as human beings within an environment of which we are constantly aware. In other words, those aspects are only (some of) the bones of a body that still needs flesh and tendons for standing up. With chapters 3 and 4 I have tried to expose the mainstream lines of research on consciousness – particularly the neuroscience of consciousness – by identifying the ontological and epistemological roots that underpins their development, i.e. the appearance-reality dichotomy. After having raised all the issues that the appearance-reality fallacy brings, it seems hard to pursue both the neuro-centric and mento-centric models of consciousness without stumbling into perplexities and problems. On the other hand, even though these two extreme positions present undeniable issues, this does not mean that all the neuroscience of consciousness or all the philosophy of mind are wrong about consciousness.

5.1.1 Regional Ontologies

As pointed out in the previous chapter, the right move might be to shift from the appearance-reality fallacy, towards an ontologically, epistemologically pluralist and methodologically interdisciplinary approach to consciousness. Ontologically speaking, the rejection of any assumption about sub-structures of reality, like the mathematico-physical world of physicalism and reductive naturalism or the mental world of idealism and solipsism. Epistemologically speaking – in relation to the ontological pluralism – the rejection of the limitation to specific epistemic stances, like

the third- or first-person perspective as sources of knowledge about experience. Methodologically speaking, the openness to different approaches towards consciousness, like the neurological, the phenomenological and the conceptual ones. In other words, we need to maintain an ontological commitment to experience understood as, at the same time, the to-be-determined by different regional ontologies¹ and the determining factor of these regions² (see “key explanandum” and “key explanans” in the first chapter). These regions reflect sets of entities which are appearances of the infinite series of the world, where consequently there is no grounding reality behind the scene (e.g. the physical one or the mental one). As Landgrebe says explaining the structure of Husserl’s regional ontologies, «[i]n general, we cannot say whether an individual, allegedly perceived thing is a real thing, if we take it by itself. Its reality only shows itself when it is grasped in connection with other things [interpenetration], wherein it is affected by them and likewise affects them» (Landgrebe, 1981, p. 138). And even if there was such a grounding reality, we have seen how the ontological reduction of consciousness to the neuro-centric or mento-centric paradigm imply enormous problems. Therefore, as I already exposed in the previous chapter, I would prefer to look at the different dimensions of appearances, where the main aim is not to reduce all to one, but to connect various – and even

¹ «The common characteristics of the objects of a region, which are essential for their having become objects of that region in the first place, are contained in the categories of the region in question. These are the fundamental concepts of the region. They capture and express the a priori presuppositions under which a manifold of experienced entities can be apprehended as so belonging together as to be able to become the theme of a science. Because they constitute the specific ontological constitution [*Seinsverfassung*], i.e., the objective-character, of the objects of a science, Husserl called the philosophical disciplines which develop the categories of anyone of these regions regional ontologies. Thus a regional ontology provides the foundation for each empirical science, since it explicates the basic concepts of a science. Of course, this should not be taken to mean that such regional ontologies could be delineated prior to all empirical science. Rather, they become outlined on the basis of subsequent reflection on the conditions under which the delimitation of any province can and has come about» (Landgrebe, 1981, p. 134).

² «The reflection on the conditions under which entities are set off as a possible province of a science is a reflection on how they become objects for consciousness in such a way that such a demarcation is possible. Thus Being, in respect to the consciousness for which it becomes an object, is considered as being-an-object-for-consciousness [*Gegenstand-Sein für das Bewusstsein*]. The question is: how does an object come to givenness, and perhaps to self-givenness, to that consciousness of self-givenness which in general is designated intuition by Husserl! Thus the differentiation of regions of Being does not arise through a simple description of entities and their differences, but only through a description that considers the correlation between what is given and the manner in which it is given to consciousness. In this sense Husserl speaks of the constitution of an entity, meaning not simply its inner structure, but its way of appearing to the consciousness which is directed to it and has it as an object» (Landgrebe, 1981, p. 135).

opposite – explanations and descriptions, through a relational and interdisciplinary epistemology.

There are different regional ontologies we could select for a deep investigation of consciousness (physical, mental/phenomenal, psychological, etc.) which provides foundations for the sciences inasmuch as they determine what must belong to an object in order for it to become object of a given science (Landgrebe, 1981). Consequently, each region provides the background for different sub-sets of approaches, e.g. for the physical we have the neurological one, which in turn can be divided into several branches, neuropsychological, neurobiological, quantum biology, etc. Probably, if I had been born in the first part of the XX century, I would probably have been focused on the biology in relation to consciousness. Nowadays we know that the brain somehow is extremely relevant for the emergence and maintenance of our consciousness and we are starting to understand something about it. I think it is clear why I consider neuroscience just as problematic as unavoidable for a contemporary analysis of topics like experience, self, awareness. Therefore, in this last chapter I shall employ also findings from neuroscience, with the target of connecting it to other appearances of consciousness instead of reducing them to the neuronal.

The lived experience analysed by the phenomenological approach will be the other domain considered in this study. As explained in the fourth chapter, lived experience is not a regional ontology among others, but is their condition of possibility, and, at the same time, it is the pre-given world that needs to be investigated by phenomenological analysis. Broadly speaking, with phenomenology I mean that approach – rather than a strict method – which involves «the careful, unprejudiced description of conscious, lived experiences (Husserl's *Erlebnisse*), precisely according to the manner that they are experienced, without the imposition of external explanatory frameworks, whether these be drawn from the natural or social sciences, from religion, or even from common sense or ordinary language use» (Moran, 2018b, p. 75). So, it is not the passive acceptance of the experience as we conceive it through folk psychology or introspection, rather the disruption of the world as we see it every day (Merleau-Ponty, 1945/2013)³. The general aim is to determine the structure of any

³ «The aim of the phenomenological description cannot be simply to reproduce the original experience unaltered. [...] One aim of the phenomenological work is consequently to disclose, disentangle,

consciousness through which the world is manifest to it, and intentionality, phenomenality and self-awareness are parts of this structure. In particular, intentionality is the principal theme of phenomenology, as our consciousness primarily is an activity of «both meaning-apprehending and meaning-making, overall making sense of the world» (Moran, 2018b, p. 77). This “meaning of meaning” here used has nothing to do with the representational or higher-order theories that conceive consciousness as a process of post-phenomenal cognition or rational access to experience⁴. Every act of consciousness is in its own way sense-disclosing activity and it necessarily places us in a meaningful world⁵:

«Our embodied sensings, perceptions, feelings, moods, emotions, imaginings, thoughts, already have ‘sense’ prior to verbal articulation and may in fact be consciously experienced in an embodied manner. Our perceptual vista simply unfolds for us as we turn our eyes around to view the landscape; our hearing immediately experiences being in the midst of a noisy scene with car engines, slamming doors, human voices, barking dogs, or whatever. Even ‘white noise’ has a certain experiential significance – and is experienced as uncomfortable, distressing, and so on. All experiencing is experiencing of something and as something. We feel the day to be gloomy and we lack energy and enthusiasm. These are all forms of sense-disclosure» (Moran, 2018b, p. 77).

This interpretation of intentionality is more complex than how it is usually described, e.g. “to be conscious of something” simply as the directedness of a sense (sight, hearing, touch, etc.) towards a specific object, i.e. the standard model made by act, object and mode of presentation (Rowlands, 2013a) (just think about the arrow

explicate, and articulate those components and structures that are implicitly contained in the pre-reflective experience» (Zahavi, 2018, p. 5).

⁴ «[...] Husserl treated intentionality not narrowly as a relation between a mental act and its object, and not as a kind of representation of the outer world in the inner mind [...], but as expressing the fundamental ways embodied subjects, interacting with other subjects, apprehend the world as senseful and respond actively generating further commitments of sense. Intentionality is essentially a claim about the *sensefulness* of experience based on the irreducible, essential interrelatedness between signifying consciousness and the surrounding world of significance» (Moran, 2018b, p. 77).

⁵ «With the term “world” I do not mean exclusively the surrounding physical world made of material stimuli, which instead can be called environment, but I use world as the human reality: «‘World’ means the widest conceivable context for all human intending, not just the limits of the physical universe or ‘totality’ (Weltall), but also the limits of what is meaningful, the limits of temporality, possibility, and so on. Husserl recognises many forms of world – natural world (natürliche Welt), the world of things (Dingwelt), the spiritual world (die geistige Welt), the world of values and interests, the environing world (Umwelt), the familiar, home-world (Heimwelt) or ‘near-world’ (Nahwelt), alien-world (Fremdwelt), universal world (Allwelt), and so on» (Moran & Cohen, 2012, p. 342).

drawn from a stylized human being or head towards an external object, usually a tree or an apple). In effect, we never perceive a single object, one by one, at first as a neutral stimulus detached from everything else and only after as an object with a specific meaning, rather we perceive from the start a chair (not a geometric assembly of pieces of wood or metal) near a table inside a living room (i.e. where chairs are supposed to be)⁶. As Moran notes, our «experiences occur against a backdrop of previously meaningful experiences, in what Husserl calls the ‘harmoniousness’ (*Einstimmigkeit*) of our experience⁷, which in turn motivates certain expectations of continuity, how the future situation will pan out, and so on (Moran, 2018b, p. 6). Moreover, this kind of expansive intentionality does not imply any kind of representational relation between an external world and an inner, mental image of it. On the contrary, as Sartre sustained from the first works⁸, consciousness and the world are intrinsically connected from the start, “given at one stroke”, and none of the two can phagocyte the other. Metaphor aside, intentionality is not just about epistemology and the way in which we know the world, but restores consciousness’s direct contact with the world (Moran, 2018a), against the illusion of immanence and the bad faith of knowledge as “possession” of objects:

⁶ «In the mature Husserl, Heidegger, and Merleau-Ponty, there is a deeper focus on intentionality as pervading the whole of human existence, including the apprehension of backgrounds, contexts, “horizons,” and indeed the whole vague but genuine experience of “world” that outruns all possible experience [...] Object-intentionality is complemented by horizon-intentionality» (Moran, 2018a, p. 20).

⁷ «When we thus abstract. we retain a unitarily coherent stratum of the Phenomenon world, a stratum of the phenomenon that is the correlate of continuously harmonious, continuing world-experience. Despite our abstraction, we can go on continuously in our experiencing intuition, while remaining exclusively in the aforesaid stratum. This unitary stratum, furthermore, is distinguished by being essentially the founding stratum - that is to say: I obviously cannot have the “alien” or “other” as experience, and therefore cannot have the sense “Objective world” as an experiential sense, without having this stratum in actual experience; whereas the reverse is not the case» (Husserl, 1977, p. 96).

⁸ «You see this tree, to be sure. But you see it just where it is: at the side of the road, in the midst of the dust, alone and writhing in the heat, eight miles from the Mediterranean coast. It could not enter into your consciousness, for it is not of the same nature as consciousness. One is perhaps reminded of Bergson and the first chapter of Matter and Memory. But Husserl is not a realist: this tree on its bit of parched earth is not an absolute that would subsequently enter into communication with me. Consciousness and the world are given at one stroke: essentially external to consciousness, the world is nevertheless essentially relative to consciousness. [...] All at once consciousness is purified, it is clear as a strong wind. There is nothing in it but a movement of fleeing itself, a sliding beyond itself. If, impossible though it may be, you could enter “into” a consciousness, you would be seized by a whirlwind and thrown back outside, in the thick of the dust, near the tree, for consciousness has no “inside.” Precisely this being-beyond-itself, this absolute flight, this refusal to be a substance is what makes it be a consciousness» (Sartre, 1939, p. 130; translation by Joseph P. Fell).

«Our own being, says Heidegger, is being-in-the-world. One must understand this “being in” as movement. To be is to fly out into the world, to spring from the nothingness of the world and of consciousness in order suddenly to burst out as consciousness-in-the-world. When consciousness tries to recoup itself, to coincide with itself once and for all, closeted off all warm and cozy, it destroys itself. This necessity for consciousness to exist as consciousness of something other than itself is what Husserl calls “intentionality”» (Sartre, 1939, p. 131).

So, according to the phenomenological approach I am taking into account here, there are not two original and independent poles, the world and consciousness, which become connected in a second step, rather our existence is situated and structured from the start within a meaningful world that, however, preserves and develops its meanings through the sense-disclosing interaction of consciousness. This being-in-the-world (*In-der-Welt-sein*; Heidegger, 1927/1996) is the grounding condition of our existence as humans and this is the beginning and the endpoint of a never-ending phenomenological analysis.

5.1.2 The Relation between Phenomenology, Philosophy and Neuroscience

Now, after this brief examination of what I mean by “phenomenology”, how can the neuroscience of consciousness relate with such a different – in some circumstances, opposite – method of analysis of our experience? Neuroscience deals with correlations between neuronal and mental activity, input-output models, experiments about flashes, figures on the screen and sequences of sounds, investigations by brain-images, computational simulations and modelling from the single neuron to wide brain patterns related to cognitive functions, etc. How can this research program be connected with concepts like intentionality, being-in-the-world, self-awareness, etc., i.e. notions that intrinsically refer to something outside the context of the brain? Interestingly enough, neuroscience does not mean necessarily neurocentrism, indeed, despite having been born within the renewed crypto-Cartesianism, the neuroscientific methods can be incorporated into new paradigms as alternatives to reductive naturalism.

First, reductive naturalism is not the only possible version – although probably the most widespread – indeed there are also non-reductive or liberal naturalisms, which argue that natural sciences are fundamental but cannot exhaust the concept of nature (De Caro, 2014; De Caro & Macarthur, 2004, 2010). A non-reductive naturalism considers the methodology, epistemology and ontology offered by the naturalist paradigm necessary tools for the investigation of every natural phenomena, however also other entities, forms of understanding and methods are considered neither reducible nor incompatible to natural sciences⁹. In this sense, non-reductive or liberal naturalism is not meant as a non-scientific anti-naturalism. There are several works that propose different hypothesis of integration between the phenomenological analysis and the neuroscientific and cognitive researches, or, more generally, between phenomenology and naturalism (Brown, 2014; Carel & Meacham, 2013; Gallagher & Zahavi, 2012; Reynolds & Sebold, 2016; Daniel Schmicking & Gallagher, 2010; Winkler, 2017), oscillating with different degrees between the idea of “phenomenologizing” cognitive sciences and neuroscience, i.e. shedding light on neuroscientific research also through the first-person aspect of human experience (De Monticelli & De Vecchi, 2011; Gallese, 2011; Lutz & Thompson, 2003b, 2003a), and “naturalizing” phenomenology, i.e. translating phenomenology in the quantitative language of mathematics and with scientific protocols (Bockelman, Reinerman-Jones, & Gallagher, 2013; Dennet, 2003; Petitot et al., 1999).

Dan Zahavi has exposed four possible types of relation between phenomenology and empirical sciences in general (Zahavi, 2004). In the first case, empirical science does not need phenomenology as the latter is an obsolete method and therefore a waste of time for the contemporary lines of research. The second possibility is that empirical science instead must be constrained (and not simply

⁹ «Liberal naturalism liberalizes the tenets of strict naturalism, by accepting: (i) a *liberalized ontological tenet*, according to which there may be entities that are both irreducible to, and ontologically independent of, entities whose nature and behavior are not explainable by science but are not supernatural either; (ii) a *liberalized epistemological tenet*, according to which legitimate forms of understanding (such as conceptual analysis, imaginative speculation or introspection) exist that are neither reducible to scientific understanding nor incompatible with it; (iii) a *liberalized semantic tenet* according to which there are truth-apt factual judgments that do not concern scientifically accepted entities or properties and are irreducible to judgments regarding such entities or properties; (iv) a *liberalized metaphilosophical tenet*, according to which there are issues in dealing with which philosophy is not continuous with science as to its content, method and purpose» (De Caro, 2014, pp. 73–74).

helped) by the phenomenological method (and not vice versa). The third case distinguishes between a pure form of phenomenology – the transcendental one – which remains untouched, and a “mundane” form – psychological phenomenology – which can establish a relationship of mutual constraints with empirical science. The last option speaks of a possible fruitful exchange between transcendental phenomenology and empirical science, which could lead both to a modification of the transcendental phenomenology informed by empirical findings and to the transformation of the concept of nature and naturalization as core concepts of contemporary empirical sciences¹⁰. Since one of the main points here is to connect neuroscience and phenomenology, I shall present a specific attempt to connect the two in a way that may resemble the third or the fourth options presented by Zahavi, and this interdisciplinary movement is called neurophenomenology.

5.1.2.1 Neurophenomenology

As the word itself says, neurophenomenology tries to connect the most recent lines of research in neuroscience with classic and contemporary developments of the phenomenological analysis. The neurophenomenological approach puts into brackets all the epistemological and ontological assumptions of the naturalistic paradigm (e.g. all the oppositions between mind and world, subject and object, etc.), and considers phenomenology and neuroscience as related and dependent one each other. Moreover, neurophenomenology acknowledges that consciousness and cognition are intrinsically embodied and embedded within the world with which they constantly interact (Gallagher & Zahavi, 2012; Rupert, 2015; F. J. Varela, Thompson, & Rosch, 1991). The aim of neurophenomenology is to rethink the problems of experience in a new methodological and pragmatic way, with no metaphysical commitments or scientific presupposition about it (F. Varela, 1996b). According to this approach, for example, getting stuck in the Cartesian mind-body separation will not be helpful for neither

¹⁰ «In other words, why let the reductionists monopolize the concept of naturalism? Most revealing of all, however, is perhaps a reply given by Varela to a question that I posed to him at a meeting in Paris in 2000: The volume *Naturalizing Phenomenology* was only intended as the first part of a larger project. The second complementary volume was planned to carry the title *Phenomenologizing Natural Science*» (Zahavi, 2004, pp. 343–344).

neuroscience nor philosophy of mind, instead, starting from the original unity of experience in the world, we can find a bridge between different methods and types of experience, such as the neuroscientific and the phenomenological ones, looking for instances of practical “collaboration” and “mutual enlightenment” (Gallagher, 2012).

Of course, the core issue of neurophenomenology is how to connect the two methods together, since they come from two extremely different approaches towards consciousness. Basically, the neurophenomenological idea is to phenomenologically train the experimental subjects in order to obtain a more detailed description of their own experiences, and without imposing pre-determined theoretical or scientific categories¹¹. There are other approaches that can be interpreted as slightly different proposal to neurophenomenology or complementary methods to it. For example, when it is not possible to train the experimental subjects, the so-called “front-loaded phenomenology” has been conceived for focusing on the design of experiments, so that the experiment will be set up on the basis of the insights developed in phenomenological analyses (Gallagher, 2003). In this way, it is possible to project new experiments or to re-think old experiments which have been structured on the basis of common sense or folk psychological concepts¹².

Unfortunately, the present research – in particular the forthcoming arguments – did not have the opportunity neither to train subjects on some specific task about consciousness, nor to design a proper experiment supported by phenomenological insights. This does not allow us to define this work as neurophenomenological *stricto*

¹¹ «[...] train the experimental subjects to gain greater intimacy with their own experiences. The subjects are subsequently asked to provide description of these experiences using an open-question format, and thus without the imposition of pre-determined theoretical categories. The ensuing descriptive categories are subsequently validated intersubjectively and then used to interpret correlated measurements of behavior and brain activity. At the same time, however, it is also suggested that, say, a consideration of insights from neurobiology and dynamical systems theory can help us improve and refine the classical phenomenological analyses» (Zahavi, 2010, p. 9)

¹² «Experimental design is always based on some concept or distinction. Most often such ideas come from previous experiments. If one traces these ideas back far enough, however, one finds that some of the concepts and distinctions are based on previously operationalized concepts that are themselves drawn from folk psychology, our everyday natural attitude, or varying philosophical traditions. In this regard there may be less control involved in experimental design than the experimenters believe; specific theories lurking in the background may already introduce certain biases that are shaping the kind of conclusions that can be drawn from the experiments. To the extent that the concepts and distinctions incorporated into the experiment are generated in careful phenomenological analyses, closer to experience and less influenced by established theoretical considerations, the scientist gains a certain degree of control that is otherwise missing» (Gallagher, 2012).

sensu, nevertheless it does not mean that there is no other way to understand the relationship between neuroscience and phenomenology – or philosophy more in general. One of the aims of the next parts of the chapter is to insert the topic of consciousness within a non-reductive context – ontological, epistemological and methodological – where the empirical results are interpreted in the light of a conceptual and phenomenological analysis and where at the same time the phenomenological descriptions are informed by data obtained from neuroscience, as well as philosophical *a priori* concepts are challenged by empirical evidences. This implies that the kind of phenomenology used here cannot confine itself to the realm of pure transcendental, as well as the neuroscientific data that will be illustrated is not to be interpreted according to the neuro-centric paradigm¹³. On the contrary, I shall try to understand consciousness through the fruitful interplay (Ratcliffe, 2006) that can emerge from the collaboration between phenomenology, philosophy and neuroscience without denying their differences, according to the fourth option exposed by Zahavi¹⁴.

In the present work, phenomenology, philosophy and neuroscience are supposed to perform different roles since the respective peculiarities must be respected and preserved, in order to avoid any kind of reductionism, e.g. eliminative materialism, idealism, phenomenalism, etc. Therefore, the real challenge here is to build connections between different methods and ontologies. These disciplines have not only different methodological approaches, of course, but they answer different guiding questions. For example, while the neuroscience of consciousness is looking for *what*

¹³ «This is in line with the Merleau-Ponty's phenomenological project, neither anti-naturalist nor naturalized, but strictly connected and informed by natural sciences: By envisaging a dialogue between phenomenology and natural science (and not merely between phenomenology and the humanities/social sciences) Merleau-Ponty goes a step further than Husserl. What is interesting and important, however, is that Merleau-Ponty didn't conceive of the relation between transcendental phenomenology and positive science as a question of how to apply already established phenomenological insights on empirical issues. It wasn't simply a question of how phenomenology might constrain positive science. On the contrary, Merleau-Ponty's idea is that phenomenology itself can be changed and modified through its dialogue with the empirical disciplines. In fact, it needs this confrontation if it is to develop in the right way. And mind you, Merleau-Ponty holds this view without thereby reducing phenomenology to merely yet another positive science, without thereby dismissing its transcendental philosophical nature» (Zahavi, 2004, p. 342).

¹⁴ «The final option is to argue that transcendental phenomenology itself can enter into a fruitful exchange with empirical science. Perhaps it can even be naturalized in the sense of "contributing to the project of naturalization." But of course, if this view is advocated it has to be realized that such a naturalization will not only lead to a modification of transcendental phenomenology, at the same time it will also transform the very concept of naturalization as well as our very understanding of nature» (Zahavi, 2004, p. 343).

happens (causes) in the brain in relation to when we experience something, and *what* the function of this neuronal, psychic and behavioural activity is, phenomenology instead looks for how our experience as such is constituted, i.e. the disclosing activity through which the world is manifested, in order to identify the common and intersubjective structures of our way of being aware of the world. Hence, as briefly pointed out in the previous chapter, it may be possible to connect empirical explanations (about the causes in virtue of which a phenomenon arises) with constitutive explanations (about the conditions under which something can be what it is). In the second chapter, for instance, I tried to expose the logical condition that makes the term “consciousness” what it is, and the provisional phenomenological conditions of what being conscious of something means, i.e. intentionality, self-awareness and phenomenality. Now these conditions must be plumbed up with empirical evidences, a deep phenomenological description and a critical conceptual analysis. For all these reasons, rather than calling this approach “neurophenomenological” in the strict sense, I would call it rather “neuropsychological” in a broader sense.

5.1.2.2 Non-reductive Neuropsychology

As already shown in the previous chapters, I shall make use of historical-conceptual analysis, thought experiments, phenomenological descriptions, and now also empirical evidences. I consider neuropsychology as an umbrella term within which we can find different approaches to consciousness and cognition. Generally, the term refers to the reductionist neuropsychology introduced by Patricia Churchland at the end of the 1980s (P. S. Churchland, 1986) and may include also other authors who promote a reductionist or eliminativist approach to the mental (Bickle, 2003; P. Churchland, 2007). In other words, reductionist neuropsychology is an instance of drastic approach towards consciousness and cognition, perfectly in line with the neuro-centric paradigm. So, if we want to use empirical data – as neuroscientific findings – without fall back into the categories of brainhood, physicalism, and eliminativism, the reductionist neuropsychology *à la* Churchlands completely misses the point.

While it is easy to conceive the connection between phenomenology other philosophical methods (thought experiment, historical-conceptual and logical analysis, etc.), it is not so for the relationship between philosophy and neuroscience. Surely, in the case in which it is inclined towards a reductionist and eliminativist neurophilosophy, no problem arises given that philosophy would be incorporated and replaced by the neuroscientific approach. Nevertheless, in the present work we are looking for a non-reductionist perspective, where the relationship between different ontological, epistemological and methodological domains is its strong point. Therefore, we may want to look towards a non-reductive neurophilosophy similar to Northoff's approach, which aims to a fruitful and reciprocal interaction between philosophical and neuroscientific methods in order to achieve at the same time logical coherence and empirical plausibility (Northoff, 2004a, 2014b). This interdisciplinary approach needs a pluralist epistemological basis which is able to strongly connect the two domains; in this sense, a mere juxtaposition of new neuroscientific findings next to philosophical theses is not sufficient to justify any classic theories about consciousness (e.g. "internalism about the mind is true because we have found correlation between mental activity and the brain"). At the same time, philosophy cannot passively approve data without challenging their empirical consistency and the scientific conditions of possibility (e.g. "the NCCs are the proof that consciousness arises from the brain"). He distinguishes "replacement or incorporation naturalism" (reductive naturalism) from "cooperative naturalism" (Northoff, 2014b, pp. 62–64); the former is typical of a reductive and monopolar methodology, while the latter allows for the development of different (philosophical and scientific) methods, towards a reciprocal complementarity. For this reason, it is quite inappropriate to place the author's perspective as a polar opposite of reductivism; in fact, non-reductivism does not entail an anti-naturalist view, according to which philosophy should have complete autonomy from neuroscience. Rather, Northoff tries to stand in the middle, putting forward a neurophilosophical approach that is «[...] non-reductive primarily in the methodological (rather than metaphysical) sense» (Northoff, 2014b, p. 1).

As opposed to mind-based and brain-reductive approaches, non-reductive neurophilosophy involves a brain-based approach, moving from the essential question: "How do the brain and its neural features relate to the mind?" The reversal of the

typical question “How do the mind and its mental features relate to the brain?” produces some new perspectives about the mind-brain relationship, and there is no need to find a neural correlate for any single mental property. It is rather possible to avoid this epistemological bottleneck by integrating the brain within a philosophical context, as well as by finding new interdisciplinary approaches, while dismissing at the same time the assumption of any kind of supernatural property. Thus, the intention is to account for a study of the brain as such, not as an empirical substitute of the mind; any possible relation with the concept of mind will come after testing the approach, not as an a priori metaphysical assumption.

Also, in line with what has been pointed out in the previous chapter, different domains of research do not lead to mutual exclusions, rather philosophy and neuroscience could be considered as «different degrees of abstraction on the same underlying continuum» (p. 52). So, instead of determining which ontological or epistemological level comes first, the non-reductive neurophilosophical approach is based on a framework called “concept-fact iterativity” (Northoff, 2011, 2014b), made by transdisciplinary and continuous loops between theoretical and experimental domains. This method consists of philosophical and neuroscientific passages where the historical-philosophical concept can be the guiding light for an hypothesis generation and then it gets through a process of adaptation, operationalization (see chapter 3) and experimentation, then going back through a neurophilosophical interpretation of the neuroscientific data, by feedbacks from the empirical, correction loops and interactions with the conceptual background¹⁵. The aim then is to define what Northoff calls “neurophenomenal hypothesis” that show how the brain’s intrinsic features predispose – but are not identical to – the intrinsic features of consciousness and its phenomenal features in a necessary way¹⁶.

¹⁵ In order to provide a rigorous structure to the concept-fact iterativity, Northoff exposes some methodological principles: «The ‘principle of asymmetry’ points to an asymmetric relationship between logical and natural conditions. The ‘principle of bi-directionality’ claims for the necessity of bi-directional linkage between natural and logical conditions. The ‘principle of transdisciplinary circularity’ describes systematic rules for mutual comparison and cross-conditional exchange between philosophical theory and neuroscientific hypotheses» (Northoff, 2004b, p. 91).

¹⁶ «[...] the concept of “neurophenomenal hypotheses” describes suggestions for how particular neuronal mechanisms of the brain are related to specific phenomenal features of consciousness. The aim of this book is to develop specific neurophenomenal hypotheses for the different phenomenal features of consciousness, like qualia, first-person perspective, intentionality, unity, and so on» (Northoff, 2014c, p. xxv)

5.1.2.3 Neuropragmatism and 4E Cognition

Other authors have recently developed other neuro-philosophical positions, whose goal is to understand the mental life through the connection different domains without reducing everything at the neurological level. For example, neurophilosophical pragmatism or neuropragmatism (Solymosi & Shook, 2013, 2014) is a broad field of research which claims that animals are goal-oriented organisms whose cognition in general (i.e. consciousness included) is not explicable apart from the bodily context. Moreover, inspired by Dewey's pragmatism but also open to other theories¹⁷, neuropragmatism argues that cognition can be understood as a dynamic and adaptive system of interactions between the organism and the environment (Solymosi, 2013), i.e. in order to survive and improve its living conditions, the organism's neural connections have changed on the basis of the challenges posed by the environment (Malafouris, 2013). This is an example of how neuroscientific data can be seen in a different light, interpreting them as a function of external causes and needs, and not as manifestations of the brain's activity as such. Here the brain is conceived intrinsically embodied and its activity no longer makes sense if we look at it without considering bodily actions and the environmental context¹⁸. A fortiori, consciousness and cognition according to neuropragmatism cannot be conceived as an internal process of the brain, against both the mind-based (how the mind relates to the brain) and brain-reductive (how the brain relates to the mind). The question, instead, can be formulated kind of like this: how does the brain in the body interact with the environment and how does the environment modify the human organism? Interestingly, neuropragmatism – as pragmatism did before – tries to get through the paradigm of neuro-centrism and its hidden crypto-Cartesianism, proposing instead an ecological approach to cognition

¹⁷ «Neuropragmatism allies easily with theories of active perception; somaesthetics; naturalizing intention; ecological psychology; ecological cybernetics; social cognition and social epistemology; neurosociology; extended mind; neurophenomenology; and radical embodied cognitive science. Even aspects of connectionism and dynamic systems theory may contribute to the proper synthesis of these positions if excessive representationalism is avoided» (Solymosi & Shook, 2014, p. 16).

¹⁸ «Neuropragmatism offers a philosophical intersection for coordinating this pluralistic effort. The prefix 'neuro' does not portend a reductionistic agenda is intended by the term neuropragmatism. Quite the opposite: the anti-reductionistic, pluralistic, and interdisciplinary tradition of pragmatism remains securely at the heart of neuropragmatism. All the same, a philosophical position on cognition and mind must cohere with the best neuroscience available» (Solymosi & Shook, 2013, p. 2).

and consciousness that uses non-reductive concepts like “resonance”, “attunement”, “anticipation” (van Dijk & Myin, 2019), where organism and environment are not connected by simple input or output relations, but rather by dynamic, circular relations (Charles, Golonka, & Wilson, 2014).

So, neuropragmatism is based on an anti-Cartesian paradigm (Solymosi, 2011) that focuses on the interaction between brain, body and the environment. This is quite the same paradigm which underpins another philosophical and scientific position, the 4E cognition movement. 4E cognition (embodied, embedded, enactive and extended) is an field of interdisciplinary research (philosophy, neuroscience, psychology, social sciences, etc.) based on the idea that cognition is structured by dynamic interactions between the brain, body, environments (in both physical and social sense) (A Newen, De Bruin, & Gallagher, 2018). It brings together a series of approaches against neuro-centric approaches like computational and cognitive theories of mind (Gallagher, 2018). Against classical cognitivism, the 4E theories argues that the mind is not a passive box receiving neutral physical stimuli from the environment through the body (Menary, 2010a). Rather, the mind seems intrinsically linked to bodily actions towards the environmental context (embedded, enactive) (Di Paolo & Thompson, 2014), moreover cognitive processes are not bound by the skull or even skin, but actively incorporate environmental structures such as symbols, tools, artefacts, media, cultural practices, norms, groups, or institutions (Theiner, 2017; Wilson, 2017). The body – and not just the brain – processes information, in a sensory-motor way (embodied). Representations can be action-oriented and body-skill-relative and the body itself plays a representational role (enactive) (Noë, 2004, 2009; O'Regan & Noë, 2001). According to some radical theories, there are no mental representations at all in a classical sense, rather, sensory-motor contingencies, bodily affects, postures and movements enter into cognition in a non-representational way (Hutto & Myin, 2012, 2013, 2017). The 4E movement cannot be conceived (only) as a neurophilosophical approach in a narrow sense, since it involves different disciplines and it is made by several theories with different methodologies, nevertheless the main point is that the body situated within the environmental context is central for the emergence of cognitive processes and for our understanding of what the mind is.

As well as neuropragmatism, the 4E approach goes beyond the limits of neurocentrism and, in turn, does not fall into the parallel paradigm of mento-centrism. Indeed, the mind here is not an extra-natural property, nor a psychic entity that arises from the body, instead both neuropragmatism and 4E cognition understand the mental as a dynamic process composed by body-environment interactions and loops¹⁹. In this sense, these anti-Cartesian positions extend the issue of consciousness (and cognition in general) beyond the brain and the head, including the body and its relations with the environment. Tibor Solymosi exposed three metaphors about consciousness that can be used for distinguishing between eliminativism/reductionism, 4E cognition and neuropragmatism: consciousness-as-digestion, consciousness-as-dancing, consciousness-as-cooking (Solymosi, 2011). The metaphor of digestion is related to reductive neurophilosophy, which conceives the mind as identical to the brain or as an activity of the brain (functionalist version) and nothing more, as digestion is a particular activity of the gut and nothing more²⁰. As already seen in the previous chapter, this metaphor only emphasizes the brain and anything about the body or the environment is mediated and filtered through it. The metaphor of dance (related to 4E cognition) has been proposed first by Alva Noë²¹, and it indicates that consciousness is not something that happens to our nervous tissues but something that we do, with and through our body that is incorporated within an environment, as the act of dance

¹⁹ «The Cartesian claim that mind and body have entirely different properties is demonstrably false. Lingering claims that consciousness has unnatural properties similarly rest on philosophical confusions and ignorance of brain science. Mental activity, conscious and unconscious, is a natural process involving the nervous system – as such it is entirely open to scientific inquiry. [...] When thus seen, we are better speaking not of mind as a noun but of mind as a verb: an organism does not have a mind, rather an organism minds. Indeed, our scientific activity should not be inquiring into the mind but into the process of minding. Mentation goes beyond the cranium, suspended in a cultural medium of communicating humans» (Solymosi & Shook, 2013, pp. 8, 15).

²⁰ «[...] the mind is the brain (at least until some technical concern is raised that leads to an onslaught of qualifications). It is either a matter of strict identity or a matter of correlation, as exemplified by the search for the neural correlates of consciousness. Other times, the emphasis is not so much on correlation between mind and brain as it is a functional relation. The mind is what the brain does. The mind is not some sort of substantive thing, but an activity performed by the brain or central nervous system. The popular analog is to digestion and the gut. Digestion is the process performed by a properly functioning GI tract. Mentation, likewise, is the process performed by the brain (or central nervous system)» (Solymosi, 2011, p. 350).

²¹ «In a way our problem is that we have been looking for consciousness where it isn't. We should look for it where it is. Consciousness is not something that happens inside us. It is something we do or make. Better: it is something we achieve. Consciousness is more like dancing than it is like digestion. [...] Human experience is a dance that unfolds in the world and with others. You are not your brain. We are not locked up in a prison of our own ideas and sensations. The phenomenon of consciousness, like that of life itself, is a world-involving dynamic process. We are already at home in the environment. We are out of our heads. [...] You can no more explain mind in terms of the cell than you can explain dance in terms of the muscle» (Noë, 2009, pp. xii, xiii, 48).

is something we actively do with our body in relation with the dancefloor and the other people dancing with us. Since the dancing metaphor seems to be «more about the body than the brain or the environment» (Solymosi, 2011, p. 360), Solymosi proposes, in relation with neuropragmatism, the metaphor of cooking, which should unify the two previous metaphors. Indeed, cooking is something that we actively do (as motor coordination when we dance), during which we select ingredients and prepare them for a better nutrition (as digestion but extended, including the fact that we do things for reasons instead of passively receiving things), also there is no cooking without a kitchen (relevance of the environment)²².

It is now clear that both neuropragmatism and 4E movement are instances of how even neuroscience can be used against the neuro-centric view. The point here is not to neglect the neuroscientific findings about the brain in relation to our conscious experience, but to relocate them within a wide-ranging interpretation, including the whole body and the environment. Moreover, they are not two competitive approaches, rather they can collaborate with each other, contrary to reductionism that attempts to unify everything under one method and scientific domain. However, these approaches, despite extending the issue of conscious experience through the body and the environment, risks to remain stay within the confines of physicalism and reductive naturalism. While I agree about the necessity of extending the context of the mental all over the body and environment without producing a new hypostatization of the mind, this does not mean that we should reduce the analysis of experience to biological, evolutionary and instrumental relations between brain, body and environment. Of course, this ontology is way more elaborated than the body-machine

²² «In cooking our food, we become more selective of ingredients and means of preparation that yield greater nutrition. Moreover, cooking is not something that simply happens to us in the way that the GI tract's activity just happens. Cooking is something we do. Consciousness (or mentation generally) is no different. [...] Digestion is just intradermal. Dancing is just a body in coordinated motion. Cooking, however, cuts across these differences and unifies them. Cooking is digestion extended, requires coordinated bodily activity, and modifies the environment. Cooking adjusts organisms more robustly than digestion or dancing alone. For this metaphor of consciousness and mind to be effective, however, an evolutionary perspective must be diligently held. [...] Cooking, unlike digestion, involves doing things for reasons (known to the cooks) as opposed to things just happening (for no apparent reason). In this shift from physiological adaptation, eventually leading to evolutionary adaptation, to the deliberate alteration of the environment to resolve problems surrounding digestion and nutrition, we witness not only the shift from digestion to cooking but also the transaction from mere awareness of the here and-now to the self-awareness characteristic of humans. [...] From the kitchen appliances and utensils to the plethora of ingredients, there is no cook without a kitchen: no mentation without a situation of organic-environmental transaction» (Solymosi, 2011, pp. 357, 360, 363, 364).

one, since the organism is conceived in a bio-psycho-social way, but, according to these views, once we find the causal relations between the organism and the environment, we may possess a complete description of consciousness, and this perspective is still related to a strong reductive physicalism. With regards to the 4E approach, the investigation of the way in which we experience the world through embodiment, interaction with and reactions to the environment, and extension of cognitive skills through body and instruments can potentially shed light on how intentionality is structured; however, without a deep analysis of the phenomenology of the body, I do not see how the 4E movement could deal with issues like self-awareness and phenomenality. Indeed, there are much more 4E works about cognition, action, perception, emotions, etc. rather than works about consciousness in narrow sense (see Newen et al., 2018).

Moreover, let us consider the issue of subjectivity: according to neuropragmatism, subjectivity is not reducible to brain activity, rather it is a particular kind of interaction between the organism and the environment²³. I completely accept this point; however, this is where other new (old) problems arise. First, it is necessary to understand how the interactions between organism and environment are accompanied by something it is like to be in that interaction, which is a question really near to the hard problem of consciousness. Secondly, what differentiates the interaction between a rolling stone and the slope on which it is located and the interaction between our subjectivity as organisms and the surrounding environment? Someone may argue that the movement of the stone is not goal-directed while our actions are, this indeed is one of the features that differentiates an organism from an inanimate thing. But then, what differentiates the interaction between human organism and environment from the interaction between the tree and the soil in which it is

²³ «No pragmatism would seek to ‘reduce’ felt qualia to nervous activity or anything else to prove that they are natural. [...] Qualitative feelings happen where nervous systems achieve certain degrees of complexity in their transaction with their respective bodies. Subjectivity need not be treated as anything spookily “unnatural.” The mysteriousness of subjectivity quite vanishes. Subjectivity and perspective are precisely what would be naturally expected when specific brains generate specific experiences. You have a very different perspective from anyone else, because you are directly experiencing through your unique nexus of your brain, body, and world, and not from mine or any other’s nexus. [...] The subject-matter is experience, conceived as the dynamic interaction of organism and environment: “Things interacting in certain ways are experience” (Dewey 1925/LW1: 12); experience is “the manifestation of the interaction of organism and environment” or simply “an interaction of organism and environment” (Dewey 1939, 531)» (Solymosi & Shook, 2013, pp. 8–9, 16).

planted? The tree produces goal-directed actions, such as the lengthening of the roots in search of water. If so, the pragmatic concept of experience would be very broad and would fall outside the scope of the following research concerning human experience. Some might argue that the water-as-a-thirst-quencher for humans takes on a different level of meaning than the water sought by the plant, which is not likely to experience it or “think” of it as thirst quencher. However, the organism-environment processes prepared by neuropragmatism (see the cooking metaphor) are based on a merely evolutionary ontology, i.e. grounded on concepts like, the necessity to survive, adapt, improve the living situation, construct efficient habits, fitness, etc., and there seems to be no room for the existential meaning of experience and the world. As I already pointed out and I shall demonstrate later, an exhaustive account of experience cannot lack an existential analysis of human consciousness as process of sense-making that deals not only with a world of tools with our feeling something it is like to be-in-the-world, i.e. to exist as a subject within a intersubjective world of given meanings. Therefore, neuropragmatism may be complemented and informed by a phenomenological analysis able to look at human existence.

To sum up, starting from the need for an alternative approach to consciousness that was characterized by interdisciplinarity and ontological pluralism, I discussed the relationship between phenomenology, philosophy and neuroscience. I have shown how the intersection of different domains of knowledge and methods is important to overcome a merely neuro-centric or mento-centric view of consciousness. I have exposed some of the main alternatives, such as neurophenomenology, the non-reductive neurophilosophical approach, neuropragmatism and the 4E movement. In the next part of the chapter I will expose a potential neurophilosophical account of consciousness, that will make use of philosophical-conceptual analysis, phenomenological analysis and various neuroscientific studies. For this reason, I consider it is difficult to place the present work within a specific category among those shown above. As will be noted later, the method will be similar to Northoff's iterativity between neuroscience and philosophy, from the ontological point of view I will take the pluralism of neuropragmatism and 4E cognition, while regarding the investigation of the experience as such I will rely on phenomenological analyses and neurophenomenological studies.

5.1.2.4 Philosophy of Neuroscience and Neuroscience of Philosophy

By referring to the famous distinction in neuroethics between ethics of neuroscience and neuroscience of ethics (Roskies, 2002, 2016), I shall propose the same subdivision in neurophilosophy, between philosophy of neuroscience and neuroscience of philosophy. Usually, in particular in reductive neurophilosophy, neuroscience of philosophy is the only one domain taken into consideration, with the main goal of translating the philosophical concepts in neuroscientific language, while using (not always) philosophy of neuroscience only as a supplementary tool for secondary issues. Thus, by neuroscience of philosophy I mean the inclusion of neuroscientific results into the development of philosophical theories. The data collected from neuroscience may be useful to strengthen and corroborate conceptual and analytical theses, although not at the level of logical justification. As seen in the first chapters, these empirical results regarding the neuronal mechanisms of the brain may develop numerous new questions on concepts typical of the Western culture, e.g. in this case consciousness, but also personality, identity, mental representation, morality, freedom, etc.; when these concepts are questioned by neuroscientific data, philosophy is called to consider the empirical domain for a comparison with the conceptual domain that should not produce neither the uncritical acceptance of the results, nor their complete rejection.

Indeed, it is hazardous to apply neuroscientific findings upon philosophical issues without constantly referring to methodological, epistemological and conceptual investigations of philosophy of neuroscience. The way in which neuroscientific findings can complete, test or develop philosophical theses can be determined by a well-formed epistemological bridge between the two disciplines, and this is one of the tasks of philosophy of neuroscience. By philosophy of neuroscience I mean a specific branch of philosophy of science, as the critical examination of the epistemological and methodological conditions on which neuroscience and cognitive sciences are based, e.g. regarding explanatory models and methods of investigation (see for example chapter three). As said, it is not enough to merely accept the neuroscientific findings and give them a philosophical interpretation, as if they were pure, pre-conceptual, presuppositionless data. Rather, as seen in Northoff's iterativity process or in the

neurophenomenological method, concepts, phenomenal experience and neuroscientific findings are intrinsically bounded together during research by mutual constraints, feedback loops, etc. When neuroscience of philosophy is supported and completed by philosophy of neuroscience, then neurophilosophy can proceed in a non-reductive, interdisciplinary, and pluralist way. At this point, therefore, it is also possible to conceive proper methods and epistemological foundations for research that are neither neuroscientific nor philosophical, but specifically neurophilosophical²⁴. In the next part of this chapter and in the following one, I will expose a neurophilosophical account of consciousness of this kind.

5.2 Conceptual and Phenomenological analysis: Non-mentalistic and Non-physicalist account of Consciousness

So far, I have argued for the necessity of an analysis of consciousness that could involve different regional ontologies and several methodologies. Now, I shall come back to the structures of consciousness delineated in the second chapter: intentionality, phenomenality and self-awareness. As I tried to explain, an account of consciousness based on a neuro-centric approach presents many issues, but also the mento-centric approach is quite perplexing. How can we organise an analysis of consciousness without hypostatizing it in a property-based or content-based way? I believe that, again, Sartre could be helpful.

5.2.1 Intentionality with No Contents: Anti-idealism and Phenomenological Realism

I already pointed out the way in which intentionality is conceived by Sartre, as the main modality through which we find ourselves immersed in a world from the start. If

²⁴ «Philosophy of neurophilosophy is about methodological, theoretical and foundational issues in neurophilosophy. These issues concern, for instance, different kinds of possible methodological approaches in neurophilosophy, types of neurophilosophical explanations, and the relation between neurophilosophy and philosophy» (Northoff, 2014b, p. 316).

we want to use a metaphor, instead of imagining intentionality as a grasping-and-absorbing activity that the subject performs towards objects, it is more like an explosion, bursting towards the world, towards the objects (Sartre, 1939). There is nothing inside consciousness, since intentionality as “being conscious of something” is not a process of cataloguing objects inside the mind but openness to and encounter with the world²⁵. Against internalism, any idea or image is overproduced by the reification of consciousness, so that when we say “I am conscious of this chair” we already have objectified the experience of “there is a consciousness of this chair” (Sartre, 2004b). So, there are not “things” that are both in the world and in the mind, and where consciousness is supposed to link these two “realms” or to be one of the two: «The first procedure of a philosophy ought to be to expel things from consciousness and to reestablish its true connection with the world, to know that consciousness is a positional consciousness of the world» (Sartre, 1956, p. li). Thus, there is nothing in consciousness, and everything is outside of it, but everything is at the same time an appearance of consciousness – not in a mentalistic sense opposed to the beings-in-themselves, as we saw in chapter 4, but as objects of awareness²⁶. Even imagined things are outside consciousness, in such a way that if I imagine or remember a table, this becomes immediately an object of the particular act of consciousness (imagination or memory), i.e. a transcendent object which is exterior to the act of consciousness as such, in this case a table that appears to me as imagined or remembered (Sartre, 2004a). As we have already seen in the previous chapter, the

²⁵ «All at once consciousness is purified, it is clear as a strong wind. There is nothing in it but a movement of fleeing itself, a sliding beyond itself. If, impossible though it may be, you could enter “into” a consciousness, you would be seized by a whirlwind and thrown back outside, in the thick of the dust, near the tree, for consciousness has no “inside.” Precisely this being-beyond-itself, this absolute flight, this refusal to be a substance is what makes it be a consciousness» (Sartre, 1939, p. 130).

²⁶ «The transcendental field, purified of all egological structure, recovers its former limpidity. In one sense, it is a nothing, since all physical, psycho-physical and psychical objects, all truths, and all values are outside it, since the me has, for its part, ceased to be part of it. But this nothing is everything because it is the consciousness of all these objects» (Sartre, 2004b, p. 25). «All consciousness, as Husserl has shown, is consciousness of something. This means that there is no consciousness which is not a positing of a transcendent object, or if you prefer, that consciousness has no “content.” We must renounce those neutral “givens” which, according to the system of reference chosen, find their place either “in the world” or “in the psyche.” A table is not in consciousness, not even in the capacity of a representation. A table is in space, beside the window, etc. The existence of the table in fact is a center of opacity for consciousness; it would require an infinite process to inventory the total contents of a thing. To introduce this opacity into consciousness would be to refer to infinity the inventory which it can make of itself, to make consciousness a thing, and to deny the cogito. The first procedure of a philosophy ought to be to expel things from consciousness, to know that consciousness is a positional consciousness of the world» (Sartre, 1956, p. li).

search for the location of the objects of consciousness as if they were somewhere inside the head or the mind is based on unfounded assumptions (illusion of immanence; Sartre, 2004a).

This idea of consciousness devoid of objects as an activity of pure directedness towards the world has been interpreted by Rowlands as the Intentionality Thesis, i.e. all consciousness is intentional²⁷, and the No content Thesis, i.e. necessarily, any object of consciousness is outside consciousness, which is structured by conscious acts, e.g. thinking, imagining, remembering, perceiving, etc. (Rowlands, 2013b, 2016). This opens to an anti-substantial and anti-idealist account of consciousness, where consciousness is neither a sort of non-physical thing/psychic object/qualitative datum nor the foundation of the things in the world (idealism). Consciousness (*pour-soi*: being-for-itself) is intrinsically the disclosing activity (intentionality) to something that is transcendent to it, i.e. something which is not consciousness (*en-soi*: being-in-itself), and this means that consciousness as intentionality necessarily needs to be supported by a being outside of it (Sartre, 1956, p. lxi)²⁸. In other words, there is no consciousness without a world, understood as a collection of potential or actual objects for consciousness, which instead is a mode of existence, i.e. pure activity within the world ("hole of being"; Sartre, 1956, p. 79) on which it depends, while the world does not similarly depends on consciousness²⁹.

In this sense, Sartre can be conceived as bearing either an anti-idealist and externalist position on consciousness (Rowlands, 2003b) and a realist position on the

²⁷ Someone might argue that pre-reflective non-positional consciousness does not involve intentionality as being conscious of something, however, as we have seen in the second chapter, «Sartre does not think of non-positional consciousness as a separate form of conscious act but, rather, as something that is necessarily built into any positional (i.e. intentional) conscious act» (Rowlands, 2013b, p. 530n).

²⁸ «In Sartre's terminology: being for-itself requires and presupposes being-in-itself. Being for-itself has no content. We might think of it as a hole in being-in-itself. A hole cannot exist without its edges, but these edges are not part of the hole. The existence of a hole guarantees the existence of something outside it. In this sense, consciousness (the *pour-soi*) provides an ontological proof of being-in-itself (the *en-soi*). The converse dependence does not hold. Being-in-itself is complete and self-contained. NCT, therefore, in addition to entailing the basic distinction between being-in-itself and being-for-itself also entails that idealism is false: "To say that consciousness is consciousness of something is to say that it must produce itself as a revealed-revelation of a being which is not it and which gives itself as already existing when consciousness reveals it." [Sartre, 1956, p. lxi] For this reason, Sartre characterizes his position as a "radical reversal of idealism"» (Rowlands, 2013b, p. 528).

²⁹ «[...] [C]onsciousness can serve as the "ontological proof" of the existence of an in-itself. In fact, if there is consciousness of something, then it is necessary at the start that this "something" have a real being—that is, a being not relative to consciousness» (Sartre, 1956, p. 506).

world (Duncan, 2007; McCulloch, 1997), which is therefore against any attempt to multiply substances or to add intermediary or mediating entities between the thing-in-itself and our experience, like sensations/sense-data, mental images, etc³⁰. Surprisingly, this view is in line with reductive approaches to consciousness, which consider consciousness from a perspective which cannot invoke any other substance apart from the physical world, however the similarity with the Sartrean account ends here³¹. Indeed, we must keep in mind that the appearance-reality fallacy represents the boundary line between scientific and phenomenological account of consciousness, so that Sartre's realism cannot but be a phenomenological realism that deals with the infinite series of appearances of being, and not merely with the physical abstraction of it³².

5.2.2 Phenomenality I: Sartre and the Intentionalist approach

Now, Sartre's account of consciousness may seem similar to externalism which claims that certain types of intentional states are necessarily related to the environment, but in a really radical version of externalism, in the sense that every intentional state, i.e. conscious state, must be related to something that is external to our own individuality. So, in this sense, we might compare Sartre's Intentionality Thesis and No Content

³⁰ «Sensation, a hybrid notion between the subjective and the objective, conceived from the standpoint of the object and applied subsequently to the subject, a bastard existence concerning which we cannot say whether it exists in fact or in theory-sensation is a pure daydream of the psychologist. It must be deliberately rejected by any serious theory concerning the relations between consciousness and the world» (Sartre, 1956, p. 315).

³¹ «In the brief defense of phenomenology he advanced against the charge of idealism at the end of *Transcendence*, Sartre wrote of "metaphysical materialism" as an "absurdity." Half a decade earlier, in his contribution to *Les Nouvelles littéraires*, he had written: "Useless ... is the determinism which oddly enough attempts to create a synthesis of existence and being. We are free as you like, but helpless." Sartrean realism is incompatible with a metaphysically ordered and determined mechanism. It is a materialism of sorts, but one replete with contingency, and thus one without any formal guarantees. It is a realism-all-the-way-down» (Duncan, 2007, p. 102).

³² «Sartre was not, in fact, a commonsense realist about the world. But he was a phenomenological realist. The reality and objectivity of objects is to be understood in terms of their being constituted by a potentially infinite series of appearances, appearances that subsume and incorporate what we might regard as the "hidden" aspects of objects (e.g., their chemical structure). Appearances are, for Sartre, transcendent items rather than being parts of, or constructions out of, consciousness. The independence of things from consciousness and their ontological priority over consciousness is essential to Sartre's view, forming the basis of what he calls his "ontological proof," and his characterization of his view as a "radical reversal" of idealism» (Rowlands, 2010b, p. 226).

Thesis with intentionalism (Crane, 2009) also called tracking or strong representationalism (not to be confused with Classical Representationalism³³), according to which the content of perceptual experience derives from external objects rather from internal mental entities (Dretske, 1996, 2003; Lycan, 1998; Tye, 2002), and the phenomenal character is not a feature of experiences but supervenes on or is identical to our intentional content (intentional states like beliefs and thoughts)³⁴. Also, against internalist positions like the sense-datum theory, intentionalism argues that «representation of a certain kind suffices for a sensory quality, where the kind can be specified in functionalist or other familiar materialist terms, without recourse to properties of any ontologically “new” sort» (Lycan, 2019). This leads to the Exhaustion Thesis, which means that «a state’s phenomenal character is exhausted by its content. The exact meaning of this exhaustion thesis is that for every phenomenal character P there is some content C such that a state with P is nothing more than a phenomenal state with C as content» (Seager & Bourget, 2007, p. 262). In this sense, experience is transparent, i.e. when I experience a red tomato, I experience the colour as a feature of the tomato, not as an intrinsic property of my experience³⁵. The problem here is that such strong representationalist and «externalist theories of intentional content are ill suited to capture the qualitative character of such states, which possess a fine-grainedness and significance that is left out by an account in terms of purely

³³ «CR [Classical Representationalism] endorses the conclusion that what we directly perceive are mental entities variously called “ideas,” “impressions” or “sense-data.” This view has been called “representationalism” because such internal impressions act as pictures or signs of the external world. While not without insight, CR has several unpalatable consequences. First, it literally cuts us off from the world and each other. No one has ever been directly aware of anything but their own minds, hence no two people have ever been perceptually aware of the same thing. [...] Perhaps the main problem with CR’s posit of mental entities which possess all the features we are aware of in perception is that it seems to be irreconcilable with a physicalist world view. [...] MR [Modern Representationalism] rejects the assumption that leads to CR: not every perceptual experience requires that there be something perceived with qualities matching its phenomenal character. Instead, perceptual experience is understood in terms of representation, or intentionality.[...] This conception of experience should not be conflated with that of CR: according to MR, perception is representational in the sense that it is intentional, not in the sense that it is mediated by “internal pictures.”» (Seager & Bourget, 2007, pp. 261–262).

³⁴ «You experience them as being qualities of the surfaces. None of the qualities of which you are directly aware in seeing the various surfaces look to you to be qualities of your experience. You do not experience any of these qualities as qualities of your experience» (Tye, 2002, p. 138).

³⁵ «When Eloise sees a tree before her, the colors she experiences are all experienced as features of the tree and its surroundings. None of them are experienced as intrinsic features of her experience. Nor does she experience any features of anything as intrinsic features of her experience. And that is true of you too. There is nothing special about Eloise’s visual experience. When you see a tree, you do not experience any features as intrinsic features of your experience. Look at a tree and try to turn your attention to intrinsic features of your visual experience. I predict you will find that the only features there to turn your attention to will be features of the presented tree, including relational features of the tree “from here”» (Harman, 1990, p. 39)

objective properties» (Klausen, 2004, p. 14) or rather, they try to reduce the phenomenal-subjective question to intentionality³⁶. As I will explain later, for an exhaustive theory of consciousness phenomenal content may be consistent with the intentional content, but it is not possible to avoid the reference to a phenomenal manner of presentation of the object, i.e. subjectivity³⁷.

As Zahavi has briefly pointed out, «[I]ike latter-day representationalists, Sartre consequently takes phenomenal properties to be properties of worldly objects. For the very same reason, it makes no sense to say that they are located inside the head» (Zahavi, 2008a, p. 357). According to Sartre, the world is full of smells, sounds, colours, shapes, figures which are already totalities. There are not primary qualities subjectively fulfilled by secondary qualities, the two categories are abstractions that in the world are mutually implicated. Qualities are not manifestations of some neutral substances nor are they held together by the mind.

«Too often quality has been conceived as a simple subjective determination, and its quality-of-being has then been confused with the subjectivity of the psychic. [...] A quality does not objectivate itself if it is subjective. Supposing that we had projected the unity of an object-pole beyond qualities, at most each one of them would be given directly as the subjective effect of the action of things upon us. But the yellow of the lemon is not a subjective mode of apprehending the lemon; it is the lemon. And it is not true either that the object X appears as the empty form which holds together disparate qualities. In fact the lemon is extended throughout its qualities, and each of its qualities is extended throughout each of the others. It is the sourness of the lemon which is yellow, it is the yellow of the lemon which is sour. We eat the color of a cake, and the taste of this cake is the instrument which reveals its shape and its color to what we may call the alimentary intuition. Conversely if I poke my finger into a jar of jam, the sticky coldness of that jam is the revelation to my fingers of its sugary taste. The fluidity, the tepidity, the bluish

³⁶ «Is it really possible to investigate intentionality properly without taking experience, the first-person perspective, semantics, etc., into account? And vice versa, is it possible to understand the nature of subjectivity and experience if we ignore intentionality? Or do we not then run the risk of reinstating a Cartesian subject-world dualism that ignores everything captured by the phrase "being-in-the-world"?» (Zahavi, 2003, p. 66).

³⁷ «One might agree with the strong representationalist that sensory qualities are identical with intentional contents, but also contend that the latter intentional content properties cannot be characterized without reference to sensory qualities, so despite the identity there cannot be reduction without circularity» (Lycan, 2019).

color, the undulating restlessness of the water in a pool are given at one stroke, each quality through the others; and it is this total interpenetration which we call the this» (Sartre, 1956, p. 186).

The qualitative character, e.g. the yellowness of the lemon, the fluidity of water, are immediately given with the manifestation of the object of intentionality, it is not added by a subjective process to the lemon as made by primary qualities only. Rather, it is through the yellowness, the sourness etc. that the lemon is manifested. The qualities are some of the infinite series of manifestation of the things, and not a qualitative aspect of the experience. The yellowness that figures in one's experience is not a mental quality, it is the colour of the lemon itself, so there is no need to hypostatize intermediate entities like qualia, sense-data, etc.

Now, it seems that Sartre's phenomenological ontology and contemporary intentionalism have something in common regarding the attribution of the phenomenal properties to the intentional content without the intervention of subjective entities. However, there are also important differences, which have to do with the concept of representation and the subjective aspect of the experience. First, intentionalism, despite arguing that a state, in order to be conscious, does not need to be targeted by higher-order representations or abstract symbol manipulation, but only needs to instantiate an intentional object, this kind of approach to consciousness still deals with representations, i.e. intentionality and the intentional objects are explained in terms of representations and representational content (Lycan, 2019). This means that the perceived object, e.g. the red tomato, is an intentional object in the sense that it is a representation not the actual object in the environment. In other words, perceptual experience for intentionalism is not the direct manifestation of our surroundings, rather the representational ground for inferences about those surroundings. Instead, for Sartre (and for the phenomenological background we have developed in the previous chapter) experience is the direct "presentations"³⁸ of the things in the world by means of their appearances, so that the relation of intentionality is not a representing activity but a revealing/disclosing activity, the way Rowlands formulate it: disclosing activity as an

³⁸ «Rather than saying that we experience representations, it would be better to say that our experiences are presentational, and that they present the world as having certain features» (Zahavi, 2003, p. 75).

act that reveals an object to a subject *iff* the act is that in virtue of which object falls under empirical mode of presentation for subject (Rowlands, 2015, p. 16).

5.2.2.1 Perception and Imagination: the Disjunctivist approach

In this respect, it might be helpful to briefly focus on the difference between perception and hallucination (understood as a particular form of imagination). Indeed, for intentionalism the only difference between a veridical perceived and hallucinated object is that the former, besides being an intentional object is “also” an actual object in the environment, while the latter is “only” an intentional object. But the representational and phenomenal content of the two objects, e.g. a perceived red tomato and a hallucinated red tomato, is the same, so that it is difficult if not impossible to discriminate the two experiences without a verification of the actual presence of the intentioned object in the world. Therefore, the representing relation is not a concrete relation with the world, since the object can be represented whether or not it exists, and maintaining the same representational nature whether it is a perceived or imagined/hallucinated object (Crane, 2006). This is a position that is hardly suitable for Sartre, since one of the main topics of his early works were the different nature of perception and imagination (Sartre, 1940/2004a, 1936/2012). According to the author of *L’imagination* and *L’imaginaire*, the difference between perception and imagination is not a matter of true/false or good/bad correspondence to the objective world³⁹. Rather, perception and imagination are two experiential acts of different

³⁹ «Perhaps this way of proceeding would have led sooner or later to this observation: in spite of metaphysics there is between image and perception a difference of nature. But most authors have envisioned the question quite differently. They did not ask themselves if the psychic formations are not immediately given to consciousness for what they are; they placed themselves at the metaphysico-logical point of view of truth. They tacitly transformed the discrimination that every consciousness spontaneously makes between image and perception into a distinction between the true and the false. Thus Taine was able to say ‘that perception was a true hallucination’. Still truth and falsehood are not conceived of here as intrinsic criteria in the manner of Spinoza. It is a matter of a relation to the object. We are before a world of images. Those that have an external correspondence are called true or ‘perceptions’. The others are called ‘mental images’. We see the trick of prestidigitation: the givens of the intimate sense (*sens intime*) are transformed into external relations between a content of consciousness and the world, and the immediate distinction between the contents is replaced by a classification of these contents with respect to something other than themselves» (Sartre, 2012, pp. 180–181).

nature, since they involve a different attitude towards the world, i.e. a different kind of intentionality⁴⁰. Taken together, Sartre and intentionalism have in common the idea that phenomenal content is not a product or reaction of subjectivity, rather depends on the intentional object; nevertheless, they differ in how they conceive the relation with the intended object, where the former argues for experience as a direct relation to worldly objects while the latter understands experience as represented objects with represented properties which are only intentionally present in experience (Lycan, 2019).

From this perspective, the Sartrean account of consciousness is more similar to disjunctivism (Webber, 2001) which argues that perceptual experiences necessarily involve external things as constituents, providing direct reference to the perceived objects and that are intrinsically different from an hallucination, in the sense that there are no hallucinatory “perceptions” at all (Soteriou, 2016). Indeed, perception makes us relationally aware of external objects in the environment, while hallucinations are kinds of experience with an object different from the mind-independent objects of perceptual experience⁴¹. Disjunctivism agrees with intentionalism that our experience is not made of internal mental images (sense-datum theories), but disagree about the

⁴⁰ «The imaging consciousness of the object includes, as we noted above, a nonthetic consciousness of itself. This consciousness, which one could call transversal, has no object. It posits nothing, refers to nothing, is not knowledge (*connaissance*): it is a diffuse light that consciousness emits for itself, or – to abandon comparisons – it is an indefinable quality that attaches itself to every consciousness» (Sartre, 2004a, p. 14).

⁴¹ «[...] intentionalism and disjunctivism provide different theories of how perceptions relate us to external things, and of why they are non-neutral and immediate. Intentionalism about perceptual experiences does not distinguish between perceptions and hallucinations when addressing these issues. It maintains that perceptual experiences – whether they are veridical or not – make us intentionally aware of those entities. Accordingly, they are understood as presenting us with external objects without requiring their existence. They are non-relational appearances of things which do not involve the latter as their constituents. Moreover, intentionalism characterises the form of intentionality common to both perceptions and hallucinations as distinctively perceptual. This means that they involve a specific intentional attitude towards their objects which takes the latter to be actually present before us. This explains why they are committal to aspects of our actual environment, and why they have authority over our respective beliefs. Given that perceptions and hallucinations are said not to differ in respect of how they make us subjectively aware of things, intentionalism concludes that they share the same phenomenal character and are therefore of the same fundamental kind of experience. By contrast, disjunctivism about perceptual experiences (as it is defended by Martin) maintains that there is an essential difference between the character of perceptions and the character of hallucinations, to the effect that they belong to different fundamental kinds of experience. The difference in question is that, while perceptions make us relationally aware of objects, hallucinations do not. That perceptual awareness is relational means that the character of perceptions – and thus the perceptions themselves – are constituted by the external objects of which they make us aware. This explains why we find those objects when reacting on our perceptions, and also why the latter are non-neutral about the presence of the former in the situation before our senses» (Dorsch, 2010, p. 175).

fact that our experience is a re-presenting activity, rather disjunctivists claim that the objects in the world are given, presented to the subject as immediately available – “this” or “that” object “here” or “there”. If I am looking to a yellow lemon and after a moment a deity remove it from my visual field but preserving the cause of the experience – e.g. let us suppose the neural firing, for argument’s sake – I may not notice any change, still the pre-removal experience is completely different in nature from the post-removal one (Byrne & Logue, 2009)⁴². Similarly, Sartre holds that images, hallucinations and dreams are forms of consciousness completely different from perception (Sartre, 1956, p. 26), so that these imaginative experiences cannot be altered or malfunctioning instances of perceptual experiences (like illusions⁴³), since the intentional attitude of the imagining consciousness is completely different in nature from the perceiving intentionality⁴⁴. The positional consciousness of imaginative experience – which is completely different from the perceiving intentionality of objects as existing and present – can be directed towards objects as non-existent, absent or existent elsewhere, i.e. as “nothingnesses”, starting from material (shapes, images,

⁴² «Your “experience” is apparently not something that can be identified independently of its objects. So you cannot do better than reply: “I am having an experience of that object.” What is that yellow ovoid object that is present in your experience? It is, of course, overwhelmingly natural to take it to be the lemon. But is this right? After the deity has removed the lemon, it is no longer “right there” and “available.” If the nature of your experience has not changed, then the very same yellow ovoid object remains present in your experience. Therefore, this object is not the lemon. [...] Contraposing, if the very same yellow ovoid object does not remain present in your experience, then the nature of your experience has changed. Before the deity removes the lemon, it is present in your experience; after the removal, something else is going on entirely—no physical object is present in your experience, and maybe no object of any sort is. So the pre- and postremoval experiences are radically different» (Byrne & Logue, 2009, p. viii).

⁴³ «The dehydrated desert traveller who mistakes the hazy air above distant hot sand for water, for example, or the person who mistakes a distant tree for a person, is perceiving rather than hallucinating: the thetic component of the experience attempts and fails to track the manifest qualities of experience, mistakenly seeing the haze as water or the branches and trunk as arms and torso. This failure to track qualities is a feature of illusion but not hallucination: 'My perception can deceive me, but not my image' (Sartre 2004a, p. 9)» (Webber, 2001, p. 196).

⁴⁴ For example, we actually learn nothing from imagined objects because they are not products of a synthesis of different revealing acts of the world, rather a production of consciousness by itself: «[...] the object of the perception overflows consciousness constantly; the object of the image is never more than the consciousness one has of it; it is limited by that consciousness: nothing can be learned from the image that is not already known. In perception, knowledge is formed slowly; in the image, knowledge is immediate. We see now that the image is a synthetic act that links a concrete, not imaged, knowledge to elements more properly representative. An image is not learned: it is organized exactly as the objects that are learned, but, in fact, it is given whole, for what it is, in its appearance. If you turn a cube-image in thought to amuse yourself, if you pretend that it presents its various faces to you, then you will not be more advanced at the end of the operation: you will not have learned anything» (Sartre, 2004a, p. 8, 9).

patterns, suggestions) used as *analogon*⁴⁵. Thus, even imagination (and its results like images, hallucinations, dreams, etc.) is not a pure activity of consciousness by-itself but needs a background material as an analogic springboard for the production of non-perceived objects, indeed it should always be kept in mind that «consciousness is born supported by a being which is not itself» (Sartre, 1956, p. lxi)⁴⁶.

To sum up, I first focused on the issue of phenomenality that I exposed in the second chapter, and I compared the Sartrean phenomenological approach with the intentionalist one, finding some interesting common points, like the union of phenomenal and intentional content, but also important differences, like Sartre's direct phenomenological realism against intentionalists' representations. Thus, I showed other similarities with another contemporary position of philosophy of mind – disjunctivism – which argue for an intrinsic and necessary difference between perception and hallucination, similarly to Sartre's distinction between perceptual and imaginative experiences. I am not sure whether Sartre's phenomenological realism may be considered as disjunctivism or intentionalism, rather I think that his approach cannot be condensed in these ways, and also that this issue falls outside the aims of the present work. Nevertheless, I believe that it can be useful to compare and analyse Sartre's principal claims with contemporary tools from philosophy of mind, in order to clarify better at the conceptual level. Now, I shall come back to the comparison between Sartre and intentionalism in order to highlight other details about the issue of phenomenality.

5.2.3 Phenomenality II: Self-awareness and the problem of Subjectivity

Previously, I have explained that the comparison between intentionalism and Sartre's account of consciousness presents not only similarities but also intrinsic

⁴⁵ «Presently my intention appears; I say: 'This is the portrait of Pierre' or, more briefly: 'This is Pierre'. Then the portrait ceases to be an object, it functions as matter for an image. This invitation to perceive Pierre has not disappeared, but has entered into the imagined synthesis. To tell the truth, it is the invitation that functions as an analogon and it is through it that my intention is directed at Pierre. I say to myself: 'Look, it's true, Pierre is like that, he has these brows, this smile.' All that I perceive enters into a projective synthesis that aims at the true Pierre, a living being who is not there» (Sartre, 2004a, p. 22).

⁴⁶ The material or analogue is not itself observed as in perception, it is not 'contemplated' (PI: 49), which would require the thetic component of the experience to be concerned only with tracking its qualities. Instead, it forms the intuitive basis for the 'quasi-observed' intentional object (Webber, 2001, p. 193).

differences. Now, I shall present another gap between the two position that will allow me to introduce an important issue of phenomenality, which I briefly exposed in the second chapter: subjectivity. According to intentionalism, since the qualitative aspects supervene or are identical to the intentional content, it is relatively easy to account for the phenomenal content, e.g. with physics that can provide us with an account of that physical property (for example, the yellowness in terms of surface reflectances). Hence, for intentionalism if the phenomenal character is a question of intentionality and intentionality can be reduced in physical terms, there is no hard problem of consciousness anymore (Zahavi, 2003). We have already seen that this reductive move it is not acceptable by a phenomenological and pluralist approach like this. In particular, although Sartre agrees in believing that the phenomenal qualities belong to the objects of experience and not to the experience itself, this does not mean that for him consciousness is completely reducible to intentionality in causal and physical terms, rather the opposite⁴⁷. In order to delineate the main difference between the two positions, I will briefly turn back on the example of the experience of a chair. As Sartre says in *The Transcendence of the Ego* the sentence “I am conscious of this chair” is already a product of reflective consciousness, which postulates the presence of an “I-subject” that has intentionality towards a chair, i.e. a mento-centric perspective, so that he alternatively proposes a sentence like “there is a consciousness of this chair”. Why not “there is a chair”? After all, the chair is a totality as such, and all the possible qualitative characters (woodiness, shininess or roughness, shape, etc.) supervene or are identical to the intentional object (the chair). If he had proposed “there is a chair” – which of course he does not accomplish – Sartre would have lost the structure of positional consciousness that allows the appearance of being to the subjects without

⁴⁷ Dretske’s and Tye’s intentionalistic interpretation of phenomenal qualities has the great advantage of staying clear of any kind of immanentism. As already mentioned, it also bears a certain resemblance to views found in phenomenology. This is in particular the case for the analysis offered by Sartre. Sartre is renowned for his very radical interpretation of intentionality. To affirm the intentionality of consciousness is, according to Sartre, to deny the existence of any kind of mental content (including any kind of sense-data or qualia) (Sartre 1956). There is nothing in consciousness, neither objects nor mental representations. It is completely empty. Thus, for Sartre, the being of intentional consciousness consists in its revelation of transcendent being. Sartre consequently takes the phenomenal qualities to be qualities of worldly objects, and certainly not to be located within consciousness. However, from the fact that consciousness is nothing apart from its revelation of transcendent being (or as Tye and Dretske would probably say, from the fact that it exhausts itself in its representation of external reality), Sartre would never infer that intentional consciousness is therefore no problem for reductionism. On the contrary, in his view, it is exactly the emptiness (or non-substantiality) of consciousness, that demonstrates its irreducibility (Zahavi, 2003, p. 76).

ever exhausting it. In other words, he would have lost the fact that an intentional object is always given to someone through some appearances, and this is the point that is not considered by intentionalism, which instead would have chosen the “there is a chair” sentence because its aim is to reduce the phenomenal character to the physical states of the intentional object.

However, if what I argued in the fourth chapter is true, the intentional object is not made by physical states, rather it can be conceived as manifested to a consciousness (also) by physical states. Taken together, on one hand we have the mentalistic position, which maintains that experience is to connect external objects with internal subjective states, and which consequently leads to issues like the nature of qualia and the hard problem. On the other hand, we have the intentionalism, which shifts all the qualitative states outside in the world, so that we are not aware of the qualities of our experience but of the properties of the perceptual object, nevertheless producing other issues. For example, saying that there being something it is like to see a yellow lemon is just one’s consciously seeing it (Harman, 1990; Tye, 2002) does not exhaustively explain what being aware of something means. Moreover, there is no way to know how the object I perceive, with its phenomenal characters, is given to me and for me as my own experience, since phenomenal characters are related to a subject but at the same time they come from the intentional content. In other words, if consciousness simply consists of in being conscious of what one perceives, how can I be aware of that as the one who is perceiving such qualitative properties? If I do not experience any features of anything as intrinsic features of my experiences (Tye, 2014), how can I be conscious of the external objects as for me (for-me-ness), as parts of mine own experience (mineness), as limited to my perspective (perspectivalness)? The only way seems to be the introspection my own experience, without creating an inner mental state⁴⁸, similarly to what has been explained in the previous chapter against mento-centrism.

⁴⁸ «But in so doing, you are not aware of any inner object or thing or event. When you introspect your visual experience, the only particulars of which you are aware are the external ones making up the scene before your eyes. You are not aware of those objects and a further inner object or episode. Your awareness is of the external surfaces and how they appear. The qualities you experience are the ones the surfaces apparently have. Your experience is thus transparent to you. When you try to focus upon it, you ‘see’ right through it, as it were, to the things apparently outside and their apparent qualities. But when you introspect, you are certainly aware of the phenomenal character of your visual experience. [...] Visual phenomenal qualities or visual qualia are supposedly qualities of which the subjects of visual experiences are directly aware via introspection. Tradition has it that these qualities are qualities of the

«On the basis of introspection, you know what it is like for you visually on the given occasion. Via introspection, you are directly aware of a range of qualities which you experience as being qualities of surfaces at varying distances away and orientations and thereby you are aware of the phenomenal character of your experience. By being aware of the external qualities, you are aware of what it is like for you. This is not to say, of course, that you infer the phenomenal character of your experience from your awareness of the external qualities. Obviously, no reasoning is involved. Still, by attending to what you experience outside, as it were, you know what it is like inside. So, your awareness of phenomenal character is not the direct awareness of a quality of your experience. Relatedly, the phenomenal character itself is not a quality of your experience to which you have direct access» (Tye, 2002, p. 139).

However, when we do not introspect our own conscious acts, we are not aware of features of our visual experience, of the visual experience itself, we cannot attend to features of the visual experience, but the only features of which we are aware are external features (Tye, 2014). If so, how can for-me-ness, mineness and perspectivalness structure my own experience? To make this clear, it is necessary to make a phenomenological distinction between what the object is like for the subject (i.e. what does the object feel like to the perceiver) and what the experience of the object is like for the subject (i.e. what does the perceiving feel like to the perceiver) (Zahavi, 2003, p. 78). In this sense, the hard problem, the what-it-is-likeness of experience is surely referred to objects to the world and not to ineffable, non-physical, private qualia (as intentionalists argue), however the mineness of the intentional act is an experiential property⁴⁹. As Zahavi notes, this distinction, which is not considered by intentionalism – is purely conceptual, that is, «[b]oth the worldly properties of the

experiences. Tradition is wrong. There are no such qualities of experiences. [...] Still there are qualities of which the subjects of visual experiences are directly aware via introspection. They are qualities of external surfaces (and volumes and films), if they are qualities of anything. These qualities, by entering into the appropriate representational contents of visual experiences, contribute to the phenomenal character of the experiences. So, they may reasonably be called “phenomenal qualities” in a less restrictive sense of the term» (Tye, 2002, pp. 139, 141–142).

⁴⁹ «In contrast to the redness of the tomato or the bitterness of the tea (both of which are worldly properties), the mineness characterizing the perception of the redness or bitterness is not a worldly property, but an experiential property. When asked to specify ‘what the experience of the object is like for the subject,’ this is exactly one of the features to mention. In short, the experiential dimension does not have to do with the existence of ineffable qualia, it has to do with this dimension of first-personal experiencing. The ‘what it is like’ question has two sides to it: ‘what is the object like for the subject’ and ‘what is the experience of the object like for the subject.’» (Zahavi, 2003, p. 80)

appearing object and the experiential properties of the modes of givenness are part of the phenomenal dimension. They are not to be separated, but neither are they to be confused» (Zahavi, 2003, p. 82). I think that intentionalism, in its effort to clean consciousness out of all the mental properties, substances and entities in general, risks to throw out the baby (subjectivity) with the bathwater (qualia, sense-data, etc.). For this reason, we need a deeper phenomenological analysis.

5.2.3.1 Pre-reflective Non-positional Consciousness

As said, Sartre has some points in common with contemporary intentionalism, regarding the phenomenal features of the objects of intentionality.

«All consciousness is positional in that it transcends itself in order to reach an object, and it exhausts itself in this same positing. All that there is of intention in my actual consciousness is directed toward the outside, toward the table; all my judgments or practical activities, all my present inclinations transcend themselves; they aim at the table and are absorbed in it. Not all consciousness is knowledge (there are states of affective consciousness, for example), but all knowing consciousness can be knowledge only of its object» (Sartre, 1956, pp. li–lii).

Indeed, subjective phenomenal features are not properties of our mind or our experience, but properties of the things in the way they are manifested to us.

«So all at once hatred, love, fear, sympathy – all those famous “subjective” reactions that were floating in the malodorous brine of the mind – are pulled out. They are simply ways of discovering the world. Things are what abruptly unveil themselves to us as hateful, sympathetic, horrible, lovable. Being dreadful is a property of this Japanese mask: an inexhaustible and irreducible property that constitutes its very nature – and not the sum of our subjective reactions to a piece of sculptured wood» (Sartre, 1939, p. 131).

Though, at the same time, Sartre strives to develop an account of consciousness that is able to consider the transphenomenality of being, i.e. the fact that the objects of consciousness are always outside and transcendent with all their qualitative characters, without falling into the materialistic and reductive versions of intentionalism, where

“being conscious of something” is reduced to physical relations⁵⁰. In this sense, Sartre’s aim is to strike a balance between the transcendent being of the world and the consciousness’ intake (self-awareness and phenomenality)⁵¹, so that it could be possible to avoid both idealism and mechanistic materialism⁵².

First, claiming that when we are conscious of some object with qualitative properties (red tomato) we are aware only of the phenomenal character of that object (redness of the tomato) does not explain anything about what being aware means. Sartre goes to the roots of the question and, inspired by Heidegger, argues that existence (the for-itself, consciousness) is essentially an activity of negation, i.e. nothingness in the sense of the surpassing activity (*depasserment*) towards and within the world that permits the determination of itself as self-transcendence⁵³. In other words, consciousness as disclosing activity surpasses itself (intentionality and no content) towards the world, so that it encounters being as something which is not itself, which is not consciousness. «It is impossible to construct the notion of an object if we

⁵⁰ «By providing a twostep reduction of consciousness to tracking, tracking representationalism offers an attractive physicalist theory of consciousness. The first step reduces consciousness to intentionality and manners of representation. [...] The second step reduces intentionality to a species of tracking relation, and manners of representations to broadly physical properties. The reduction of intentionality to tracking is somewhat plausible since many of the currently popular theories of intentionality are tracking theories. Combining the two steps, then, offers us a reduction of phenomenal consciousness to broadly physical properties. The technicolour of experience is just a matter of tracking relations between internal states and properties in the world, when such relations have the right additional physical properties to be conscious» (Bourget & Mendelovici, 2014, p. 219).

⁵¹ «The question, which I tried to solve during my whole life, was: how to give man his autonomy, and at the same time his reality among real objects, avoiding idealism and without falling on the other hand into a mechanistic materialism?» (Sartre, 1977, p. 153; translation mine).

⁵² In this regard, I also believe that some terms used in philosophy of mind like “internalism” and “externalism” – as well as “mind-dependent” and “mind-independent” – would be for Sartre, *pace* Rowlands, mere theoretical abstractions based upon the appearance-reality fallacy that does not apply to the phenomenological approach: «Considering the way in which phenomenologists conceive of intentionality, of the mind-world relationship, I think it is questionable whether it really makes much sense to classify their views as being committed to either internalism or externalism. Avoiding the two terms obviously won’t solve all the problems, but might at least permit us to avoid letting our investigation be guided by misleading metaphors. The mind is neither a container nor a special place. Hence it makes little sense to say that the world must be either inside or outside of the mind» (Zahavi, 2008a, pp. 371–372).

⁵³ «All determination for Heidegger is surpassing since it supposes a withdrawal taken from a particular point of view. This passing beyond the world, which is a condition of the very rising up of the world as such, is effected by the Dasein which directs the surpassing toward itself. The characteristic of selfness (Selbstheit), in fact, is that man is always separated from what he is by all the breadth of the being which he is not. He makes himself known to himself from the other side of the world and he looks from the horizon toward himself to recover his inner being. [...] Dasein is "outside of itself, in the world"; it is "a being of distances"; it is care; it is "its own possibilities," etc. All this amounts to saying that Dasein "is not" in itself, that it "is not" in immediate proximity to itself, and that it "surpasses" the world inasmuch as it posits itself as not being in itself and as not being the world» (Sartre, 1956, pp. 17, 18).

do not have originally a negative relation designating the object as that which is not consciousness» (Sartre, 1956, p. 173). Instead of being an I-subject that determines the external objects (“I am conscious of this chair”) or being a pure intentional and phenomenal content of the world (“there is a chair”), this negative relation permits, at the same time, the determination of the object as part of an infinite series of appearances and of myself as the subject to which this series is manifested:

The reality of that cup is that it is there and that it is not me. We shall interpret this by saying that the series of its appearances is bound by a principle which does not depend on my whim. But the appearance, reduced to itself and without reference to the series of which it is a part, could be only an intuitive and subjective plenitude, the manner in which the subject is affected. If the phenomenon is to reveal itself as transcendent, it is necessary that the subject himself transcend the appearance toward the total series of which it is a member (Sartre, 1956, p. xlvii).

In this way, we can interpret the sentence “there is consciousness of this chair” in the sense that there is an experience of a chair as a transcendent object given to someone. So, being conscious of something means being directed towards something “that it is there” (red tomato) and also that it is not me.

The appeal to the recognition of subjectivity from what it is not a subjectivity is what is missing in radical intentionalism. How does Sartre justify the emergence of such a structure without falling into the hypostatization of some mental activity? For being aware of something that is out there, and it is not me, we must be aware also of our own conscious act. But if we put subjectivity as a *primum* we fail to meet the aim of consciousness as pure transcendence towards something which is not consciousness⁵⁴. Sartre’s pre-reflective non-positional consciousness, which I already described in the second chapter, is the solution of this problem and the main difference from radical intentionalism. After having described consciousness as pure intentionality towards being, he adds:

⁵⁴ «To say that consciousness is consciousness of something means that for consciousness there is no being outside of that precise obligation to be a revealing intuition of something - i.e., of a transcendent being. Not only does pure subjectivity, if initially given, fail to transcend Itself to posit the objective; a "pure" subjectivity disappears» (Sartre, 1956, p. lxi).

However, the necessary and sufficient condition for a knowing consciousness to be knowledge of its object, is that it be consciousness of itself as being that knowledge. This is a necessary condition, for if my consciousness were not consciousness of being consciousness of the table, it would then be consciousness of that table without consciousness of being so. In other words, it would be a consciousness ignorant of itself, an unconscious-which is absurd. This is a sufficient condition, for my being conscious of being conscious of that table suffices in fact *for me* to be conscious of it. That is of course not sufficient to permit me to affirm that this table exists in itself-but rather that it exists *for me* (Sartre, 1956, p. lii; emphasis mine).

Being conscious of something without being conscious of being so is not possible (see chapter two for details), indeed being conscious of something means necessarily that this something exists “for me”. This “for me” is not a substantial subject because it directly derives from the appearance, but it is similar to those concepts of for-me-ness, mineness and subjective givenness exposed by contemporary phenomenologists (Zahavi, 2019a; Zahavi & Kriegel, 2015) and linked to Sartre’s pre-reflective non-positional awareness, which is the only possible subjectivity:

«What can properly be called subjectivity is consciousness (of) consciousness. But this consciousness (of being) consciousness must be qualified in some way, and it can be qualified only as revealing intuition or it is nothing. [...] In particular there is no point in replying that in fact subjectivity implies objectivity and that it constitutes itself in constituting the objective; we have seen that subjectivity is powerless to constitute the objective» (Sartre, 1956, pp. lxi, lxii).

Therefore, subjectivity as pre-reflective consciousness is a feature of the intentional relation (intentional content as being for the subject), rather than a constituent of it (Levine, 2016), and is neglected by radical intentionalism which hopes to eliminate the hard problem by avoiding the existence of qualia⁵⁵. But,

⁵⁵ «I am aware not simply of the intentional objects and properties presented by my experience, but also of features of my experience, or rather of my ongoing activity of experiencing. [...] We could summarize this last point by saying that the extreme transparency of awareness thesis neglects that constitutive feature of experience we can call (following Husserl and Sartre) prereflective self-consciousness. In my visual experience of the wine bottle, I am explicitly aware of the bottle, but also implicitly aware of my visual experience of the bottle. This sort of implicit awareness is a kind of self-consciousness (I am implicitly aware of the visual experience as mine). But it is not a reflective or introspective self-consciousness, because there is no phenomenally conscious reflection or introspection that takes the experience as its object» (Thompson, 2008, p. 403) «The decisive difficulty for reductionism is not the

subjectivity, while not dealing with qualia, as a feature of intentionality is not part of an inner mental life but is «essentially oriented and open toward that which it is not, and it is exactly in this openness that it reveals itself to itself. What is disclosed by the cogito is, consequently, not a self-contained immanence or a pure interior self-presence, but an openness toward alterity, a movement of exteriorization and perpetual self-transcendence» (Zahavi, 2003, pp. 73–74)⁵⁶.

5.2.4 Intentionality goes Adverbial

So, the pre-reflective self-awareness is the necessary condition by means of which every intentional content – with its phenomenal character – is given to us as something which is not us, and thus giving rise to the feature of subjectivity; otherwise, without this surpassing activity, consciousness would be dispersed around the objects of the world, simply put, the concept of consciousness would be superfluous and we could not understand how the phenomenal content of the object is given to us. As I explained in the second chapter, pre-reflective awareness is non-positional, i.e. it is not directed to an object, and its non-intentional nature does not invalidate the Intentional Thesis, i.e. every consciousness is intentional, because this kind of consciousness exists only as a necessary condition of positional, intentional consciousness. Someone might argue that if pre-reflective self-awareness is non-intentional, then it cannot be something related to the objects of the world; therefore, it must be an internal state of our consciousness, like a mental state about our intentional states; and this would go against the No Content Thesis, showing that radical intentionalism is the right account of consciousness. However, this argument falls short to address pre-reflective self-

existence of epiphenomenal qualia, qualia in the sense of atomic, irrelational, ineffable, incomparable, and incorrigible mental objects. And the hard problem does not disappear if one (rightfully) denies the existence of such entities, and if one, so to speak, relocates the phenomenal from the 'inside' to the 'outside.' The hard problem is not about the existence of non-physical objects of experience, but about the very existence of subjective experience itself, it is about the very fact that objects are given to us» (Zahavi, 2003, p. 75).

⁵⁶ For Sartre, the same is true with emotions, understood as a specific kind of openness towards transcendence instead of an internal reaction of the subject: «For the majority of psychologists everything happens as though the consciousness of emotion were primarily a reflective consciousness; [...] And certainly, it is always possible to become aware of emotion as a fact of consciousness, as when we say: I am angry, I am afraid, etc. But the fear does not begin as consciousness of being afraid, any more than the perception of this book is consciousness of perceiving it. The emotional consciousness is at first non-reflective, and upon that plane it cannot be consciousness of itself, except in the non-positional mode. The emotional consciousness is primarily consciousness of the world [...] In a word, the emotional subject and the object of the emotion are united in an indissoluble synthesis. The emotion is a specific manner of apprehending the world» (Sartre, 2015, pp. 56, 57).

awareness/non-positional consciousness, since this latter has a different nature compared to positional consciousness⁵⁷.

In order to show how the No Content Thesis is preserved, I shall use the adverbialist interpretation of the Sartrean account of consciousness made by Rowlands (Rowlands, 2013b, 2016). Adverbial theories of perceptual consciousness hold that phenomenal qualities are modifications of the experience itself rather than properties of the object of intentionality, nevertheless without involving any kind of subjective sense-data (Crane, 2017). In this sense, adverbialism agrees with intentionalism that there is no need for strong ontological commitment about mental states as in sense-datum theories; contrarily to intentionalism, it denies that the phenomenal character supervenes or is identical to the intentional content, rather it is non-relational (not linked to the object) and intrinsic to the subject⁵⁸. The adverbial theories are so-called because instead of understanding the conscious act as “I am perceiving a red square”, they say “I am perceiving squarely and redly”, i.e. using the adverb to explain the modification of the experience. There is no implication of hypostatized entities in the experience, there are only the subject who instantiates the experience and the ways this experience is modified.

Now, someone might argue that adverbialism is incompatible with intentionalism, since the former shifts the phenomenal character into the subjective modification of experience, while the latter consider phenomenality as determined by the intentional content only. This would be true if the objective here was to merge the two theories on the same consciousness structure. However, as Rowlands explained, we need to keep separate positional and non-positional consciousness, where the first is to be aware of objects while the second is to be conscious (of) consciousness in a

⁵⁷ «How do we reconcile this interpretation with NCT, and Sartre’s frequent characterizations of consciousness as “nothingness”, “total emptiness” – as something with the “entire world outside it”? [p. 15]. If the structure of pre-reflective consciousness is understood by analogical extension from reflective consciousness, then consciousness must be something for pre-reflective consciousness to be consciousness (of) it. If consciousness is really to be “nothingness”, then pre-reflective consciousness must have an entirely different nature. The question, then, is precisely what this nature must be» (Rowlands, 2016, p. 115).

⁵⁸ «Adverbialism is an attempt to get rid of mysterious mental objects and properties by appeal to manners of perceiving. In so doing, adverbialism seems to deny altogether that perception is a relation. Perception emerges in this theory as a subjective act that can be done in different manners, rather than as a relation to an outside world. Thus, adverbialism seems to react to sense-data theory by denying something that most other theories accept: the relational status of perception» (Kind, 2018, p. 110).

certain mode or manner (Rowlands, 2013b). So, positional consciousness deals with the intentional objects, like the tomato and its qualitative characters (redness, shininess, etc.) – but also with my own mental states as objectified by a positional reflective consciousness, e.g. thinking about me seeing a red tomato. Instead, non-positional consciousness qualifies the act of positional consciousness rather than supplying a description of the object of the act, in this sense it is completely different from the introspection prepared by intentionalism as awareness of external features. Non-positional consciousness does not posit positional consciousness as intentional object (pardon the pun) but modifies or qualifies my directedness to the world and nothing more, so that the No Content Thesis is preserved⁵⁹. Thus, while the best way to describe positional consciousness is through intentionalism, the best way to explain non-positional consciousness is through adverbialism, i.e. not in relation to the phenomenal character of experience as in classical adverbialism⁶⁰.

Classical adverbialism claims that the redness of my experience is not a property grasped by introspection but a certain way of making experience of something, i.e. “redly”. Similarly, regarding our account of non-positional consciousness, the sense of ownership of the experience, the fact that the objects of experience are given for me and I feel like they are presented to me and only to me specifically, all these characteristics that we previously called for-me-ness, mineness, or subjectivity are not properties that I understand through introspection, i.e. a particular kind of positional reflective consciousness; these characteristics, as seen with Sartre, qualify my conscious intentional acts but they are not intentional objects, in an adverbial sense I

⁵⁹ «In particular, it carries no commitment to the idea—an idea that Sartre must reject—that when I am pre-reflectively aware, my consciousness is directed towards my consciousness in any more substantial sense. Pre-reflective consciousness characterizes my directedness to the world. It does not indicate a substantial form of directedness towards my own consciousness over and above modifying my directedness towards the world» (Rowlands, 2013b, p. 535).

⁶⁰ In this way, some typical criticisms of classical adverbialism that focus on the phenomenal content of consciousness are avoided, such as the many-property problem which highlights the difficulty of an adverbial consciousness to distinguish conscious states with various objects and properties: «Consider someone who senses a brown square and a green triangle simultaneously. The adverbial theory will characterise this state of mind as “sensing brownly and squarely and greenly and triangularly”. But how can it distinguish the state of mind it is describing in this way from that of sensing a brown triangle and a green square?» (Crane, 2017). However, since this kind of adverbialism applied to Sartrean account of consciousness does not deal with phenomenal content but with the qualification of conscious acts as such which are taken one by one, the problem does not arise: «An adverbial interpretation of pre-reflective consciousness is not at all the same thing as an adverbial interpretation of phenomenal content – for the simple reason that an account of pre-reflective consciousness is not the same thing as an account of phenomenal content» (Rowlands, 2016, p. 102).

see/imagine/perceive things “minely”, in the sense that «[w]hen I have positional consciousness of an object, I am non-positionally aware of this because, and precisely because, I experience the intentional object not simply as an object with various properties but, crucially, as something that is presented to me» (Rowlands, 2016, p. 117). There is no introspection, no reflection, no mental properties or sense-data; I am “non-positionally” aware of my experiences in the sense that my non-positional consciousness modifies and qualifies the directedness to the world⁶¹.

Taken together, in this chapter I focused on the necessity of an alternative approach to consciousness, which could avoid both the mento-centric and neuro-centric paradigms, promoting instead interdisciplinarity and ontological pluralism. I presented different approaches characterized by the attempt to put together philosophy, neuroscience and phenomenology. Regarding this, I defined the present work as a kind of non-reductive neurophilosophy in a broad sense, since it includes also an important contribution from Sartre’s phenomenology. Then, I started the conceptual and phenomenological analysis of consciousness, taking up the provisional description indicated in the second chapter, and I deepen it through additional phenomenological investigation from Sartre, comparing it with some relevant position in contemporary philosophy of mind. The result is an account of consciousness which attempts to overcome, on one hand, the appearance-reality fallacy through a concept of intentionality as direct relation towards the world, and, on the other hand, the hypostatization of mental states through the interpretation of subjectivity as pre-reflective non-positional consciousness in an adverbial way. Now, it is time to give bones, flesh and tendons to this theory, focusing on the relation between the body and the world, from both phenomenology and neuroscience.

⁶¹ «The same phenomenon can also be explained in adverbial terms. When I have experiences, I have them *minely*. Their mine-ness is an adverbial modification of the act rather than an introspectively grasped property of an object of that act. This adverbial modification is not a brute, unexplainable feature of the experiential act. Rather it stems from the nature of positional consciousness itself and, in particular, the way this sort of consciousness presents its objects. The adverbial modification results from the fact that the objects of my experiences – the intentional objects of my positional acts of consciousness – are presented, precisely, as things that are *for-me*» (Rowlands, 2016, p. 116).

6. Experience as Embodied-Being-in-the-World and the Environment-Body-Brain Relation. A Neurophilosophical account

This last chapter represents the final synthesis of the present dissertation, where I expose in detail an account of consciousness understood as embodied-being-in-the-world (phenomenological perspective) and environment-body-brain alignment (neuroscientific perspective). The aim indeed is to give bones, flesh and tendons to the interdisciplinary approach previously exposed, focusing on the relation between the body and the world, from both phenomenology and neuroscience. First, I will rely on Sartre's early phenomenological studies (from *The Imagination* to *Being and Nothingness*), in particular to re-evaluate his complex analysis of corporeity, often unfairly overshadowed by the studies of Husserl and Merleau-Ponty. In this sense, I will analyse bodily consciousness as intrinsically situated and connected with the world, through a body that is often "surpassed" and "passed by in silence", to use Sartre's words. In order to build a bridge with contemporary studies of consciousness and the body, I will compare Sartre's phenomenological analysis with the key concepts of the 4E Cognition. This will also lead me to consider the possibility that, in order to understand how we exist our body, we need to understand first how we practically live (in) the world, in search of a new paradigm for the study of consciousness that is not founded neither on the brain nor on the body alone, but on our relationship with the world. This relationship may be called "embodied-being-in-the-world" and I will expose and discuss it, highlighting its characteristics in relation to how we experience the world in a spatio-temporal way. Finally, in order to be faithful to my interdisciplinary project, I will analyse consciousness from a neurophilosophical point of view, discussing how the consciousness-world connection is conceived by some of the most relevant neuroscientific theories of consciousness, consequently relating the embodied-being-in-the-world condition to the concept of "world-brain relation"

proposed by Georg Northoff. From here, I will discuss Northoff's theory and give an account for what I call the environment-body-brain alignment as the empirical description of the ontological concept of world-brain relation, and also as a necessary precondition for the maintenance of consciousness. This will involve a final test for the theory here exposed, both at a phenomenological and a neuroscientific level, a sort of *experimentum crucis* that can help us to understand if I can really provide a description of our experience as embodied-being-in-the-world and environment-body-brain alignment. In this regard, therefore, I will propose two cases in which consciousness seems to exist in absence of relationship with the body and the environment, namely the Locked-in Syndrome, a rare and severe clinical condition, and the dreaming state intended as a case of "disconnected consciousness". These paradigmatic cases can offer us an empirical, phenomenological and conceptual test to the theory presented in this research, a neurophilosophical and neurophenomenological account of experience that is, above all, based on the concept of being in connection with the surrounding world. This concept of consciousness-in-the-world will suggest the need to reconsider consciousness not merely in terms of brain- or mind-relation, but intrinsically as world-related.

6.1 Consciousness as Embodied-Being-in-the-World

So far, I have focused on the philosophical and phenomenological analysis of consciousness. I have exposed an account of consciousness mainly based on the Sartre's works, connecting his phenomenology with contemporary studies in philosophy of mind. I have developed the structure of consciousness as intentional and devoid of mental contents, connecting these concepts with the intentionalist and externalist account of phenomenal experience. I then tried to overcome some possible limits of a radical intentionalism that could not completely give account of the subjectivity of experience, by focusing on Sartrean non-positional consciousness and the adverbialist interpretation made by Rowlands. The result is a concept of consciousness as directedness towards the world without ontological commitment on mental entities and sense-data, but with the feature of qualifying any single act of experience subjectively.

Nonetheless, following the project of an interdisciplinary and ontologically pluralist account of consciousness, these analyses still are too conceptual and theoretical. We need a deeper investigation of what concretely these consciousness acts mean. I claimed that intentionality is the condition that connects us with the world, but how do intentionality concretely work? We are not talking about evanescent entities but human beings, whose existence is determined by their being situated into the world, a world they encounter from the beginning not through an imaginary arrow that starts from the head towards the things, but through their bodies. Even more, as I will explain, not only do they meet the world, but the world meets them. For this reason, we need now to focus on consciousness as body- and world-related. Consciousness is relation to the world through the body so it must involve embodiment and embedment. As Rowlands says, intentional acts as well as mental processes in general are an amalgamation of neural, bodily and environmental structures, processes and transformations, so that they endorse embodiment and embedment (Rowlands, 2010b). Now, the aim is to analyse this amalgam of different processes and define the necessary conditions for consciousness as body in the world. In order to do this, first I will go on with the phenomenological analysis of bodily consciousness and the condition of being-in-the-world, and finally I will expose some neuroscientific findings that provide a theory of how consciousness is structured through an environment-body-brain relation, instead of a mere neuro-centric framework.

6.1.1 Why Sartre?

The theme of the body has already been briefly mentioned in chapter two, where I used the example of body awareness as a form of pre-reflective non-positional consciousness (M. Boyle, 2018). Here I shall deepen the question further, showing how the body is a necessary condition for the development of consciousness as intentionality towards the world. Classical and contemporary phenomenology take into account the body as one of the main themes for a complete and reliable description of experience, since there is no consciousness of the objects of the world without a body, and even during instances of so-called disconnected consciousness, e.g. dream or imagination, we cannot conceive us completely devoid of a body.

Now, for this analysis I will make use of Sartre's phenomenological ontology of the body. Someone might ask why Sartre and not Merleau-Ponty since this latter is known as, perhaps, the most important phenomenologist on the topic of the body, indeed some of his major works are specifically dedicated to the body and its relation with the world (Merleau-Ponty, 1964, 1945/2013). It is likely that Merleau-Ponty's works focused primarily on the body have overshadowed the bodily consciousness analysed by Sartre in the third part of the second chapter of *Being and Nothingness*¹. However, recently several authors have argued for the importance of Sartre's contribution on the phenomenology of the body, in relation with contemporary studies of embodied, embedded, extended and enactive cognition, and against a neo-Cartesian interpretation of his studies² (Boulé & O'Donohoe, 2011; M. Boyle, 2018; Moran, 2011; K Morris, 2009; Wider, 1997, 2016). Following this line of research, I shall demonstrate how his conception of the body could be extremely useful for understanding the relation between consciousness, body and world.

Another possible criticism could be that which would see Merleau-Ponty as the principal classical phenomenologist interested in combining phenomenological analyses with scientific findings, when instead Sartre is often depicted as an old-fashioned philosopher. This would lead to doubting Sartre as the most suitable phenomenologist to cite in a study that wants to be mainly interdisciplinary. But this is a mistaken criticism, since Sartre seizes information deriving from the psychology of his time and early neuroscience, particularly in his early works (Sartre, 1940/2004a,

¹ «Despite this burgeoning interest in the body, Sartre's contributions remain relatively neglected. [...] There are no doubt explanations for this neglect: his principal chapter on the body is tucked away, seemingly almost as an afterthought, in Part III (of four parts), chapter 2 (of three chapters) of *Being and Nothingness* [...]; and many commentators take Merleau-Ponty's critiques of Sartre at face value» (K. J. Morris, 2010, pp. 3–4). «If we look at this outline as giving a list of parts and chapters that could be read more or less independently and that are united by the loose structure that is found in the expository writing which is evident in Merleau-Ponty's own works, then it does seem that Sartre has little regard for the human body. But if we keep in mind Sartre's overall methodology, which is to reveal the degree to which the world is human because of the advent of bodily consciousness within matter, it becomes clear why and how the book pivots about the chapter on the body in part 3. In Sartre's methodology, a direct, expository approach to the body would effectively mean that we examine our body as an object» (Catalano, 2010, p. 32) see also Moran, 2010.

² Sartre's account of the body is subtle, complex, and many layered. It is of course largely a promissory note for a kind of larger study which he himself never carried out. However, his view of the body is a healthy corrective to what is often seen as his rather crude Cartesianism. Sartre is often criticized for his metaphysical claims concerning the gulf between different orders of being (for-itself and in-itself), but, in fact, his phenomenological account overcomes this crude 'Cartesian' dualism by stressing our embodied-being-in-the-world (Moran, 2011, p. 290).

1936/2012, 1939/2015). Moreover, despite the categorical refusal to identify consciousness with the brain³, it has never denied the close relationship between brain, body as whole and consciousness⁴. For these reasons – and as I will explain below – I consider Sartre’s phenomenological analysis of the body a crucial contribution to the understanding of consciousness as an “embodied-being-in-the-world” (Moran, 2011).

6.1.2 The Body Surpassed towards the World

I discussed at length the structure of consciousness as intentionality, both in the second and present chapters, and we can now say that consciousness, i.e. experience, is a revealing or disclosing activity (Rowlands, 2018) in the sense that the world is revealed to us by way of appearances through the body. Simply put, when I see a red tomato, a portion of the world is revealed or disclosed to me as red tomato (which is not me nor the yellow lemon near to it). This revelation is performed through the senses of my body, in this case the seeing act, and all the other bodily functions which contribute in being – and not merely have – a body, e.g. proprioception, vestibular sensation,

³ In *Sketch for a Theory of the Emotion*, Sartre exposed an instance of mereological fallacy *ante litteram*: «The theory of cerebral sensibility appeals to a cortical disturbance that is unverifiable. Sherrington made some experiments on dogs, and one can certainly praise his operational dexterity. But these experiments taken by themselves prove absolutely nothing. Simply because the head of a dog practically isolated from its body still gives signs of emotion, I cannot see that we have the right to conclude that the dog is feeling a complete emotion. Besides, even supposing that the existence of a corticothalamic sensitivity were established, it would still be necessary to ask the previous question: can a physiological disturbance, whatever it may be, render an account of the organized character of an emotion?» (Sartre, 2015, p. 34) Moreover, he was critical of the idea that the brain contained images and perceptions, like an empirical version of illusion of immanence: «The classical conceptions of aphasia and cerebral localization do not withstand criticism. The image memory is something other and more than a simple cerebral reviviscence. The brain could not have as a function the storing of images. Perception is a direct contact with the thing. [...] one speaks to us about images and ideas ‘in the brain’. And we do not know what that means. Does it have to do with a physiological hypothesis presenting the brain as an organ that functions like the heart or the liver in the unity of a biological synthesis? Does it have to do with a psychological theory concerning the indivisibility of the psychic state? Does it have to do with the two at the same time? And then to what extent does one negate the reality of the image? Must one understand that the image is a ‘metaphysical reality’, an ‘abstract’, the way the individual is for certain sociologists? In this case, one would have to understand simply that it has no functional reality, that it is never independent. » (Sartre, 2012, p. 72, 103).

⁴ Interestingly, in a late interview (1975) he has stated that the next step in the studies of consciousness could be an account of the relation between consciousness and the brain: «That is a problem I studied for several years at the time of *L’Être et le Neant* and afterward. I tried to deal with it, but I did not finish it. I was not interested in doing a study on the relations of consciousness and the brain, because I first wanted to define consciousness. That is enough for one man’s lifetime. I wished to define it as it presents itself to us, for you, for me. In so doing, I wanted to posit a definite object which others would then have to try to explain within a materialist system, that is, to study its relation to the brain» (Sartre, 1981, p. 40).

interoception, kinaesthesia, exteroception, and so on. So, if intentionality is a revealing act that puts in relation the human being with the world, this relation necessarily passes through the body. This is the point that will be discussed in this section, exposing also a comparison between Sartre's bodily consciousness and the core concepts of 4E Cognition.

First, we must clarify what this relation between consciousness/intentionality and the body means, because one could easily misunderstand it as a relation between the conscious subject – be it a mental substance or the brain – and the body-machine, where someone can control the body as a bridge to the world that otherwise would remain unreachable for the conscious subject alone⁵. As we have already seen, this position is not valid anymore for the purpose we have here, i.e. elaborate an alternative approach towards experience that is not based on reductive accounts of consciousness and body. Perhaps, the term “relation” might be misleading because it involves two different poles – in this case, consciousness and body – taken together by some kind of nexus, e.g. physical, psychophysical, phenomenal, and so on. In effect, this is not the case since for Sartre consciousness is nothing but bodily consciousness, i.e. there is no difference between the intentional acts and the processes of the body, so that «embodied consciousness is not to say that consciousness happens to ride around inside the body» (Cox, 2006, p. 54). On the other hand, it would be a mistake to confuse this intrinsic relation between body and consciousness with the identity of the two, as if Sartre wanted to pursue a behavioural approach towards consciousness, or with the unification of two different dimensions⁶.

For instance, Sartre consider the problem of the unification of mind and body a false problem, i.e. what philosophers of mind call the “mind-body problem”; Sartre,

⁵ Interestingly, this might seem a Cartesian-type paradigm of the body, moreover, as we have previously seen, even some neuroscientific approaches interpret the body as a machine controlled by the brain. However, Descartes in the *Discourse on the Method* already conceived the body and the mind as different but unified substances: «I had described after this the rational soul and shown [...] that it is not sufficient that it should be lodged in the human body like a pilot in his ship, unless perhaps for the moving of its members, but that it is necessary that it should also be joined and united more closely to the body in order to have sensations and appetites similar to our own, and thus to form a true man» (René Descartes, 2003, pp. 39–40).

⁶ «In one sense therefore the body is a necessary characteristic of the for-itself; it is not true that the body is the product of an arbitrary decision on the part of a demiurge nor that the union of soul and body is the contingent bringing together of two substances radically distinct. On the contrary, the very nature of the for-itself demands that it be body; that is, that, its nihilating escape from being should be made in the form of an engagement in the world» (Sartre, 1956, p. 309).

instead of proposing a solution to it or claiming that it cannot be solved, he argues for its dissolution. Indeed, he shows how the attempt to put together our embodied experience and our material body – or the existed body and the body seen as an object – is an absurd attempt to unify two ontologically, and not merely epistemically, different dimensions, i.e. the body-for-me and the body-for-the-others⁷. This does not mean that we have two substances, like for substantial dualisms, but that we are embodied and united from the start of our experience – necessarily – however, we cannot perceive this form of embodiment unless we completely change it from actual to objectual consciousness (see Rowlands), making connections between the body as lived and the body as a corpse. For the same reasons, Sartre rejects the cruciality of the double sensation – as exposed by Husserl and Merleau-Ponty – which implies the intertwining between the perceived and perceiving body parts, for example the touched and touching hands (Moran, 2010b, 2010a)⁸. Thus, when we speak of body-consciousness relation according to Sartre, this relation must be developed outside any physicalist or dualist paradigm, because body and consciousness are two different diffractions of the same phenomenon, where the body stands for the sign (the mean)

⁷ «The problem of the body and its relations with consciousness is often obscured by the fact that while the body is from the start posited as a certain thing having its own laws and capable of being defined from outside, consciousness is then reached by the type of inner intuition which is peculiar to it. [...] Actually if after grasping "my" consciousness in its absolute interiority and by a series of reflective acts, I then seek to unite it with a certain living object composed of a nervous system, a brain, glands, digestive, respiratory, and circulatory organs whose very matter is capable of being analyzed chemically into atoms of hydrogen, carbon, nitrogen, phosphorus, etc., then I am going to encounter insurmountable difficulties. But these difficulties all stem from the fact: that I try to unite my consciousness not with my body but with the body of others. In fact the body which I have just described is not my body such as it is for me. [...] Of course the physicians who have taken care of me, the surgeons who have operated on me, have been able to have direct experience with the body which I myself do not know. I do not disagree with them, I do not claim that I lack a brain, a heart, or a stomach. But it is most important to choose the order of our bits of knowledge. So far as the physicians have had any experience with my body, it was with my body in the midst of the world and as it is for others. My body as it is for me does not appear to me in the midst of the world. Of course during a radioscopy I was able to see the picture of my vertebrae on a screen, but I was outside in the midst of the world. I was apprehending a wholly constituted object as a this among other thises, and it was only by a reasoning process that I referred it back to being mine; it was much more my property than my being» (Sartre, 1956, p. 303).

⁸ «When a doctor takes my wounded leg and looks at it while I, half raised up on my bed, watch him do it, there is no essential difference between the visual perception which I have of the doctor's body and that which I have of my own leg. Better yet, they are distinguished only as different structures of a single global perception; there is no essential difference between the doctor's perception of my leg and my own present perception of it. Of course when I touch my leg with my finger, I realize that my leg is touched. But this phenomenon of double sensation is not essential: cold, a shot of morphine, can make it disappear. This shows that we are dealing with two essentially different orders of reality. To touch and to be touched, to feel that one is touching and to feel that one is touched-these are two species of phenomena which it is useless to try to reunite by the term "double sensation." In fact they are radically distinct, and they exist on two incommunicable levels» (Sartre, 1956, p. 304).

through which intentionality (the meaning) expresses itself. There is no meaning without the sign/mean but you must surpass the sign (which remains in the background) in order to see the meaning of it⁹.

This metaphor lead to another important point of the Sartrean phenomenology of the body, that is, the so-called lived body introduced by classical phenomenology (*le corps-existé* with the words of Sartre) is identical with the for-itself (consciousness), therefore is not something of which I constantly have consciousness, but it is “surpassed”, i.e. nihilated, towards the world, as the pre-reflective non-positional consciousness of the revealing intentionality to the world (see bodily awareness in chapter two). The body-for-itself, «is never a given which I can know. It is there everywhere as the surpassed; it exists only in so far as I escape it by nihilating myself» (Sartre, 1956, p. 309). Therefore, it is nothing but the unseen condition of possibility of being-in-the-world:

«Therefore my body is a conscious structure of my consciousness. But precisely because the body is the point of view on which there can not be a point of view, there is on the level of the unreflective consciousness no consciousness of the body. The body belongs then to the structures of the non-thetic self-consciousness. Yet can we identify it purely and simply with this non-thetic consciousness? That is not possible either, for non-thetic consciousness is self-consciousness as the free project toward a possibility which is its own; that is, in so far as it is the foundation of its own nothingness. Non-positional consciousness is consciousness (of the) body as being that which it surmounts and nihilates by making itself consciousness - i.e., as being something which consciousness is without having to be it and which it passes over in order to be what it has to be. In short, consciousness (of) the body is lateral and retrospective; the body is the neglected, the "passed by in silence." And yet the body is what this consciousness is; it is not even anything except body. The rest is nothingness and silence» (Sartre, 1956, p. 330).

⁹ «Consciousness of the body is comparable to the consciousness of a sign. The sign moreover is on the side of the body; it is one of the essential structures of the body. Now the consciousness of a sign exists, for otherwise we should not be able to understand its meaning. But the sign is that which is surpassed toward meaning, that which is neglected for the sake of the meaning, that which is never apprehended for itself, that beyond which the look is perpetually directed» (Sartre, 1956, p. 330).

Thus, the body is the orientation, the point of view and the permanent structure of the conscious being within the world¹⁰, in such a way that everything I perceive or even imagine is not somewhere without any reference to me. The world is not in front of me like an external and detached set of things, I am within it with a body that gives me the coordinates towards the things of the world, so that the red tomato is on my left, while the door is behind me¹¹.

6.1.2.1 Embodiment, Embedment, Extendedness, Enactment

This condition of being-in-the-world is not that consciousness is “engulfed” in a body which is in turn “engulfed” in the world (Sartre, 1981) like separate elements inserted one into the other, rather the body is the manifestation of consciousness and there is nothing more about conscious experience but its intrinsic relation with the world. The body provides the “situation” for the foundation of the possibilities for the for-itself (my possibilities in being-in-the-world, orientation and point of view). But, again, the body exists as surpassed and transcended in seeking to experience the world. This is similar to the concept of embedment used by the proponents of 4E cognition to describe the fact that our cognition is always “situated” and dependent on contingent factors from the environment.

Thus, as we have seen above, the body is not simply causally relevant together with the brain for cognitive and consciousness processes, as instead happens in the 4E’s embodiment (Wilson & Foglia, 2017). But we are not merely embodied, we are also embedded in the sense that our embodied being expresses itself through and

¹⁰ The body «[...] is therefore in no way a contingent addition to my soul; on the contrary it is a permanent structure of my being and a permanent condition of possibility for my consciousness as consciousness of the world» (Sartre, 1956, p. 328). For the impossibility of disembodiment see (Legrand, 2010).

¹¹ «But when we say that the for-itself is-in-the-world, that consciousness is consciousness of the world, we must be careful to remember that the world exists confronting consciousness as an indefinite multiplicity of reciprocal relations which consciousness surveys without perspective and contemplates without a point of view. For me this glass is to the left of the decanter and a little behind it; for Pierre, it is to the right and a little in front. It is not even conceivable that a consciousness could survey the world in such a way that the glass should be simultaneously given to it at the right and at the left of the decanter, in front of it and behind it» (Sartre, 1956, p. 306).

within environmental interactions, so that the body as point of view is “surpassed”, “transcended”, “passed by in silence” towards the world and upon which, in turn, I cannot have a point of view¹². In other words, when we are conscious of the red tomato, we can see it, touch it, taste it, weight it with the hands, etc.; during these conscious acts of intentionality, at no time the body is revealed, the only intentional objects is the red tomato, we do not see our eyes seeing, our fingers touching, our sense of taste tasting. When I weight the tomato, I feel nothing but the tomato’s weight, in the same way as when I write my hand vanishes behind the words I am writing, «it is lost in the complex system of instrumentality in order that this system may exist. It is simply the meaning and the orientation of the system» (Sartre, 1956, p. 323). Of course, this does not mean that we cannot perceive our own body (see chapter two), indeed while I am weighing the tomato I can concentrate at some point on my hand holding the tomato, but this does not mean that I am perceiving my lived body. At that specific moment my consciousness is positing my hand from the point of view of the other as psychic object (Sartre, 1956, p. 347), as a body part among other’s body parts, as a tool among a world of tools¹³. As Moran pointed out, «[f]or Sartre, I cannot see the sensitivity of the hand or even the ‘mineness’ of my hand» (Moran, 2011, p. 275), because the subjectivity of my body is pre-reflectively and non-positionally given, as I explained in the previous paragraphs. For example, while weighing the tomato I am not feeling the effort of my hand and arm (skin, bones, tendons, muscles, etc.) during that action, I am precisely feeling the weight of the tomato, i.e. the resistance of a thing against me. We always perceive the things in the world (their resistance, hardness, softness, etc. Sartre, 1956, p. 304)¹⁴, not the lived body as such, which instead is invisible

¹² See also the difference between body schema and body image exposed in the second chapter.

¹³ «The Other's body appears to me here as one instrument in the midst of other instruments, not only as a tool to make tools but also as a tool to manage tools, in a word as a tool-machine. If I interpret the role of my body in relation to my action, in the light of the knowledge I have gained of the Other's body, I shall then consider myself as disposing of a certain instrument which I can dispose of at my whim and which in turn will dispose of other instruments all functioning toward a certain end which I pursue» (Sartre, 1956, p. 320).

¹⁴ «The body is lived and not known. This explains why the famous “sensation of effort” by which Maine de Biran attempted to reply to Hume's challenge is a psychological myth. We never have any sensation of our effort, but neither do we have peripheral sensations from the muscles, bones, tendons, or skin, which have been suggested to replace the sensation of effort. We perceive the resistance of things. What I perceive when I want to lift this glass to my mouth is not my effort but the heaviness of the glass-that is, its resistance to entering into an instrumental complex which I have made appear in the world» (Sartre, 1956, p. 324; see also Peckitt, 2010).

because integrated within the world (embedding)¹⁵ and it can be experienced as an object only in specific situations, like reflection, i.e. when we posit ourselves as intentional content, or illness and other psychosomatic situations, i.e. when the body appears as an obstacle that influences our being-in-the-world condition (Costa & Cesana, 2019; Leder, 1990) or when I am looked by the other and «I exist for myself as a body known by the Other» (Sartre, 1956, p. 351)¹⁶.

The Sartrean description of the body can be also seen through the concepts of enactment and extendedness. Enactment means that the cognitive processes can be understood as dynamical interaction between the embodied mind and various affordances given by the environment. There are three principal ways of interpreting enactivism (D. Ward, Silverman, & Villalobos, 2017). The first one is the sensorimotor approach (Engel et al., 2013; Noë, 2004; O'Regan & Noë, 2001) according to which consciousness and cognition – but mostly cognition – are not something that human beings “have” but “do” through actions within the world. According to the stronger version, the actual moving body is essential for experience (Noë, 2004), while the weaker version argues that bodily movements as such are not necessary for the cognitive process, but the corresponding skills and practical knowledges a person has built up in the past, which are enabling condition for performing a movement (Shapiro, 2011). The second version is the autopoietic enactive approach (Di Paolo & Thompson, 2014; Thompson, 2007; F. J. Varela, 1997), interprets the body not simply in a neuromuscular sense, but as a system with specific abilities to regulate its interaction with the world (Di Paolo & Thompson, 2014). In this sense, cognition is understood in a broader way, as sense-making process of an autonomous system that tries to preserve itself as a precarious form of life within a world full of potential interactions. The body, therefore, is essential as the bearer of interaction with the world in order to regulate the whole cognitive system as an autonomous sense-maker. The third version, called radical enactive cognition – REC – (Hutto, 2017; Hutto & Myin, 2013, 2017), analyses the mind in terms of dynamic patterns of adaptive environmental

¹⁵ «This is the body experienced not as a thing but as a ‘non-thing’, almost transparent medium for my experience of the world. In this regard my body is surpassed towards the world. [...] This non-apprehended body swims in the world, as it were, unnoticed» (Moran, 2011, p. 271).

¹⁶ «This intuitively felt body becomes obtrusive in illness (I become dizzy and lose my balance), failure (the stone is too heavy to lift), disability (I cannot feel my leg), or in psychosomatic conditions (in anorexia nervosa I experience my body as too gross), or, as Sartre emphasises with great force and originality, in the look of the other» (Moran, 2011, pp. 272–273).

interactions, trying to pursue an even stronger anti-cognitivism and anti-representationalism, rejecting the sense-making processes that “insert” meaning and significance in the organism/environment relationship and proposing a more basic biodynamic relation between organism and environment without any propositionally specifiable content (D. Ward et al., 2017) .

Sartre presents elements that can be compatible with the enactivist approaches (Wider, 2016). As said, Sartre anti-representational and anti-internalistic conception of consciousness can be compared to the “structural coupling” of organism and environment exposed by the sensorimotor approach (O'Regan & Noë, 2001; F. J. Varela et al., 1991), which argues for a close relation between perception and action (Noë, 2004), as Sartre also did in *Being and Nothingness* (Campos, 2017)¹⁷. Regarding REC, while sharing the strongly anti-representational and contentless account of cognitive processes, Sartre's position, as already seen, focuses on the phenomenological and subjective dimension of corporeity and its relationship with the world, which is not taken into consideration into the (merely) biological and physical relationships between organism and environment proposed by radical enactivism. Also, for Sartre every conscious act towards the world neither is constituted by a passive reception of external stimuli (as enactivism in general argues) nor only by physical interactions, but also by a revealing act of meanings within an organized world; for example, in my daily experience nails and hammers are related from the start, because my consciousness is a pure and contemplative act of knowledge of colours, shapes, textures etc., but it makes linkages and references between things, my and other's body and their functions in the world¹⁸. In this sense, we may find some

¹⁷ «Thus the world from the moment of the upsurge of my For-itself is revealed as the indication of acts to be performed; these acts refer to other acts, and those to others, and so on. It is to be noted however that if from this point of view perception and action are indistinguishable, action is nevertheless presented as a future efficacy which surpasses and transcends the pure and simple perceived. [...] Perception is naturally surpassed toward action; better yet, it can be revealed only in and through projects of action. The world is revealed as an “always future hollow,” for we are always future to ourselves» (Sartre, 1956, p. 322).

¹⁸ «Therefore if I conceive of my body in the image of the Other's body, it is an instrument in the world which I must handle delicately and which is like a key to the handling of other tools. [...] Objects are revealed to us at the heart of a complex of instrumentality in which they occupy a determined place. This place is not defined by pure spatial co-ordinates but in relation to axes of practical reference. “The glass is on the coffee table;” this means that we must be careful not to upset the glass if we move the table. The package of tobacco is on the mantle piece; this means that we must clear a distance of three yards if we want to go from the pipe to the tobacco while avoiding certain obstacles—end tables, footstools, etc.—which are placed between the mantle piece and the table. In this sense perception is in no way to be distinguished from the practical organization of existents into a world. Each instrument

similarities with the idea of sense-making cognition as proposed by the autopoietic enactive approach, but I think this can be seen only as a weak suggestion. Indeed, the concept of sense-making claimed by the proponents of the enactive approach refers to the mind/life continuity produced by interactions between living systems and environment (at different levels: unicellular, animal, collective, socio-technical; Barandiaran, 2017; Di Paolo & Thompson, 2014), while Sartre meant a pragmatic relation between human experience and the environment as “human reality”, i.e. a “world” as constituted by meanings made for and by humans¹⁹.

This last point allows me to consider the concept of extendedness in relation to Sartre. The Extended Mind Thesis was introduced by Clark and Chalmers (Clark & Chalmers, 1998; Menary, 2010b), and claims that the separation between brain, body and environment is just an assumption, and that there is no reason to confine our cognition within the skin and skull (Menary, 2006). The mind cannot be described only by means of neural activity, rather our mental processes are distributed also in the environment using instruments. In this way, human cognition finds support in the external world through tools that become part of our memory, of our computing capacity (e.g. notebook, paper and pencil, smartphone), of our communication capacity (e.g. computers and brain-computer interfaces), and so on. In *Being and Nothingness* there are elements that show how consciousness transcends itself towards the world not only through the body, but also through the adaptation to various instruments:

«My body is everywhere: the bomb which destroys my house also damages my body in so far as the house was already an indication of my body. This is why my body always extends across the tool which it utilizes: it is at the end of the cane

refers to other instruments, to those which are its keys and to those for which it is the key. But these references could not be grasped by a purely contemplative consciousness. For such a consciousness the hammer would not refer to the nails but would be alongside them; furthermore the expression “alongside” loses all meaning if it does not outline a path which goes from the hammer to the nail and which must be cleared. The space which is originally revealed to me is hodological space; it is furrowed with paths and highways; it is instrumental and it is the location of tools» (Sartre, 1956, pp. 321–322).

¹⁹ «We know that for Heidegger the being of human reality is defined as “being-in-the-world.” The world is a synthetic complex of instrumental realities inasmuch as they point one to another in ever widening circles, and inasmuch as man makes himself known in terms of this complex which he is. This means both that “human reality” springs forth invested with being and “finds itself” (*sich befinden*) in being-and also that human reality causes being, which surrounds it, to be disposed around human reality in the form of the world» (Sartre, 1956, p. 17).

on which I lean and against the earth; it is at the end of the telescope which shows me the stars; it is on the chair, in the whole house; for it is my adaptation to these tools» (Sartre, 1956, p. 325).

However, the way in which Sartre consider the extension of the body through the tools is different from the concept of extendedness proposed here above. According to the 4E Cognition, the mind extends towards the instrument in a functional way, that is, extra-bodily mechanisms that can replicate or improve the abilities and the functions of the body can be considered extension of mental cognition. In this sense, the mind is only contingently embodied and extended through the body, so that we can conceive a progressive substitution of the body with extra-bodily tools and parts, because here the extendedness is merely functional. According to Sartre, instead, the fact that our “body is everywhere” maintains a phenomenological and pragmatic meaning, in the sense that consciousness – expressed as bodily consciousness – finds itself in strict correlation with the world and every single thing in the environment can be perceived as a bridge (or as an affordance; Gibson, 1979) for performing this world-body connection. Moreover, according to the extended mind thesis, our mental processes goes from the brain-body towards the environment by means of instruments, instead for Sartre is quite the opposite, that is, it is the world that presents occasions of extension to the human body-consciousness which is always and already embodied, situated and extended; and this extension can be either a sort of body enhancement («[...] glasses, pince-nez, monocles, etc., which become, so to speak, a supplementary sense organ» Sartre, 1956, p. 329) or a danger to the body itself («Thus my body is indicated originally by instrumental complexes and secondarily by destructive devices. I live my body in danger as regards menacing machines as for manageable instruments» Sartre, 1956, p. 325).

Taken together, Sartre’s account of consciousness, using the 4E Cognition’s vocabulary, intrinsically implies embodiment – not only as a causally relevant element – and embedment – our experience is always related to a contingent world-situation – also it shows elements of enactment – perception and action are two sides of a coin – and of extendedness – understood differently from the extended mind thesis though. In particular, the last points indicate how the fact that our experience is always

embodied is not the centre of the Sartrean analysis, because embodiment, embedment, enactment, and extendedness exist as a function of the relation with the world.

6.1.3 The Primacy of the World

So far, I have presented Sartre's theory of the body and I showed how consciousness and body are intrinsically related, so that we may say that there is no essential difference between body and consciousness for the French phenomenologist (Sartre & Contat, 1975). However, I also pointed out that in most of our existence our body is nothing but transcendent towards the world, so that the body exist as the form through which we are in the world. Now, it is perhaps easier to understand why the body, despite being a central theme within the Sartrean thought, is found in *Being and Nothingness* only after the chapters dedicated to the transcendence and the existence of others. While today's theories of embodiment generally consider the body and bodily intentionality as a movement that goes from the brain to the body and towards the world, for Sartre instead intentionality is mainly a revelation of the world that manifests itself to us and that at the same time shows us as a embodied being: «We must note in addition that this appearance of the body does not give us the body as it acts and perceives but only as it is acted on and perceived» (Sartre, 1956, p. 358). The world is certainly manifested thanks to the body, so that we can perceive it, feel it, observing it, etc., but the body as perceived is a late result of reflection after that the world has "indicated" our body with things and instruments (Sartre, 1956, pp. 319–320)²⁰. We are constantly immersed in the world and in this way we exist our body as intentionality towards the world, instead, every time we focus specifically on our body, we objectify it and we abstract a little from the surrounding world but we only be able

²⁰ «Thus any sense in so far as it is-for-me is an inapprehensible; it is not the infinite collection of my sensations.since I never encounter anything but objects in the world. On the other hand if I assume a reflective point of view on my consciousness, I shall encounter my consciousness of this or that thing-in-the-world, not my visual or tactile sense; finally if I can see or touch my sense organs, I have the revelation of pure objects in the world, not of a revealing or constructive activity. Nevertheless the senses are there. There is sight, touch, hearing» (Sartre, 1956, p. 316). In the words of Catalano: «From a phenomenological perspective, the human body as such is a late appearance, and this claim is true even of the lived body. To repeat again, in our engaged activities we pass through our bodies. By using what he calls 'pure reflection', the phenomenological task is to move the lived body out of the background into our philosophic awareness. Thus, once the body has been separated from the world, the phenomenological task is to describe how we proceed from a lived awareness of the body to a conceptual understanding of it» (Catalano, 2010, p. 34).

to see another object of the world, i.e. our body²¹. In this sense, perception is not given before or after the perceived objects, but perception and object perceived are contemporaneous, and bodily consciousness reveals the objective characteristics of the world²²:

«A sense is not given before sensible objects. [...] Neither is it given after sensible objects; for in that case it would be necessary to suppose a world of incommunicable images, simple copies of reality the mechanism of whose appearance was inconceivable. The senses are contemporaneous with objects; they are things "in person" as they are revealed to us in perspective. They represent simply an objective rule of this revelation. Thus sight does not produce visual sensations; neither is it affected by light rays. It is the collection of all visible objects in so far as their objective and reciprocal relations all refer to certain chosen sizes-submitted to all at once-as measures, and to a certain center of perspective. From this point of view the senses must in no way be identified with subjectivity. In fact all variations which can be registered in a perceptive field are objective variations» (Sartre, 1956, p. 318).

Thus, we may say that according to Sartre the world comes first with perception, where the latter has nothing to do with subjective and mental states like qualia²³, rather we must say that «[s]ensation and perception are not first ways of getting to know about the world or the existence of others; [...] they are that through which we are a being-in-the-world. The realization that our consciousness is in the form of a body with senses is simultaneously the awareness that our world and our presence in it are situational» (Catalano, 2010, p. 35). There is first a world that is not the environment

²¹ «[...] it is necessary that I lose myself in the world in order for the world to exist and for me to be able to transcend it. Thus to say that I have entered into the world, "come to the world," or that there is a world, or that I have a body is one and the same thing. In this sense my body is everywhere-in the world» (Sartre, 1956, p. 318).

²² Our body is first out in the world; it is in things having color, sound, textures, and as being here rather than there. Because of the depth of the body, things have depth. It is the body that establishes the condition for the possibility of perception, and it is the body that gives depth to the horizon of perception. Phenomenologically, we move from the body as it differentiates matter into a world to other people as bodies that in turn reveal our body to us as the lived center of relations (Catalano, 2010, pp. 32–33).

²³ «Sensation is, for Sartre, primarily a mode of the body's being-in-the world, a mode that opens these doors into reality rather than others, for example, the door into color. By reducing this selectivity to an intellectualist position, Merleau-Ponty is able to dismiss it with the neat objection that all constituting makes the object constituted lucid, and thus all depth, surdness and mystery of the world is removed. But if the relation of consciousness to the world is that of matter to matter, that is, the matter of an active and organic consciousness delineating its world, then Merleau-Ponty's objections are not to the point, or, at least, they are not to the point in the way he is making them» (Catalano, 2010, p. 28).

first detached from the human being and then linked to it somehow, rather a world that is originally relation to us in the form of beings-in-the-world. In this sense, the relationship comes before the poles themselves (Costa & Cesana, 2019, p. 38). The body is “existed” only as connected and directed towards a world to experience and the world in turn manifests itself only to the senses of the body. Moreover, contrary to common sense, it is not the senses that reveal the body, but it is the world as transcendence that indicates the body and gives us the bounders of our existence:

«This [inapprehensible] given which I am without having to be it – except in the mode of non-being-this I can neither grasp nor know, for it is everywhere recovered and surpassed, utilized for my assumed projects. On the other hand everything indicates it to me, every transcendent outlines it in a sort of hollow by its very transcendence without my ever being able to turn back on that which it indicates since I am the being indicated» (Sartre, 1956, p. 327).

Here we can see an interesting difference from Merleau-Ponty, since for Sartre the body-for-itself is indicated as an item within the structure of our being in the world, i.e. condition of possibility of the revealing act, but, at the same time, becoming a situated element which follows the rules of the world and which is never a pure and absolute centre of reference (Sartre, 1956, p. 327)²⁴. In other words, if for Merleau-Ponty the body is phenomenologically prior to the system of the world and represents the ability of surpassing towards the world («I am a body that rises up toward the world» (Merleau-Ponty, 1845/2013, p. 78)), for Sartre the body is the “perpetually surpassed” towards the world (P. Hoffman, 2006)²⁵:

²⁴ «In both perception and action, then, the body is indicated, and defined, as an element of an objective system. And just as our perception of distance turned out to be determined (albeit not causally) in conformity with the objective laws of perspective, so too our bodily actions are viewed as either appropriate or inappropriate to the instruments and tasks at hand due to our understanding of the strictly objective relations between instruments and our active bodily members. Nothing could be further from Merleau-Ponty’s view, where the body’s perceptual and practical assessments of what it copes with are prior to, and independent of, our understanding of any such objective system. To avoid possible confusion we must stress that for Sartre this objective system of which my body is an (emptily indicated) element is never displayed for me from the vantage point of “absolute objectivity” or “pure rationality” (Sartre 1956: 319)» (P. Hoffman, 2006, p. 259).

²⁵ «The body is necessary again as the obstacle to be surpassed in order to be in the world; that is, the obstacle which I am to myself. In this sense it is not different from the absolute order of the world, this order which I cause to arrive in being by surpassing it toward a being-to-come, toward being-beyond-being» (Sartre, 1956, p. 326).

«Far from the body being first for us and revealing things to us, it is the instrumental-things which in their original appearance indicate our body to us. The body is not a screen between things and ourselves; it manifests only the individuality and the contingency of our original relation to instrumental-things. [...] And if I am an instrument in the world, this is because I have caused instruments in general to-be-there by the projection of myself toward my possibles. It is only in a world that there can be a body, and a primary relation is indispensable in order that this world may exist. In another sense I am separated from it by the infinite density of the world; it is given to me by a reflux of the world toward my facticity, and the condition of this reflux of the world toward my facticity is a perpetual surpassing» (Sartre, 1956, pp. 325–326).

6.2 The Worldhood Paradigm

So far, after having discussed in the previous chapter the concept of intentionality, phenomenality and self-awareness, I have tried to make the issue more concrete, arguing for a theory of embodied consciousness as revealing act of a world, i.e. as the condition of embodied-being-in-the-world, in order to move away from both the mentalistic/mento-centric perspective of experience as mind-related and the brainhood/neuro-centric perspective of experience as brain-related. In addition, I showed how consciousness, despite being intrinsically embodied, is nothing if it is not directed towards a world and incorporated and embedded into it. Thus, I may say that this account of consciousness as embodied-being-in-the-world moves away also from potential body-centric perspective, which could understand corporeity as the main and first object and relation of our experience, consequently losing sight of the importance of the body as the way in virtue of which we experience the world²⁶. It is the world that indicates me as an embodied conscious agent, not vice versa, and the embodied conscious agent make the world appear²⁷, in the sense that the world must be conceived

²⁶ As for a blind person experience is given by means of the cane as the object “in virtue of which” the person is conscious of the world and not “as the object of” the experience (Rowlands, 2010b), similarly experience is given to us by means of the body as the object “in virtue of which” we are conscious of the world and not “as the object of” the experience.

²⁷ «We know that for Heidegger the being of human reality is defined as “being-in-the-world:” The world is a synthetic complex of instrumental realities inasmuch as they point one to another in ever widening circles, and inasmuch as man makes himself known in terms of this complex which he is. This means both that “human reality” springs forth invested with being and “finds itself” (*sich befinden*) in

not as an object created by the mind, the brain or even the body, but as the ontological condition of consciousness²⁸.

Thus, consciousness, i.e. experience of something, needs to be-in-the-world, so that there is no consciousness without a world, where with “world” I do not mean only the surrounding environment, but the pre-intentional and pre-given lifeworld as the condition of both the existential world of meanings and instruments and the physical world of atoms and molecules. Seen in these terms, in order to provide a possibly complete account of consciousness, we should ground our researches on a – I would call it – “worldhood” paradigm, rather than merely on brainhood, corporeity, or mento-centrism, as these are abstractions derived from the being-in-the-world condition. It is not possible to conceive a conscious being (*Dasein*, in Heideggerian terms) as a worldless subject taken from the outside, since consciousness is transcendence towards a world that is already given and into which we already find ourselves²⁹, therefore, the way in which the condition of being-in-the-world is structured intrinsically influences the way in which we experience.

If we dismiss the perspective from the brain or the mind or the body, and we take a broader view from the world, we may start understanding consciousness as intrinsically world related. In this sense, in order to understand how we experience something, we do not have to focus merely on the brain activity nor on the mental introspection, nor even on the body modifications, rather we may analyse how our relation to the world is structure. We need to be careful here, because taking a world-

being-and also that human reality causes being, which surrounds it, to be disposed around human reality in the form of the world» (Sartre, 1956, p. 17).

But human reality can make being appear as organized totality in the world only by surpassing being.

²⁸ «Consciousness is consciousness of something. This means that transcendence is the constitutive structure of consciousness; that is, that consciousness is born supported by a being which is not itself. This is what we call the ontological proof» (Sartre, 1956, p. lxiii).

²⁹ «Experiential life is, as Heidegger says, as such world-related, it is always already living in the world; it does not have to seek it out. Life-experience is literally speaking “worldly tuned,” it always lives in a world, it is properly speaking a worldlife, and it always finds itself in a lifeworld. These ideas are forcefully articulated in the following quote from *Being and Time* (1927): “When *Dasein* directs itself towards something and grasps it, it does not somehow first get out of an inner sphere in which it has been proximally encapsulated, but its primary kind of Being is such that it is always ‘outside’ alongside entities which it encounters and which belong to a world already discovered. Nor is any inner sphere abandoned when *Dasein* dwells alongside the entity to be known, and determines its character; but even in this ‘Being-outside’ alongside the object, *Dasein* is still ‘inside,’ if we understand this in the correct sense; that is to say, it is itself ‘inside’ as a Being-in-the-world which knows” (Heidegger, 1996, p. 58)» (Zahavi, 2019b, p. 116).

based approach does not mean that we refuse the contribution from brain sciences, philosophy of mind, 4E cognition, etc. – quite the opposite. A world-based approach means that we should take into consideration all these layers of knowledge and their respective regional ontologies, trying to put together within a broader perspective. This means that the analysis of the concept of consciousness must be observed within the context of our immersion and projection into the world, and only through the description of brain activations or body modifications and behaviour. So far, I have discussed the concepts of 4E cognition, and I explained how a Sartrean approach can be useful for immersing again hypostatized mental concepts into lived experience (chapter 4) and for relocating the body within the world (chapter 6). We will see later how this can work also with neuroscience, for now, I shall look in detail how different forms of consciousness can be understood as different ways of embodied-being-in-the-world.

6.2.1 Being-in-the-World as Spatio-temporality

I have exposed how the body is crucial for our experience, in the way that consciousness is impossible without it insofar the body determinates the way in virtue of which we experience the world. This means that the body is a relational structure and its modification is here conceived not merely as the cause of a change of the way we perceive our body, but primarily as an alteration of the relation with the world (Costa & Cesana, 2019). Indeed, the world is an organized system of possibilities for me as a bodily consciousness and the realization of these possibilities depends on the kind of relation I have with the world. We always are in situation, i.e. embodied and embedded in the world, we are not isolated heads or isolated body in front of a world, so that a damage to the body must be understood as a modification of this situation, as we will see later.

It may be noted that our path has been developed as a movement from what is more conceptual and structural, e.g. intentionality, phenomenality, self-awareness, etc. towards, from time to time, more concrete forms of experience, e.g. the body, consciousness of the world, etc. The next step will be to identify the matrix of these concrete forms of experience, that is, to understand how the experience as a

relationship with the world is given. We have already seen that property-based and content-based theories do not seem to account for experience as intrinsically the relationship between body and world, indeed, for example, arguing that consciousness emerges from integrating information in the brain does not explain anything about how we experience the world, at most it can explain how a part of our co-relation with the world is formed relative to the functioning of the brain. Thus, in the same way as the broken leg example, we should conceive the brain activity not as the origin and container of experience (brainhood), but as one of the key points of our relationship with the world.

Thus, so far, I have argued that we need to focus on the relationship rather than on the poles of the relation alone, i.e. world and body, including the brain. Now, the aim is to understand what this relationship is made of, concretely. We need therefore to step a foot outside the pure phenomenological and conceptual domain, but without completely leaving it. I argued that this relationship intrinsically is the activity of consciousness, it is not some kind of property or feature, is the very same activity of consciousness as intentionality towards the world. At the same time, we have seen that this intentionality is not possible without an already pre-given world as a phenomenological and concrete condition of embodied-being-in-the-world. So far so good. But now, what are intentionality and embodied-being-in-the-world made of? Cognitivist and representationalist approaches fail to explain how we engage with the environment because they need to add intermediate entities (ideas, representations, symbol manipulation, etc.) for connecting us to the world. As I briefly pointed out above for 4E cognition and neuropragmatism, these basic structures of embedment and situatedness are grounded on pragmatic sensorimotor loops, embodied and extended cognition, action-based relationships between environmental affordances and bodily skills, adaptive and meaning-producing abilities. As I tried to show some of these proposals can be endorsed, but we need to go deeper and identify a common matrix. Once again, the phenomenological analysis can be useful, before proceeding towards neuroscientific support.

6.2.1.1 Temporality and Spatiality

Sartre, in his phenomenological path that goes from the most general to the most concrete, when he needs to define in a practical way how consciousness develops in relation to the objects of the world, he follows and reinterprets the teachings of Husserl and Heidegger concerning the fact that consciousness unfolds essentially in a temporal and spatial sense. Here, with time and space I do not mean anything like the physical time and space as building blocks of nature, the homogeneous and pure extension of Cartesian *res extensa*, nor merely the spatio-temporal coordinates within which the objects of the world are situated (“the chair is next to the table”). A rock is in-the-midst-of-the-world, i.e. an inert presence as a passive object among other object, while a consciousness is in-the-world which implies an involvement with a world that has to be engaged, revealed and differentiated³⁰. To avoid confusion, it would therefore be better to talk about temporality and spatiality. I already exposed how spatiality is rendered through the body, so not as an abstract, centreless, independent space (“world without men”; Sartre, 1956, p. 307) but as hodological space, i.e. a lived situation as a field of forces of the experiencers in relation with the world³¹. Similarly, consciousness is not simply given in time like a thing, but is temporalization, since our experience is the act of relate things together with us through temporality, which is not an universal, Newtonian container, nor an extrinsic law of nature, but the way in which a bodily consciousness relates with the things. In other words, before cognitively manipulating information and symbols, before mentalizing the external stimuli in representations, even before bodily interacting with the environment through action-perceptions, the world is given to experience through the temporalization and spatialization of the things in relation with us³². Metaphorically speaking, to keep the rhythm of a music with your hands, the foot or the whole body like in dance, you do

³⁰ «I can not say either that I am here or that I am not here, in the sense that we say “that box of matches is on the table;” this would be to confuse my “being-in-the-world” with a “being-in the midst of the world.” [...] People consider that we exist first as contemporary with this chair or this table, and they work out the meaning of the temporal by means of the world. But if we thus place ourselves in the midst of the world, we lose all possibility of distinguishing what no longer is from what is not» (Sartre, 1956, pp. 60, 110).

³¹ «For human reality, to be is to-be-there; that is, “there in that chair,” “there at that table,” “there at the top of that mountain, with these dimensions, this orientation, etc.” It is an ontological necessity» (Sartre, 1956, p. 308). The fact that our spatial experience is constructed through continuous interrelations between the potential movements of the body and the affordances offered by the environment is also supported by neuroscientific findings (Djebbara et al., 2019).

³² For example, «[k]inesthesia, as a movement, presupposes time, so that changes within the visual field are also experienced as temporal changes: what was there before is now in another place, and what is now found here is found first there and it would be in that other position if I moved in a certain direction» (Costa & Cesana, 2019, p. 226; my translation).

not need to cognitively count the timing (“one, two, three, four, one...”), in fact it is usually the movement of the body that helps to follow the music; however, even before keeping up the pace with the body you must feel the beat first, that is, your body (not just your ears) must be receptive to the sound that surrounds you and must tune in and synchronize with the music. As we shall see, this attunement is spatio-temporally related, because it implies the alignment between the environmental stimuli, the body (e.g. the heart rhythm) and the brain frequencies.

It is not possible to think about a conscious activity without temporality and spatiality, it is an ontological necessity. Temporality and spatiality are different from the psychic temporality and spatiality, i.e. the products of impure reflection where time and space are objectified in the form of duration³³ and three-dimensional space, so they have nothing to do with our cognitive ability to think about the things in the quantitative timeline and within spatial context³⁴. “Hodological space” is a pre-reflective way of transcending towards the world relative to our own centre of reference (“the glass on my left”), while “original temporality” is the pre-reflective mode of intentionality in which we unify all the different appearances of things not into a series of “now” and “here” (psychic temporality and spatiality) but into an “organized structure” of an “original synthesis” of past, present and future (Sartre, 1956, p. 107)³⁵. Through this original synthesis, in which the body as the role of necessary centre of reference, the objects of experience are temporally and spatially

³³ «Thus psychic temporality is an inert *datum*, closely akin to Bergson's duration, which undergoes its intimate cohesion without effecting it, which is perpetually temporalized without temporalizing itself, in which the irrational and magic interpenetration of elements that are not united by an ekstastic relation of being can be compared only to sympathetic magic acting from a distance-an interpenetration which hides a multiplicity of already formed “nows.” These characteristics do not result from any error on the part of psychologists or from a lack of knowledge; they are constitutive of psychic temporality, which is the hypostasis of original temporality» (Sartre, 1956, p. 167).

³⁴ «Thus Temporality is not a universal time containing all beings and in particular human realities. Neither is it a law of development which is imposed on being from without. Nor is it being. But it is the intra-structure of the being which is its own nihilation-that is, the mode of being peculiar to being-for-itself. The For-itself is the being which has to be its being in the diasporatic form of Temporality» (Sartre, 1956, p. 142).

³⁵ «Here we are then in the presence of two temporalities: the original temporality of which we are the temporalization, and psychic temporality which simultaneously appears as incompatible with the mode of being of our being, and as an inter-subjective reality, the object of science, the goal of human acts (in the sense, for example, that I do everything possible to “make Annie love me,” to “endow her with love for me”). This psychic temporality, which is evidently derived, can not stem directly from original temporality; the latter constitutes nothing other than itself. As for psychic temporality, it is incapable of constituting itself, for it is only a successive order of facts. Moreover psychic temporality could not appear to the unreflective for-itself, which is pure ekstastic presence to the world. Psychic temporality reveals itself to reflection, and reflection must constitute it» (Sartre, 1956, p. 159).

extended and embedded as into a wider spatio-temporal field³⁶. This synthesis is not a container of past, present and future contents of consciousness (No Content Thesis), rather a mode in which consciousness makes manifest itself and the world («we are the temporalization»; Sartre, 1956, p. 159)³⁷. Taken together, our embodied-being-in-the-world is every time given according to spatial and temporal modes. Thus, we always temporalize and spatialize the things we apprehend from the world and ourselves as situated in the world. This generally happens in a pre-reflectively way, so that we do not need to reflect upon our experience in order to put them in time and space, the world of experience is intrinsically temporalized and spatialized.

6.2.1.2 Brainhood versus Worldhood on the interpretation of experience

One way to highlight how we experience the lived time and space is by analysing cases in which this space-time relationship with the world becomes problematic or fails. Now, I shall go through some instances of experience that can be described as alterations of the spatio-temporal relation to the world. I will start with some examples where it is clear to understand the alteration, then I will focus on other example, in which the spatio-temporal interpretation is less manifested and can be misunderstood. But first, it is necessary to make a distinction between “malaise” (*malessere*) and “morbid form” (*forma morbosa*) (Costa & Cesana, 2019, p. 112). Malaise is an existential condition and refers to a disorder of the being-in-the-world, for which the subject suffers from an alteration of his relationship with the world. For example, an illness or body injury could deprive the subject of many possibilities of

³⁶ This point is not only argued by Sartre, but is also an essential key core of other phenomenologists like Husserl and Merleau-Ponty. «Temporally-extended objects can never appear if lived purely in the present, in the ‘here and now’ so to speak. Rather, this experience of here and now, which Husserl calls ‘primal impression’ (e.g. as expressed in the tone of a melody), must be embedded in a temporal horizon or field. This implies that every incoming tone is retained in consciousness (as retention x1 of the formerly primary impression of the tone x), and thereby gradually modified: with every new tone, the retention of the former tone sinks deeper in the background of consciousness, losing its liveliness and clarity; the former perceived tone thus gets part of a continuity of momentary gradated retentions. In turn, every incoming tone must also be part of a horizon of upcoming or anticipated tones, a horizon of the protention of “constantly gradated coming”» (Wehrle, 2019)

³⁷ In other words, the for-itself exists as the result of the past-self surpassed by the for-itself as such. Also, the for-itself is going toward its possible future-self, which continuously goes beyond present as ex-future.

action in the world and modify her/his relationship with the future (e.g. degenerative diseases). The morbid form, on the other hand, affects the empirical conditions that allow the manifestation of a being-in-the-world and it concerns the objective body and its characteristics in relation to the world. For example, a person in vegetative state has no form of malaise, because the morbid form has directly deprived her/him of her/his being-in-the-world. «In this sense, one can have a morbid form without having a malaise and, *vice versa*, one can have a serious morbid form (for example a tetraplegia) which however does not generate a malaise, as the subject adapts, that is, changes the order of her/his wishes» (Costa & Cesana, 2019, p. 113). I will go through both kind of alteration/disruption of the embodied-being-in-the-world condition.

I would start examining some cases in which our space-time relationship is altered due to sensorimotor disorders that could generate malaise. To give a first simple example, if I break my leg and now it is in a cast, my perception of the house stairs will necessarily change from a spatial and temporal point of view, the stairs will immediately manifest itself as difficult to scale, or as a spatially obstructive and temporally demanding obstacle³⁸. Temporality and spatiality here do not refer to the geometric and temporal coordinates of the stairs compared to other objects in the world, but to our bodily experience of the world. In this case the “background disappearance” stops and the body becomes cumbersome, opaque (not transparent anymore) and puts into light the relation with the environment that usually we do not notice (Leder, 1990). Then, the meaning of the stairs is not a property attached to the object but is the sense of the relationship I entertain with it. When a stairlift is installed, my situation could change, and the stairs would no longer be perceived – or perceived in a lesser way – as a spatiotemporal obstructive object, since my consciousness of my bedroom-to-be-reached is easily available again. So, this is not just a question of

³⁸ «Here I am at the foot of this crag which appears to me as “not scalable.” This means that the rock appears to me in the light of a projected scaling-a secondary project which finds its meaning in terms of an initial project which is my being-in-the-world. Thus the rock is carved out on the ground of the world by the effect of the initial choice of my freedom. But on the other hand, what my freedom can not determine is whether the rock “to be scaled” will or will not lend itself to scaling. This is part of the brute being of the rock. Nevertheless the rock can show its resistance to the scaling only if the rock is integrated by freedom in a “situation” of which the general theme is scaling. For the simple traveler who passes over this road and whose free project is a pure aesthetic ordering of the landscape, the crag is not revealed either as scalable or as not-scalable; it is manifested only as beautiful or ugly» (Sartre, 1956, p. 488).

changing the body as such, but of the bodily relationship with the world which manifests itself as a set of our possibilities and intentions³⁹.

Now, let us imagine a person with brain lesions that has provoked irreversible damage to her/his ability to move (spasticity), feel (weakening of some afferent pathways) and communicate (difficulties in oral expression and writing). From a neuro-centric perspective, there is no relevant problem for consciousness, because the brain areas dedicated to the level and content of consciousness has not be involved in the lesion, so the ability to perceive the surrounding environment is supposed not be changed. However, the world manifests itself to the patient in a completely different way from the pre-injury period, as some things lose their role of possibility and instrument for relating with the world, while other people are revealed as distant, too fast compared to the patient's actions, therefore difficult to reach since communication is no longer efficient. This example may be useful for understanding how perception of the environment and of the others is not merely a matter of cognitive processing and categorizing. Rather, the perceived world is always revealed from a certain perspective “for me” (see chapter four), therefore «it is the manifestation of something only as it is experienced as referring to a body position» (Costa & Cesana, 2019, p. 210; my translation), my hodological body⁴⁰. In this regard, Costa and Cesana argue that neurorehabilitation is not a matter of restore the broken pieces of the brain at the pre-injury period (usually it is not an available option), as if brain and body were like a car

³⁹ «But this is precisely what separates Sartre’s and Merleau-Ponty’s competing accounts of the experience of that big rock, encountered as an obstacle. Sartre’s analysis conforms to the general framework we have just brought out. To encounter the rock as an obstacle or not (“scalable” or “non-scalable”), two contributing factors must be brought to bear upon the situation. First, there must be the individual’s “projected scaling” (Sartre 1956: 488). If I don’t entertain such a project when looking at the rock (if, in Sartre’s example, I am a “traveller” simply contemplating the rock), the rock will be neither “scalable” nor “non-scalable.” Second, the rock manifests itself as “scalable” or “non-scalable” “only within an instrumental-complex which is already established” (Sartre 1956: 482). And this second condition is, we recall, determined by the purely objective laws and rules governing the “dynamic order” of instrumental-complexes in general, including my own body as the (emptily indicated) instrument that I am. To put it plainly, if my climbing equipment, as well as the objectively ascertainable condition of my body, is adequate to the task of scaling that rock, then, given my project of climbing, the rock will manifest itself as “scalable” rather than “non-scalable.”» (P. Hoffman, 2006, p. 260).

⁴⁰ Let us take the specific case of Multiple Sclerosis: «The subject does not recognize his own movements anymore, the environment becomes impossible to master, his own walking unfamiliar, so that his own body identity is compromised. Thus it becomes evident how a motor disorder has on the one hand the meaning of an alteration of being in the world, since it subtracts an entire sphere of possibilities of action, on the other it has the meaning of a modification that goes to undermine the body identity of the subject, therefore the Self, since he no longer recognizes his own movements and appearance. Thus, the patient can observe his own handwriting with horror, and say with shame: “This is not my writing!”, that is, this isn’t me!» (Costa & Cesana, 2019, p. 212; my translation).

to be fixed, rather it consists on the reorganization of the relation between world and body, in order to articulate new possibilities towards the world (Costa & Cesana, 2019, pp. 63, 76–77).

Now, let us go through another kind of examples mainly related to morbid forms of being-in-the-world which may cause also malaise. In effect, someone may argue that these previous instances work only because they refer to sensory-motor contingences and interactions between environment and body, e.g. broken leg or spasticity, which are the only examples that can be explained through a “worldhood” approach. There are instead mental disorders like depression and mania that are something strictly related to the brain only, specifically to some neural mechanism that needs to be fixed. Being depressed or maniac has nothing to do with our mode of being embodied-in-the-world, one may say. However, also these cases can be examined by a “worldhood” approach. It has been showed how several mental disorders, which it is generally said they depends on a problem inside the brain only, actually are better described as modifications of lived embodied functions, and alteration of the being-in-the-world and being-with-others conditions. Recently, psychopathology and neuropsychiatry have made use of phenomenological and philosophical insights for a broad-spectrum and complete description of the pathologies listed here (Bizzari, 2018a; Fuchs, 2013b, 2019; Fuchs & Schlimme, 2009; Northoff, 2011, 2016b). In particular, depression and mania, instead of being considered cases of inner cognitive disorders, they have been investigated as related to the distortion of patients’ experience of the world in a bodily and temporal sense, where mania involves the acceleration of the way in which the world is experience, so that the environment becomes extremely slow compared to the subjectivity of the patient (Binswanger, 1964; Ghaemi, 2007; Martin, Gergel, & Owen, 2018; Moskalewicz & Schwartz, 2018), while depression, on the other side, involves a general slowing down of the ability to follow external stimuli, so much so that the world is experienced as too fast and difficult to interact with (Fuchs, 2013a; Ghaemi, 2007; Stanghellini et al., 2017). This phenomenological interpretation finds support in recent neuroscientific studies based on a spatiotemporal model of the brain, and that in cases of depression and mania the spontaneous activity of the brain seems to “slow down” or “accelerate” by

misaligning itself from the world and body (imbalance in interoceptive-exteroceptive processing) (Northoff, 2016e, 2016f; Northoff et al., 2011, 2018).

According to this interpretation model, the brain surely is still a crucial element to investigate in order to understand what happens when this kind of mental disorders appears, however the way in which the brain activity is considered is different from the classic cognitivist approach. Instead of asking “what’s on her/his mind/brain?”, the worldhood approach questions the structure of how the world appears to the patient, therefore asking “what kind of world does she/he live in?” (Costa & Cesana, 2019). These two questions represent different approaches towards the brain. According to the cognitive neuropsychology, the brain takes a constructive role, i.e. the brain associates the isolated and neutral stimuli (input), creates the spatio-temporal matrix and gives meaning to the world (output). On the other side, according to the “phenomenological neuropsychology” (Costa & Cesana, 2019), the brain takes a receptive (not merely passive) role, i.e. it is directed towards a world made of an wide context of possibilities of action, objects and meanings, where the environmental stimuli give shape to the brain during its formation and the brain in turn is crucial for the integration and selection of the objects of the experience of the world.

One of the main aims of this research has been to argue for an account of consciousness as embodied relation embedded into the world, i.e. embodied-being-in-the-world, and I am trying to demonstrate with several examples that this relation is spatiotemporal and that a study of the brain alone or of the body alone will not be useful for an exhaustive description of how we experience. These last examples on mental disorders allow me to introduce the question of how the space-time matrix is concretely formed. So far, I have provided a description of how experience develops essentially as world-disclosing in a space-time sense and I have described the phenomenological-existential conditions of this consciousness as embodied-being-in-the-world. Now, what are the material conditions of this existential condition? That is, what happens concretely in the human being so that we can give an experience? Indeed, we have seen that the condition of being-in-the-world can be disturbed by different types of malaise, however this experiential condition is formed on a neurological, corporeal and environmental basis without which there would be no possibility of experience (such as, for example, in cases severe morbid form). If we

want to pursue the aim of an integrative, pluralist and interdisciplinary account of consciousness, we cannot stop at the philosophical-phenomenological definition phase. It is necessary to continue towards the empirical conditions of consciousness, to integrate the phenomenological and conceptual analyses with the neuroscientific results. Of course, this cannot be a simple juxtaposition of empirical data alongside phenomenological descriptions. As happened with the examination of consciousness from a non-mento-centric point of view (chapter five), so the interpretation of neuroscientific data will have to be non-neuro-centric (chapter three).

6.3 Neurophilosophy of Consciousness

Before continuing with the neuroscientific hypotheses, there is an epistemological question to be resolved. Notwithstanding that a theory of consciousness cannot be limited to presenting the neuronal correlates in the brain, how is it possible to suggest to the neuroscience to implement a non-neuro-centric paradigm? It seems a rather paradoxical request, like asking biology not to be too concentrated on the concept of life (*bíos*) or physics not to deal exclusively with nature (*physis*). However, this potential criticism completely misses the point regarding the non-neuro-centric approach, indeed, as I tried to show in the previous chapters, the goal is not to avoid or limit the use of neuroscientific data within consciousness studies, but to insert them into a broader and non-reductionist interpretative practice.

Therefore, in light of what has been observed up to now regarding the description of the phenomenological structures of consciousness and its intrinsic relationship with the world and the body, it seems necessary to find an alternative to the neuro-centric view that analyses the brain as an isolated and autonomous thing which, at the same time, is supposed to be the ontological ground of our own consciousness, selfness and personhood (the brainhood paradigm). To achieve this goal, we need an approach that, while remaining adherent to the rigid analysis of brain activity, is concerned with linking neuroscientific results to the activity of other parts of the body (heart, digestive tract, afferent and efferent pathways, etc.) and environmental stimuli in a different way from the input-output model. Now I will briefly analyse three main neuroscientific theories of consciousness: Integrated

Information Theory, Global Neuronal Workspace Theory, and Temporo-spatial Theory of Consciousness.

6.3.1 World and Body in neuroscientific Theories of Consciousness

6.3.1.1 Integrated Information Theory

Integrated Information Theory (IIT) claims that if a system has the intrinsic potential to integrate information, then we can assume presence of consciousness (Tononi et al., 2016). Indeed, according to IIT, an experience does not require the involvement of body and world, language, introspection or reflection, it can exist without spatial frames of reference or a sense of the body and the self (Laureys & Tononi, 2009a), and it is not reduced to attention or memory (Tononi, 2017b). More precisely, Tononi and colleagues argues that information coming from the environment surely is an important element for the brain, shaping the quality and structure of the conscious experience we have, through the evolutionary history of the human brain (Joshi, Tononi, & Koch, 2013), as well as the development and the plastic changes of the individual (Tononi, 2003, 2008). However, the brain as such, or even better a precise conceptual structure within it, i.e. cause-effect structure, can be the only sufficient condition for consciousness. In this respect, Tononi and colleagues make the example of a dreaming brain as isolated from the external world, in a body completely paralyzed and unresponsive, but still producing a world of conscious experiences (Nir & Tononi, 2010; Sarasso et al., 2014; Tononi, 2012)⁴¹.

Hence, consciousness as integrated information is intrinsic and solipsistic, and it does not necessarily require any extrinsic relation outside the system, nor functions or purposes towards something else (Tononi, 2008). Indeed, a system with the right

⁴¹ «[...] consciousness – in the sense of having an experience – does not require sensorimotor loops involving the body and the world, does not require language, introspection or reflection, can do without spatial frames of reference and perhaps even without a sense of the body and the self, and does not reduce to attention or memory» (Laureys & Tononi, 2009a, p. 376).

architecture for a sufficient integration of information could exist in complete detachment from the external world, keeping producing its own internal states (Oizumi et al., 2014). Furthermore, according to IIT, the brain is a sufficient non-necessary condition for consciousness, since that, in principle, any artificial or natural being which integrates information can be conscious, even without a brain (Oizumi et al., 2014; Tononi, 2017b)⁴².

6.3.1.2 Global Neuronal Workspace Theory

The GNWT is a cognitive theory of consciousness, according to which conscious access is global information availability (Dehaene & Naccache, 2001). The theory argues for the existence of recurrent and competitive top-down/bottom-up loops that accumulate information from perceptual, motor, attention, memory and value networks, then sharing and broadcasting back information to lower-level processors. This broadly distributed network permits conscious access to the unified and synthesized information, i.e. consciousness of some specific content (Dehaene et al., 2011).

This theory does not claim for the necessity of some particular brain area, because it is the global state of information availability as such, i.e. neuronal workspace through long-range cortico-cortical fibers, that is necessary for consciousness. According to the theory, a neuronal marker useful to determine the level of information availability is the late component P300, more in particular P3b (in fact, P3a component may reflect automatic and non-conscious attention) (Dehaene et al., 2011). P3b is observed mainly during subjective reports and the diffused and simultaneous involvement of multiple

⁴² Another “brain-centered” (p. 211) position focused on the intrinsic activity of the brain seems to claim a similar position (Farisco, Laureys, & Evers, 2017). According to Farisco et al., brain is not an isolated system nor a mere input-output machine, rather the brain is constantly evaluating through its intrinsic activity, modeling and predicting the world. In addition, feedback and feedforward activity of the brain are so strongly related as to make it difficult to draw a clear-cut separation between intrinsic and extrinsic activity (Soddu et al., 2009), since the first is the necessary condition for the onset of the second, which in turn modulates the former (Farisco et al., 2017). However, this theory, considering consciousness as a product and property of a functional and structural network of the brain, claims that the only concrete condition for the rise of experience is the intrinsic and resting state brain activity, and those brain activities, under specific configurations, can be conscious as such.

area activations⁴³. According to the theory, the world plays an “accessory role”, since consciousness can be mainly described as a process of unification, selection and broadcasting of information, i.e. “just brain-wide information sharing” (Dehaene 2014: 165). The world can be understood as the set of stimuli that comes into the brain through the sensory system and is processed by a multitude of unconscious processes, in order to produce a faithful representation of the environment for the following high-cortical processing, i.e. conscious access.

6.3.1.3 Temporo-spatial Theory of Consciousness

The Temporo-spatial Theory of Consciousness (TTC) focuses primarily on the temporal and spatial features of the brain activity, in particular on its spontaneous or intrinsic activity (the so-called resting state) conceived in spatiotemporal terms, that is, structured by various mechanisms in terms of spatial (functional connectivity), and temporal, (fluctuations in different frequencies) features (Northoff & Huang, 2017). According to TTC, the brain is understood as embedded in the body and in the environment, so that the brain’s intrinsic activity serves as a background and fundamental condition for all the stimulus-induced interactions with body and environment. In other words, the brain simply does not function according to an input-output model, but its spontaneous activity acts as a baseline that constantly tries to align itself with the outside, influencing before and after stimulus processing. In this sense, the brain is not isolated (Northoff, 2004a), rather is strongly dependent on the environmental context and this is particularly evident either in cases of abnormal consciousness caused by early traumatic childhood events or mental disorders (Northoff, 2016b, 2018a) but also in normal cases like while dancing⁴⁴. «This and

⁴³ There is currently a debate about the reliability of P3b as a valuable index of consciousness. Silverstein et al. sustain that P3b alone cannot be identified as a sufficient neural correlate of consciousness, arguing that it is elicitable also in some unconscious processing (Silverstein et al., 2015). This fed a debate about the correlation between this EP component and conscious access (Naccache et al., 2016; Silverstein et al., 2016).

⁴⁴ «For instance, a study by Duncan et al. (2015) showed that traumatic life events in early childhood impact the brain’s spontaneous activity in adulthood: the higher the degree of early traumatic childhood life events (i.e., either a greater number of traumatic events or events that are more traumatic), the higher the degree of spatiotemporal disorder (e.g., entropy, in the brain’s spontaneous activity in adulthood). This makes it clear that the brain is not an isolated organ encapsulated in the skull that communicates with the environment only through the body. For instance, the brain can shift the phase onset of its fluctuations (in different frequency ranges) in accordance with the onset of tones in the external

other examples make it clear that the brain is a highly context-dependent organ that stands in direct relation to the respective environment, with the latter impacting, modulating and sculpting the former. Taken in this sense, the brain can no longer be considered an isolated organ but a relational organ that can be characterized by what we describe as a world–brain relation or ‘environment–brain unity’» (Wagner & Northoff, 2017, pp. 344–345).

This integration is normally sustained by the body, understood as a means or a mediator for aligning/interfacing brain with world and relating interoception with exteroception (Northoff, 2011). This theory of consciousness leads us to an important distinction between the accessory and constitutive role of the world for consciousness. In the first case, the world plays a secondary role (see IIT and GNWT), being useful for the development of the brain and for the enrichment of the subjective experience, but consciousness could arise even without it, purely through brain activity. In the second instance, the world is constitutive as a necessary predisposition for consciousness, like in TTC as I will soon explain. It remains to be seen how it is related with consciousness and in which way it is relevant for the brain activity. Differently from IIT and GNWT, TTC does not argue for a neuro-centric view of consciousness, albeit claiming for the necessity of the spontaneous activity of the brain as predisposition of consciousness, i.e. the spatiotemporal structure through which external stimuli are processed and encoded (Northoff, 2014c, 2018b). This means that TTC is grounded on a brain-based approach rather than a brain-centered or brain-reductive one (see chapter three), since being based on something does not lead to the exclusive dependence on or identity with it (Northoff, 2014c). Therefore, the brain plays a necessary non-sufficient condition for consciousness, while the world acts as an enabling condition through which consciousness can arise, within the sustained

environment (see chapter 20 in Northoff, 2014b, for details). That is apparent when we listen to music and swing our perceptions and movements in tune with the rhythm. Such direct coupling between brain and environment has also been described as ‘active sensing’ or the brain’s ‘sixth sense’ (van Atteveldt et al., 2014)» (Wagner & Northoff, 2017, p. 344). I am not claiming that for the abovementioned theories of consciousness the environmental influence is not relevant for the development of a proper brain activity, nevertheless for TTC the environment plays a crucial role for the emergence of the experience, as a necessary condition in the world-brain interaction. While the other two theories consider the possibility of consciousness in a hypothetical isolated brain, excluded from body and environment, this is impossible for TTC, as the world (body included) is not a mere series of stimuli processed by the brain (neuro-centric view) but is the wider spatiotemporal context within and with which the brain interacts.

relation between environment, body and brain. Northoff's TTC offers the best neuroscientific approach to the methodological and epistemological points required by the present research, as the theory is proposed as an alternative to the classical neuro-centric paradigms. Now, I will explain how his theory can be useful to complete our account of consciousness.

For Northoff, the brain must be conceived as a “relational organ” that is part of a systematic and broader perspective over the human being and consciousness. Instead of the neuro-sensory (brain's spontaneous activity is passively modified by the world as a wax tablet) or the neuro-cognitive (brain's spontaneous activity is directly interacts with or impacts on the neural processing of external stimuli) models, he proposes a neuro-ecological model of the brain as necessarily related with body and environment (Northoff, 2016c). The neuro-ecological model considers the relationship between environment and brain as the most basic and fundamental⁴⁵, where the brain is not merely passive (neuro-sensory) nor aprioristically active over the environment (neuro-cognitive), rather it receive the stimuli from the environment and the body, progressively adapting on and depending to them (“experience-dependence”; Northoff, 2016b), while at the same time amplifying or attenuating them on the basis of specific mechanisms of the spontaneous activity⁴⁶. Particularly, one of these mechanisms is the so-called spatiotemporal alignment, which is extremely relevant for consciousness. To put simply, spatiotemporal alignment refers to the ability of the brain's spontaneous activity to tune (e.g. in terms of cerebral frequencies) with the body and the world⁴⁷, so the more one is aligned with the body and the environment,

⁴⁵ «The neuro-ecological continuum is characterized by the balance between neuronal and ecological features in the brain's neural activity. Therefore, the brain's neural activity can be described as hybrid in that it is neither purely neuronal nor purely ecological. On a more general level, the hybrid neuro-ecological nature of the brain's neural activity means that it relates world and brain—this amounts to what I call the world–brain relation. What do I mean by world–brain relation? First, the concept of world–brain relation is here understood in an empirical sense. It denotes the relation between world and brain that is constituted by the brain's spatiotemporal alignment in its interaction with the world» (Northoff, 2018b, p. 223).

⁴⁶ From a phenomenological perspective, it is the world has a system of meanings and references that gives the brain a structure, not *vice versa* (Costa & Cesana, 2019). From a neuroscientific perspective, it is the environment that influences and organized the brain's spontaneous activity, i.e. experience-dependence and ecological dependence (Northoff, 2016c).

⁴⁷ «The brain's spontaneous activity and its temporal structure seem to align themselves to the temporal structure of the body—one can thus speak of spatiotemporal alignment of the brain to the body [...] The same holds, analogously, with regard to the world—the brain's spontaneous activity and its spatiotemporal structure align themselves to the world. This is most apparent when we listen to music and dance to its rhythm—we align our brain's temporal structure of its neural activity (as in its

the more one is able to consciously discriminate and/or integrate the stimuli, or in phenomenological terms, to position intentional objects against the phenomenal background of being in the world. This alignment may be subject to alterations that may be associated with mental pathologies and disorders of consciousness (Northoff, 2018a).

6.3.2 Northoff's World-Brain Relation and the link between Neuro-ecology and Phenomenology

Thus, according to TTC, a fundamental integration within the world (and not only within the body) is crucial for the brain in order to maintain the basic predisposition for the presence of consciousness (Northoff, 2018b). This is called by Northoff the “world-brain relation” (Northoff, 2018b), an ontological and pre-phenomenal relationship as a condition of possibility for the actual emergence of consciousness⁴⁸. Therefore, in order to avoid neuro-centrism and mento-centrism, he suggests to shift the focus from brain, mind with the respective mind-body and mind-brain problems to the relationship between the world and the brain, so that experience is necessarily related to the brain, but not as an isolated organ, rather only within the world-brain

frequencies and synchronization) to the temporal structure of the music and, more generally, the world [...] Most important, empirical data suggest that such spatiotemporal alignment of brain to body and world is central for consciousness [...]. The better our brain aligns us to body and world, the more likely we can become conscious of the respective contents in body and world (see discussion in chapters 7 and 8). Hence, spatiotemporal alignment of brain to body and world is central for mental features such as consciousness» (Northoff, 2018b, p. xvi).

⁴⁸ Interestingly, the world-brain relation is not observable, nor in third-person (e.g. you do not see anything but the brain through neuroimaging techniques) nor in first-person (e.g. you cannot subjectively report the presence of a relation between the world and your brain). Nevertheless, we can indirectly access the world-brain relation «by inferring from combined experience and observation in very much the same way that Nagel suggests the need for a correspondence between the first- and third-person perspectives [...] At the same time, a proper world-brain relation may be conceived of as a necessary or predisposing condition for the possible perspectival differentiation between the first- and third-person perspectives (Northoff, 2014b). One may consequently feature the world-brain relation as preperspectival rather than as completely nonperspectival or perspectival: the prefix ‘pre-’ indicates here that the world-brain relation is the necessary condition of subsequent perspectival access to reality in terms of first-, second- and/or third-person perspectives. In contrast, it would be nonperspectival when there is no relationship at all between world-brain relation and the different perspectives» (Wagner & Northoff, 2017, pp. 348, 349).

system. According to Northoff, this shift requires nothing less than a “Copernican revolution” in neuroscience and philosophy (Northoff, 2018b, 2019)⁴⁹.

In this way, it is possible to re-think also the relationship between consciousness and brain. We have seen in the second chapter that a necessary a posteriori identity between brain (or brain activity *x*) and consciousness is not possible, and that we can, at most, conceive the relationship as a “contingent identity” (i.e. a non-identity as Kripke would say) in the sense that at certain points in time and space a particular conscious state is associated with a particular brain state, nothing more. Now, after having exposed the world-brain relation, we can see how the problem was not about the kind of relationship, rather about the way in which we conceive the brain as isolated. Insofar as the brain is taken as an isolated organ within an element-based or property-based ontology, we will not be able to go beyond contingent relationships between brain areas and conscious states, i.e. the neural correlates of consciousness. Instead, if we start analysing the brain (not only conceptually but also empirically) as part of a broader relational (rather than element-based) system, i.e. spatio-temporal world-brain relation, the neuroscience of consciousness could take a step towards understanding the necessary conditions of the relationship between brain and consciousness. Northoff argues that the brain as such cannot directly be the necessary (a posteriori) condition for consciousness⁵⁰, and this is in line with what I sustained in the second chapter. On the other side, the world-brain relation is an ontological predisposition of consciousness, i.e. necessary (and non-sufficient) a posteriori, since consciousness cannot be maintained without a proper spatio-temporal connection with body and environment. Therefore, given that, according to the neuro-ecological view, consciousness can be defined as a world-brain structure in a necessary a posteriori (i.e.

⁴⁹ «Analogous to Copernicus (albeit in a weak rather than strong sense), I argue that we need to replace our current vantage point from within mind (or within brain) to a vantage point from beyond brain. Such a vantage point from beyond brain will allow us to take the view that the brain is part of the world by being related to it, that is, by a world–brain relation comprising structure and relation as the basic units of existence and reality [...] In the same way that an intuition of Earth as being the center of the universe was rendered impossible by the dramatic shift in vantage point of Copernicus, the vantage point from beyond brain now makes impossible the intuition of mind as the center of mental features, ourselves, and the world. This enormous paradigm shift allows us to replace the mind–body problem by the world–brain problem—this shift in viewpoint amounts to a Copernican-like revolution in the spheres of both neuroscience and philosophy» (Northoff, 2018b, p. xxiii).

⁵⁰ «Importantly, the brain alone, that is, independent of its relation to the world as defined in element-based ontology (chapter 9), does not have a necessary but only contingent connection to consciousness» (Northoff, 2018b, p. 270).

empirically discoverable) way, then, by transitive property, we can indirectly conceive the brain in a necessary relationship with consciousness, as long as the brain is incorporated within body and environment⁵¹. This can be also related to the concept of consciousness as rigid designator of “being aware of something” (phenomenological point of view) and “environment-body-brain alignment” (neuroscientific point of view), where the spatio-temporal dynamics for conscious alignment amounts to the “atomic number 79” for “gold”.

In this regard, one might notice a certain similarity with the condition of being in the world, where being in the world is a necessary condition of pre-intentional possibility for consciousness as intentionality. Just as intentionality towards objects is not phenomenologically possible without the subject being first situated into a pre-given being-in-the-world, so the emergence of consciousness is not empirically given without an adequate alignment between environment, body and brain. The world-brain relation and the embodied-being-in-the-world condition seem to be two diffractions of the same consciousness phenomenon, where the former offers the conditions for any possible consciousness, i.e. there is no emergence of consciousness without a sufficiently strong relation between environment, body and brain (Zhang et al., 2018; see also morbid forms⁵²), and the latter offers the phenomenological and existential condition for consciousness, i.e. there is no normal intentionality and sense-disclosing towards the world without embedment and embodiment of the individual into a meaningful world-context (see malaises).

Moreover, it has to be noted that Northoff speaks of a world-brain relation, rather than brain-world relation, and this distinction is relevant because it indicates a difference between the way we can conceive this relation. The world-brain relation

⁵¹ «I argue that the assumption of world–brain relation as OPC renders it possible to conceive the relationship between brain and consciousness in a necessary (a posteriori) rather than contingent way: as world–brain relation provides the necessary ontological condition of possible consciousness (i.e., OPC), the brain, as ontologically defined by world–brain relation, is also necessarily (rather than contingently) connected to consciousness» (Northoff, 2018b, p. 270).

⁵² «If the world–brain relation is disrupted, becoming nonperspectival rather than preperspectival, we may lose our personal identity, including our relation to the world. For instance, patients in vegetative state, who have lost consciousness, lose the neuro-ecological intimacy of their world–brain relationship, which renders the latter nonperspectival. That, in turn, makes it impossible for them to take on any kind of perspective, neither first- nor second-person perspective (as manifest in the absence of first-person experience and second-person intersubjectivity) nor third-person perspective (as manifest in the absence of third-person observation)» (Wagner & Northoff, 2017, p. 349).

refers to the way in which the human brain with all its small spatio-temporal dynamic is nested within larger systems, like body and environment, and this means that it is the world that sets the tone of the relation, and the brain is receptive and tries to align itself to the world⁵³. The brain-world relation instead refers to the ability of the brain to connect with the world by means of sensory, affective, cognitive, social functions as well as through enactment and embodied extension. According to the neuro-ecological approach the brain-world relation must be included into the world-brain relation, with the latter being the condition of possibility of the former⁵⁴, while the classic neuro-centric approaches conceive only the cognitive brain-world relation⁵⁵, which instead I would consider as a complementary rather than a necessary condition for consciousness. This interpretation of the bidirectionality and asymmetry of the relationship between world and brain opens to similarities with what was discussed above regarding the concepts of extendedness, embodiment, enactment, embedment, etc. in relation to the Sartrean phenomenology. As said, these concepts respect the sense of intentionality as a disclosing act of the world as long as they remain within the idea that we are not heading towards a world separated from us, but it is the world that shows us by manifesting itself and engaging our body, a body that is located from the start in a contingent situation of “facticity” (*facticité*; Sartre, 1956). If instead we mean intentionality like the arrow that goes from the head towards the body and the

⁵³ «The world includes much slower frequencies, such as seismic earth waves, than does the brain. Therefore, those slower frequencies nest and contain the brain’s faster frequencies—taken in purely spatiotemporal terms, the brain is thus nested and contained within the world. For that reason, I speak of world–brain relation rather than brain–world relation» (Northoff, 2018b, p. 225).

⁵⁴ «A relation between world and brain that can be understood in a bi-directional way as well as in both empirical and ontological domains. Empirically, the brain can relate to the world on the basis of its task- or stimulus-induced activity that is related to the brain’s cognitive, sensory, motor, social, and affective functions—this can be described as brain–world relation. However, the brain can relate to the world also by its spontaneous activity that shows strong spatiotemporal alignment to the environmental context. As the brain adapts and aligns to the world as the primary origin of the world’s relation to the brain, I speak of world–brain relation rather than brain–world relation. [...] I therefore characterize the concept of world–brain relation in an ontological sense which, as in the empirical context, must be distinguished from brain–world relation. Finally, both brain–world and world–brain relation must also be distinguished on methodological grounds: brain–world relation presupposes a vantage point from within brain, whereas world–brain relation can only be taken into view by presupposing a vantage point from beyond brain» (Northoff, 2018a, p. 453).

⁵⁵ «In contrast, I do not suppose analogous relevance of brain–world relation for consciousness. Rather than predisposing consciousness, brain–world relation is central for sensory, motor, cognitive, affective, and social function with subsequent brain-based perception and cognition of the world. Without elaborating the details here, I therefore deem brain–world relation to predispose cognition rather than consciousness. To now replace world–brain relation by brain–world relation is to confuse consciousness and cognition» (Northoff, 2018b, p. 229).

environment, then we remain within a mento-centric perspective, which nevertheless expands and extends through the body.

In other world, considering consciousness as the direction from an inner mental life to the external world is just an enlarged version of the Cartesian ontology, and in this sense, both Sartre's phenomenological realism and Northoff's world-brain relation attempt to overcome this issue and be more radical, arguing that there is no intentionality without being-in-the-world, where consciousness is necessarily being-in-the-world and consciousness-in-the-world (Sartre, 1939/1970) and that the world includes within it the brain and not vice versa (Northoff, 2018b). In this sense, the example of dance no longer assumes a simple metaphorical value, but becomes a case that can be analysed empirically (for details, Northoff, 2018b, pp. 206-208, 211-213), through which it can be demonstrated that, before being able to dance and cognitively keep the pace with the body – e.g. making a particular move – in direction of the world – the dance floor – (brain-world relation), one must be receptive to the rhythm of the environmental context through the body in order to feel the beat and be in tune with it (world-brain or – as I would say – environment-body-brain relation)⁵⁶; after being synchronized with the music we can actively participate to the rhythm with our body, through whose movements we can strengthen the relationship with the world.

Now, this neuro-ecological approach towards consciousness implies that there is no original separation between brain, body and world, rather an original relational unity, so that when we speak, for example, of the brain processes information from the body and world, we are simply categorizing the single parts of a unified continuum consisting of time and space relationships:

⁵⁶ «[...] the brain's spontaneous activity and its spatiotemporal structure align themselves to the world. This is most apparent when we listen to music and dance to its rhythm—we align our brain's temporal structure of its neural activity (as in its frequencies and synchronization) to the temporal structure of the music and, more generally, the world [...]. One can thus speak of spatiotemporal alignment of the brain to the world (chapter 8; Northoff & Huang, 2017). [...] We are confronted with various types of stimuli in our environment that need to be encoded by our brain. For example, when we hear a music piece that is rhythmic, our brain seems to encode the rhythmic structure of the melody in such way that we are able to align ourselves to the melody. This allows us to participate in the melody's rhythmic structure when, for instance, we swing our arms and legs while dancing. Our brain seems to sample the rhythmic structure of the tone sequence presented, which enables our participation in the rhythmic structure of the environmental events or objects (i.e., the music piece)» (Northoff, 2018b, pp. xvi, 206).

«The neuro-ecological continuum is based on space and time: there is a continuum between the spatiotemporal structure of the brain's spontaneous activity and the spatiotemporal structure of the ecological context, the world. This spatiotemporally based neuro-ecological continuum between world and brain is based empirically on the mechanisms of spatiotemporal alignment that are central for consciousness. Together, spatiotemporal alignment and the neuro-ecological continuum provide the empirical ground, that is, the necessary empirical conditions, of what on a more conceptual and ontological level can be described as the world—brain relation» (Northoff, 2018b, pp. 445–446).

The world-brain relation has not to be conceived by a relationship between two separated poles, but as a system of inclusion where the smaller systems (e.g. brain) are inserted inside the larger ones (e.g. the brain in the body and the body in the environment)⁵⁷. Consequently, the neuro-ecological structure of consciousness cannot be located by “simple location” (see chapter three) in the brain, in the world or in the body alone, but by “complex location” into the dynamical spatiotemporal relation engaged by these – only conceptually and methodologically separated – entities⁵⁸.

The spatiotemporal relation is considered by Northoff the fundamental framework upon which different features of the world-brain relation (e.g.

⁵⁷ In this sense, world is not necessarily the environment, because it generally includes also the body, so that the world-brain relation can be conceived as an environment-body-brain relation. Moreover, in some pathological cases, the environment may disappear from the relation or the body may become the centre of conscious experience, neglecting the surrounding environment. In those cases, the body-as-object is the world which the brain is related to, as in schizophrenia (Northoff & Stanghellini, 2016).

⁵⁸ «How can we characterize such complex location? Unlike in simple location, we can no longer “localize” different existences and realities such as brain and nonbrains at one specific space-time point or event independent of each other and their respective spatiotemporal context. Instead, brains and nonbrains may be located on different positions of an underlying commonly shared spatiotemporal spectrum, a “spatiotemporal trajectory,” as I say (chapter 12). Let me give an empirical example of such complex location in terms of spatiotemporal trajectories. Take the brain's spontaneous activity. The brain's spontaneous activity shows continuous activity changes across time and space, that is, in different regions and frequency ranges. One can now locate each activity change by itself at each specific space time point or event—this amounts to simple location. Alternatively, one can consider the different activity changes in relation to each other: the degree of change in space-time points or events from one activity to another and so forth. One then locates the different activity changes in dependence on each other along the line of an ongoing spatiotemporal trajectory—this amounts to complex location in an empirical sense. Taken in an ontological context, complex location allows us to locate the brain in relation to and dependence on its respective spatiotemporal context as constituted by body and world. More specifically, the brain and its neural activity can be located in relation to the spatiotemporal features of body and world in terms of space-time relation. Complex location thus locates the brain ontologically in a way that operates across the observable space-time points or events of brain, body, and world. This distinguishes complex location from simple location that remains within the confines and boundaries of the brain and its space-time points or events» (Northoff, 2018b, pp. 343–344).

consciousness as such but also the four “E” of 4E cognition) are based. This spatiotemporal framework is not to be considered as space and time as they are taken for granted by the sciences, nor as time and space in which we arrange objects in the world through cognition. Rather, it is the intrinsic time and space that the brain “builds” in relation to the space-time dynamics of the world, so that, ontologically speaking, such construction of time and space amounts to “relational” time and space (Northoff, 2018b, chapter 9) instead of “observational” or “container” time and space⁵⁹. Empirically speaking, a brain that is not well tuned to the dynamics of the world (e.g. too fast or too slow in the brain frequencies with respect to the stimuli received) could be associated disorders or alterations of consciousness (Northoff et al., 2018; Northoff & Huang, 2017; Northoff & Tumati, 2019).

On this topic, a further link can be built with the phenomenological account of consciousness previously exposed, which is essentially based on spatialization and temporalization (see above). Of course, the space and time concepts considered by Northoff’s neuro-ecological approach are different from the phenomenological ones. Both phenomenology of temporalization and spatialization and the spatiotemporal framework of the world-brain relation are not cognitive, i.e. they are not related to the way in which we represent time and space to us. However, the former is related to the way in which we directly experience the world, i.e. we cannot be conscious of something outside a time-space context, so temporalization and spatialization refer to our lifeworld. Differently, the latter refers to the condition in virtue of which we may perceive a world in a spatiotemporal way. Thus, phenomenological temporalization and spatialization describe how we experience, while TTC describes what are the necessary pre-phenomenal conditions for us to be conscious in a spatiotemporal way. Nevertheless for both the approaches, time and space are the building blocks of

⁵⁹ «I here emphasize that one cannot take time and space for granted in neuroscience as it strongly affects how we conceive the brain and its relationship to consciousness. Time and space must be considered separate and distinct in empirical and ontological contexts or domains. Empirically, we can observe discrete points in time and space—as such this view on time and space is based on observation, I speak of observational time and space. Observational time and space are related to us as observers and how we perceive and cognize time and space. Such observational time and space need to be distinguished from the time and space the brain itself constructs in its own neuronal activity that characterizes the brain’s existence and reality. We thus need to distinguish between perception/observation/cognition of time and space in the brain on the one and construction of time and space by the brain itself. I postulate that the latter, i.e., construction of time and space by the brain, is central for consciousness» (Northoff, 2018a, p. 451)

consciousness, that is, on the one hand experience can be conceived only starting from the concepts of temporalization and spatialization of objects of/in the world, on the other hand, the neuro-ecological view of consciousness understands the brain as incorporated within a time-space relationship with body and environment.

The differences with respect to the two approaches need not necessarily be understood as a problem, given that the present research does not pursue the reduction of one entity, phenomenon or concept to another. As discussed in chapter four, neuroscience and phenomenology move from different regional ontologies and perspectives, so the goal is not to identify identities but correlations between different approaches and concepts, integrating them within a pluralistic and broader perspective. For example, the world-brain relation is certainly different in detail with respect to the embodied-being-in-the-world condition, but this does not indicate different aspects of the same phenomenon. This does not mean however that any neuroscientific or phenomenological position can be related, for example we had to eliminate IIT and GNWT because of their accessory conception of the environment and of the world in general for consciousness. Likewise, a theory like TTC that uses a neuro-ecological approach could not find fruitful connections with internalist philosophy of mind. In the case of space and time, if the neuro-ecological approach conceived them as containers for objects rather than as relational structures between objects, a connection with the phenomenology of temporalization and spatialization would be rather difficult. Northoff argues that time and space are the “common currency” of neuronal and phenomenal states (Northoff, Wainio-Theberge, & Evers, 2019) and consequently I believe that the analysis of the spatiotemporal dynamics of experience may be the connection between empirical and phenomenological analysis of consciousness.

Taken together, the Temporo-spatial Theory of Consciousness (TTC) it is the best candidate to establish a connection between the phenomenological analyses described above and the neuroscientific studies on the brain. In particular, TTC is based on the neuro-ecological view that considers the relation with the world (environment and body) as an enabling condition necessary for the formation and preservation of consciousness (world-brain relation). This is in line with what was previously argued from a phenomenological point of view, which therefore finds an empirical support for the description of consciousness as embodied-being-in-the-

world. Moreover, even Northoff's approach indicated as "Spatiotemporal Neuroscience and Psychopathology" seems to indicate the possibility of fruitful connections and confirmations with the phenomenology of time and space in consciousness (Northoff & Stanghellini, 2016; Northoff et al., 2019).

6.4 *Experimentum crucis*: consciousness without body and environment?

So far, this research led me to argue in favour of the need for an intrinsic relationship between experience and world, between consciousness and body, between brain, body and environment. This relationship, which I have tried to support both from a philosophical-phenomenological and neuroscientific point of view, develops from the need to rethink the approach to consciousness and thus to overcome the so-called neuro-centric and mento-centric approaches that I have discussed extensively in the previous chapters. Above I have presented observed examples through a worldhood approach, that is, through the observation of the changes of the brain in relation to the changes in the relationship with the world and not as pieces of a brain-machine to be replaced (brainhood approach). Following the presentation of the Temporo-spatial Theory of Consciousness, it now seems even more evident, even from an empirical point of view, that disorders and pathologies of the experience can be addressed in a broad and pluralist way rather than merely from a neuro-reductionist perspective way than the neuro-reductive one.

Nevertheless, there are situations in which the worldhood paradigm and the neuro-ecological approach to consciousness may be difficult to apply. In this respect, I will focus on another kind of examples that seems to resist these approaches. Indeed, if perhaps the structure of consciousness can be described as embodied-being-in-the-world and explained by world-brain relation, then what would happen if the relationship between body and world was interrupted but brain activity was preserved and allowed the emergence of consciousness? We have understood that in the case in which the brain is no longer able to produce a normal activity, e.g. for example due to a serious damage, the relationship with the world would be lost, although the rest of

the body is intact. Consider, for example, disorders of consciousness (DOC) such as coma, unresponsive wakefulness syndrome/vegetative state (UWS/VS) and minimum conscious state (MCS), but also the state of sleep understood as a reversible altered state of consciousness. All these are cases of disruption of the relation with the body and world by means of an alteration (normal in case of sleep or abnormal in the others) of the brain activity, therefore these cases fall perfectly within the neuro-ecological description of consciousness as – empirically speaking – environment-body-brain relation.

However, there are also clinical situations in which the patient partially or completely loses the relation with the environment, the body or both, while presumably preserving normal activity in the brain. These cases can be used to test some of the arguments held in the present research, since, at least provisionally, there may be two possible alternatives, i.e. a healthy brain (CNS) without relationship with the environment and body either experiences something, or consciousness is interrupted. In the first case, the neuro-ecological approach and worldhood paradigm they would have trouble explaining how an isolated brain can be conscious, while in the second case the brainhood and neuro-centric approaches would face difficulties. The first case is the Locked-in Syndrome.

6.4.1 Locked-in Syndrome

By definition, the syndrome is usually described as a situation in which the body is paralyzed but consciousness and cognition are fully preserved or intact (Gosseries et al., 2009; Murguialday et al., 2011; Schnakers, Perrin, et al., 2009; Schnürer, Espinosa, & Guger, 2015). In fact, the common idea of LIS is that of a living brain with all its mental features like consciousness and cognition (memory, imagery, communication abilities, etc.) perfectly intact and well-working, but locked inside a detached “dead cold body” (Kondziella, 2017) or even a “carcass” (Bauby, 1997). Despite the bodily-impaired situation, the patient has preserved basic cognitive abilities, awareness, sleep-wake cycle and capability of meaningful, but minimal, behaviour (Gosseries et al., 2009; Guilmette, 1997)⁶⁰. For this reason, LIS is generally

⁶⁰ The first scientific mention of a clinical situation resembling Locked-in Syndrome was described in 1941, subsequently diagnosed as akinetic mutism (Cairns et al., 1941). The first diagnosed case of LIS due to infarction of the brain stem was in 1947 (León-Carrión, Eeckhout, & Domínguez-Morales, 2002),

not interpreted as a disorder or alteration of consciousness, as far as it can be confused with one of them, like coma, vegetative state (VS) (Blumenfeld, 2009) or, more generally, unresponsive wakefulness syndrome (UWS), i.e. a category that groups patients who fail to show voluntary motor responsiveness in the presence of eyes-open wakefulness, which can be either transitory on the way to recovery from (minimal) consciousness or irreversible (Johnson & Lazaridis, 2018; Laureys et al., 2010; Laureys & Schiff, 2012). So, following this brief characterization of LIS, one might therefore be inclined to assume that consciousness is in the brain or is identical with some of its neuronal patterns⁶¹.

In philosophy, we can find some theoretical attempts to conceive “what it is like to be” in a state similar to LIS or to prove the existence of the mind or the soul as detached from the body through some speculative situations resembling LIS condition. The Arab philosopher Avicenna (980-1037 A.D.) in *The Book of Healing*, in order to prove the ontological difference between soul and body, created a thought experiment of the “floating man”, referring to a man suspended in the air, created at a stroke, perfectly developed, without the possibility to perceive his body (his limbs are separated and insensible) and the external objects. Despite this speculative and odd situation, the man would be able to think of himself without any reference to his body or to the environment, showing therefore signs of self-consciousness (L. E. Goodman, 1992). Actually, it is very hard to concretely find such kind of situation of the body, i.e. not only de-efferent but also de-afferent, but the principle is that of a consciousness

while the first detailed analysis of the syndrome was made by Plum and Posner in 1966 (Plum & Posner, 1966). However, before scientific and clinical studies, the syndrome has been described in literature (Dumas 2003/1846; Haan 2013; Zola 2001/1868). Some examples: in Alexandre Dumas’ novel *The Count of Monte Cristo* (Dumas, 2004/1846), Monsieur Noirtier de Villefort was in a situation of paralysis and he was described as “a corpse with living eyes”. In Emile Zola’s novel *Thérèse Raquin* (Zola, 2001/1868) there was an old woman that after two strokes remained paralyzed, with only “the language of her eyes”, and in the end she “was buried alive in a dead body”. It is true that living in a nearly complete de-efferentation can lead to the sensation of being “trapped in a badly injured body”, where the “true self” is “all inside” (Ostrum, 1994).

⁶¹ LIS is more frequently caused by selective brain damages, e.g. bilateral ventral pontine lesion or mesencephalic lesion (Cardwell, 2013; Gosseries et al., 2009; Smith & Delargy, 2005) that leaves completely intact the areas relative to the emergence and maintenance of consciousness, like the thalamo-cortical pathways. LIS can be also observed in the end-stage of amyotrophic lateral sclerosis (ALS) (H. Hayashi & Kato, 1989; K. Hayashi et al., 2016) and it has been reported as a transient state of the acute postinfectious polyneuropathy named Guillan-Barré Syndrome (GBS) (Carroll & Mastaglia, 1979; Friedman et al., 2003; Ragazzoni et al., 2000). In particular, in ALS and GBS the Central Nervous System remains essentially intact while the Peripheral Nervous System is severely damaged.

locked inside a body, although detached from it. Similarly, Anscombe introduced the thought experiment of the Sensory-Deprivation Tank, inside which, she argued, both exteroception and interoception are gone, but self-consciousness remains intact, as the presence of a thinking Cartesian ego without any reference with the body (Anscombe, 1981). Evans argued against this idea, claiming that self-conscious thoughts are not possible without referring, at least potentially, to the various ways that subjects have of gaining information about themselves (Evans, 1982).

Clinically, the locked-in syndrome leads to a selective de-efferent state, a condition in which the patient has no limb or face movements, i.e. quadriplegia and pseudobulbar palsy, and she/he cannot articulate speech, i.e. anarthria⁶². Usually, in the classical locked-in syndrome⁶³, the patient can maintain vertical or lateral eye movements or blinking (Goebel et al., 1971), used for communication. Generally, vertical ocular movements are more frequently spared because they depend on mesencephalic structures, while horizontal gaze on pontine structures (with some exceptions, e.g. Larmande et al. 1982). Locked-in state is anatomically described by isolated interruption of motor outflow pathways, and therefore is different from cognitive motor dissociation (CMD) (Schiff, 2015; Schiff & Fins, 2016), which instead is characterized by concomitant injuries across the cerebrum that may significantly limit ability to harness a potential communication channel (Curley et al., 2018).

⁶² Motor evoked potentials (MEPs) are generally absent (Bassetti, Mathis, & Hess, 1994b; Kotchoubey & Lotze, 2013) or severely impaired (Facco et al., 1989; Landi et al., 1994), due to a complete disruption of all or almost all descending motor pathways, but with normal cortical motor area activation (Cincotta et al., 1999). Residual MEPs may predict a good motor prognosis (Bassetti, Mathis, & Hess, 1994a).

⁶³ It is possible to distinguish between at least three different types of locked-in syndrome (Bauer, Gerstenbrand, & Rimpl, 1979). The incomplete state (ILIS), in which some minimal body movement is preserved, e.g. the astrophysicist Stephen Hawking who communicated with a computerized voice synthesizer using one finger (Gosseries et al., 2009). Classical locked-in syndrome (LIS), the most common condition described above: the supranuclear ocular motor pathways that lie dorsally are spared, therefore, vertical eye movements and blinking are intact. LIS is the clinical situation narrated in the best-seller *The diving bell and the butterfly* (Bauby, 1997) by the patient himself Jean-Dominique Bauby, communicating only with his eye. Total/complete/functional locked-in syndrome (CLIS/TLIS) is characterized by a dissociation between extreme motor disfunction, i.e. no movements at all, no behavioural evidences of consciousness, and preserved higher cortical functions detectable only through paraclinical tests. (Bruno et al., 2011). Even blinking and vertical movement of the eye are lost, e.g. no eye opening, no response-to-command related eye movements (following manual eye opening), no visual fixation or tracking and no visual startle can be observed in such case (Schnakers, Perrin, et al., 2009). Lastly, there can be cases of Fluctuating Locked-in Syndrome as a subtype of pontine warning syndrome, characterized by recurrent and limited episodes of motor hemiparesis, dysarthria and horizontal gaze palsy associated with basilar artery branch infarction (García-Espérón et al., 2016).

In LIS a general preservation of auditory, nociceptive, vibrotactile, visual systems can be observed, i.e. afferentation pathways (from environment to body to brain) are still preserved. Indeed, the patient is able to perceive the surrounding environment and purposefully interact with it and other people by means of eye movements (Gosseries et al., 2009). Possible selective disturbances may occur in single sensory systems. Anatomically speaking, in Classical LIS most of the autonomous, interoceptive, vestibular system and neuroendocrine interactions are still preserved (Lucci & Pazzaglia, 2015). Some patients could even assist to an improved sense of the inner body, i.e. interoception (Ostrum, 1994). Regarding body ownership, a questionnaire report on 44 LIS cases analyzed the level of discontinuity patients have towards their impaired body (Nizzi et al., 2012). The study shows that, even if the massive bodily change marks a clear-cut difference between pre- and post-LIS, these changes are generally well integrated in patients' experienced identity and in personal body representation (J. Cole, 2004). Indeed, even if impaired and motionless, the "new" body is still recognized as an essential part of patients' personal identity (J. Cole, 2004), suggesting that the personal body representation – and not the objective damaged body – remains a strong and meaningful component in patients' life (Nizzi et al., 2012)⁶⁴. This can be also supported by the fact that body ownership and agency does not derive merely from the efferent pathways (from internal signals in the motor cortex to the body) but principally from the afferent system (visual, tactile, proprioceptive, kinaesthetic, auditory, etc.) that permits to be aware of our own body in a spatiotemporal way (Pyasik, Furlanetto, & Pia, 2019)⁶⁵. These afferent pathways are at least partially preserved in Locked-in syndrome.

⁶⁴ Even in a progressive de-afferent – and not only de-efferent – situation the sense of the interoceptive body and the embodied situation are still important for the awareness of the body-self, suggesting that a global corporeal awareness may reflect an innate configuration within certain brain regions, representing the body as a coherent whole, rather than separately for each body part (Lucci & Pazzaglia, 2015). This global corporeal awareness (Lucci & Pazzaglia, 2015) can be linked to minimal eye movements, which are capable of actively engaging the body, e.g. seeing its position. Indeed, even in paralysis visual experience, memory and mental representation of the patient's own body can be still useful in order to maintain some degree of bodily awareness and proprioception (Conson et al., 2010; Dudzinski, 2001), even if with less accuracy in motor imagery comparing to healthy control (HC) (Conson et al., 2008).

⁶⁵ «For example, if someone else caresses my arm, I experience that body part as my own because I see and I feel the touches at the same time and in the same place. All in all, the stronger the spatiotemporal congruency among these signals, the higher the feeling of body ownership» (Pyasik et al., 2019).

Taken together, these findings highlight that in Classical LIS exteroception and interoception is generally still present with different degree of objective and subjective preservation, despite the nearly total de-efferentation of the body. This suggests the possibility to define a dissociation between efferentation and afferentation, i.e. efferent afferent dissociation (EAD), revealing that, between these two systems, only the latter seems fundamental for the maintenance of consciousness.

6.4.1.1 4E cognition challenged by LIS

On the basis of the brief analysis here presented, some of the theories from the 4E cognition are concretely challenged by the case of LIS. Indeed, the idea of a mind as an extended, embodied and interactive process between brain, body and environment is called into question by the empirical situation of a patient unable to move the body and to act within the world. Kyselo and Di Paolo brilliantly expose three approaches of embodiment, confronting them with the problems raised by LIS (Kyselo & Di Paolo, 2015). The sensorimotor approach or enactivism (Engel et al., 2013; Noë, 2004; O'Regan & Noë, 2001), the extended functionalist approach (Clark, 2008; Clark & Chalmers, 1998) and the autopoietic enactive approach (Di Paolo & Thompson, 2014; Thompson, 2007; F. J. Varela, 1997). The sensorimotor approach, in its stronger version according to which the actual moving body is essential for experience (Noë, 2004), cannot account for LIS, because it must justify the presence of cognition in complete musculoskeletal paralysis (Loughlin, 2014; Pepper, 2014). The weaker version of the sensorimotor approach sustains that bodily movements as such are not crucial for cognition, but the corresponding skills and practical knowledges a person has built up in the past, which are only the potential to perform a movement (Shapiro, 2011). This second version is more suitable to deal with LIS, however, Kyselo and Di Paolo argues that, as the stronger one places too strict conditions on the role of the body in cognition, the weaker one trivialized the role of the body, i.e. placing only the historical causal relevance of it, therefore it cannot be interpreted as a concrete theory of embodied cognition.

Differently, the extended functionalist approach includes the body just as a replaceable hardware component, therefore it does not explain the necessity of the

living body for cognition and the relation between the felt body and the functional body. On the contrary, during LIS the body could be replaced by any other extra-bodily mechanism with the same function, e.g. with Brain-Computer Interfaces (BCI) (Clark, 2008; A. Fenton & Alpert, 2008), therefore here cognition is only contingently embodied, i.e. there is not real embodiment. Instead, the autopoietic enactive approach sustains that the body is essentially situated in the world and without this relation there cannot be cognition, i.e. the neural body is not sufficient for cognition. Indeed, the body is understood not simply in a neuromuscular sense, but as a system of self-grounding identity and autonomy with specific abilities to regulate its interaction with the world (Di Paolo & Thompson, 2014). In this sense, cognition is understood in a broader way, as sense-making process of an autonomous system that tries to preserve itself as a precarious form of life within a world full of potential interactions. The body, therefore, is essential as the bearer of interaction with the world in order to regulate the whole cognitive system as an autonomous sense-maker. In LIS the patient loses the ability to perform movements, but the body is not only made of these muscular possibilities, in fact it can still engage relations with the world through sense organs. The patient is still sensitive to the world, therefore even in LIS the body can fill the role of center of an embodied identity and perspective on the world (Kyselo & Di Paolo, 2015)⁶⁶.

To sum up, LIS is really challenging for theories of embodied cognition given that the syndrome is a clear example of the possibility to have a body paralyzed, possessing at the same time cognitive abilities and consciousness states similar to healthy people. However, as things currently stands, it is not possible to definitively express a response about which is the best model for addressing consciousness in LIS, whether the brainhood paradigm or the worldhood paradigm. Indeed, the fact that the efferent pathways (from brain to body) are almost completely disrupted in this clinical

⁶⁶ In addition, we must also stress the fact that hitherto it has been discussed only of the analysis of embodied cognition in LIS, without question the relation between consciousness and body in LIS. Sensorimotor enactment states that consciousness, as well as cognition, is not something that happens to us, but it is enacted by what we do with our bodies (Noë, 2004, 2009), therefore it seems hard to explain conscious patients with LIS. According to extended cognition, consciousness is still brain-bound (Clark, 2009), therefore it is perfectly reasonable to be conscious in LIS and there should be nothing to discuss about. Instead, the autopoietic enactment has no clear-cut separation between consciousness and cognition, because they are interpreted as different parts of the same process of the whole living body within the world (Thompson, 2007; F. J. Varela et al., 1991).

situation do not imply a complete disembodiment and disembedding, since patient in LIS can still communicate with the eyes or BCI, and they can still perceive the environment and their body (afferentation), albeit passively as an object or even an obstacle (Leder, 1990). Therefore, Classical Locked-in Syndrome cannot be considered as a case of a conscious brain in the vat, since that embedment and some kind of embodiment are still preserved⁶⁷. At most we can indicate that radical theories of embodiment and enactment, in particular those that predict the need for musculoskeletal movement for cognitive and conscious activity, find it difficult to explain how patients in LIS are conscious and cognitively active.

6.4.1.2 LIS as “locked inside” but not outside of the world

On the basis of the afferent efferent dissociation (AED) described above, despite LIS leads to paralysis and de-efferentation the body can generally preserve its ability to perceive external stimuli at different levels, maintaining therefore an afferent connection with the world (at least in Incomplete LIS and LIS). The lack of motor functions does not mean that the body is irrelevant in the maintenance of consciousness, but only that the actual skeletomuscular movement is not necessary at least for the preservation of consciousness, while it could be still relevant for the formation of consciousness (as the weak version of sensorimotor enactment argues). Indeed, we have seen how a moving body situated into an environmental context with several affordances that invite actions (Gibson, 1979) can be the right condition for the emergence of consciousness. Also, I emphasise the fact that the body is still crucial as

⁶⁷ It is possible that patients in LIS do not have a minor degree of embodiment, but a different kind of embodiment. Indeed, in Classical LIS most of the autonomous, interoceptive, and neuroendocrine interactions are still preserved (Fuchs, 2017; Lucci & Pazzaglia, 2015), although on a smaller scale (Lenggenhager et al., 2012). Interestingly enough, some patient could even assist to an improved sense of the inner body (Ostrum, 1994). This could be since the contact with the external world during LIS is partially impaired, at least the subjective possibilities of agency through the world itself. For this reason, the inner body, which is usually not at the center of attention in the healthy subjects' daily life (Leder, 1990, 2018), could become the main center of personal awareness in patients with LIS. In this way, immobility and inaction could help intensifying the activation of bodily-formatted representations and generating a greater feeling of the body (Gallese, 2017); in the same way, for example, we stop and stay motionless for a while, in order to feel better ourselves.

the condition for the integration of the stimuli from the surrounding environment, by letting the brain receive and processes information⁶⁸.

Therefore, regarding interoception as well as exteroception, is not the moving body per se that is necessary for consciousness, rather the body as the means for the world-brain alignment, e.g. the afferent pathways of the body. Instead, UWS/VS patients, while potentially having the ability to move the body, they cannot be conscious due to their de-afferentation, i.e. disconnection with the world. In other words, the de-efferent person, e.g. LIS, is locked inside his body but can still integrate himself within the world, while the de-afferent person, e.g. VS/UWS, is not locked inside his body, rather is locked outside of the world.

6.4.1.3 Complete Locked-in Syndrome

However, in some cases the clinical situation may progressively worsen until it reaches what is termed Complete Locked-in Syndrome (CLIS), during which the movement of the eyes is no longer affable or is completely lost. It is characterized by changes in circadian rhythm, sleep pattern, and arousal with alterations of attentional performance, leading to decreased alertness, sleepiness and drowsiness more than the other types of LIS (Chaudhary et al., 2017; De Massari et al., 2013; Patterson & Grabois, 1986)⁶⁹. CLIS could lead to the potential and progressive disruption of some afferent pathways, e.g. skin mechanoreceptors above all since they are more susceptible to damage from atrophy and skin degeneration, but possibly even some joint and muscle mechanoreceptors (Murguialday et al., 2011). Also, auditory

⁶⁸ Regarding this point, Tallon-Baudry and colleagues, through a series of studies on the coupling of heart and stomach to the brain, recently argued that the brain's visceral monitoring of the body's interoceptive inputs is central for specifically the first-person perspective of consciousness, i.e. consciousness is intrinsically neuro-visceral rather than just being merely neuronal (Babo-Rebelo, Richter, & Tallon-Baudry, 2016; H.-D. Park et al., 2017; H. D. Park et al., 2014; H. D. Park & Tallon-Baudry, 2014; Tallon-Baudry et al., 2018). This is also in line with the phenomenological description of consciousness as intrinsically embodied.

⁶⁹ In Classical LIS's report only 5 out of 44 individuals tended to sleep or slept most of the time (León-Carrión, Van Eeckhout, et al., 2002). Sleep polysomnographic recordings in one CLIS patient with implanted ECoG-electrodes showed chaotic sleep patterns with unpredictable slow wave sleep episodes during the day (Chaudhary, Piccione, & Birbaumer, 2015; Murguialday et al., 2011). In another case, it was found no circadian rhythm in heart rate and body temperature and increased fragmentation of slow wave sleep across the day (Soekadar et al., 2013).

information processing and nociceptive sensitivity could be impaired in CLIS (Schnakers, Perrin, et al., 2009).

In this respect, as said above, LIS and CLIS can also be related to Guillain-Barré Syndrome, as a transient state of this acute postinfectious polyneuropathy (Carroll & Mastaglia, 1979; Medici et al., 2011; Ragazzoni et al., 2000). This syndrome is useful for defining some features of Locked-in Syndrome, because it is generally characterized by a progressive and rapid peripheral demyelination (Dimachkie & Barohn, 2013), leading to a clinical situation that we may call de-efferent/de-afferent state (McCullagh, 2005; Ragazzoni et al., 2000): complete paralysis, nerve conduction inexcitability, unresponsive alpha rhythms activity to nociceptive and auditory stimulation. In some cases, Sensory Evoked Potentials (SEPs) and Brainstem Evoked Potentials (BAEPs) can decrease or be absent for a transient period of time, showing alpha coma pattern or resembling brain death, although with no brain lesion observed (Friedman et al., 2003; Hassan & Mumford, 1991; Ragazzoni et al., 2000; Rougé et al., 2016). In other cases, it can be found ERPs consisting only of N100 and also sometimes of P200 components (with no presence of P300, generally a marker of cognitive ability), obtained by means of oddball auditory paradigm stimulation, suggesting presence of basic pre-attentive (as related to N100) and attentive (as related to P200) sensory detection, without semantic value awareness (as related to P300) (Ragazzoni et al., 2000)⁷⁰.

So, in Complete Locked-in Syndrome, the patient's clinical situation may get worse, showing complete de-efferentation, potential decrement of afferentation (pathways from sensory organs to the brain), cognitive abilities. These findings

⁷⁰ Interestingly, presence of minimal consciousness was detected in a patient with a particular form of CLIS behaviorally resembling UWS/VS (spontaneous ocular bobbing, no visual pursuit, no blink, no startle reflex to noise, no reaction to nociceptive and auditory stimulation, pupillary light reflexes, corneal and cough preserved) but with symmetrical theta band slightly reactive to passive eye-opening/closing (Rohaut et al., 2017). Instead, another case report of CLIS due to ALS shows prevailing slow alpha activity without reaction to stimulation, indicating therefore a possible disorder of consciousness (Berkhoff, Donati, & Bassetti, 2000). Regarding cognitive ERPs, a study focused on auditory stimulation with a distinction between passive and active event-related paradigms, i.e. between mere listening and counting conditions, highlighted P300 component in one patient in CLIS similar to healthy people, not only in passive listening, as already shown, but also in active counting after verbal instruction (Schnakers, Perrin, et al., 2009). Schnakers and colleagues argue that this may be considered as a case of Complete Locked-in Syndrome with intact auditory system but also with voluntary, therefore conscious, brain activity.

highlight that CLIS is not distinguished from Classical LIS only for the loss of eye movement. In addition to this, it can present other complications, e.g. decrease of attention, alertness, some possible worsening in terms of cognitive task and clouding of consciousness when afferentation progressively weakens. Also, the complete absence of behavioral index of consciousness makes it harder to distinguish CLIS from disorders of consciousness like UWS/VS. Differently from LIS, CLIS presents us with problems at the level of states of consciousness that I believe it is necessary to examine now also from a philosophical and phenomenological point of view. Indeed, based on what has been explained so far, the Complete Locked-in Syndrome could be a case in which the brain remains intact, while relations with the body and the world progressively fall short (de-afferentation). It is therefore a good candidate to distinguish the worldhood paradigm from the brainhood paradigm. In fact, if CLIS was a clinical situation in which a person continues to experience despite the loss of contact with the environment and the body, then the neuro-centric approach would be the most correct to describe consciousness. If, on the other hand, alterations or disorders of consciousness occur in CLIS, then this would be in favour of the neuro-ecological approach.

6.4.1.4 Environment-Body-Brain relation as Predisposition of Consciousness

This can be also noted in some cases of CLIS (due to ALS or GBS), in which the progressive loss of afferent connection with the world lead to a general worsening of alertness and clouding of consciousness and EEG alpha activity. As exposed above, the Guillain-Barré Syndrome is a paradigmatic example in order to show how a person in nearly complete de-efferentation and de-afferentation but with no relevant lesion in the thalamocortical pathways may sustain impairments of consciousness. Therefore, despite the fulminant GBS could lead to the idea of CLIS with “undisputed” presence of preserved consciousness and cognition (De Jong 2013: 142), this has not been actually demonstrated.

During my study period as a visiting scholar at the Royal Mental Health Center of the University of Ottawa (Canada), under the supervision of prof. Georg Northoff,

I had the opportunity to empirically investigate by computer analysis (Matlab) some datasets of electroencephalograms (EEG) of different states of consciousness (e.g. anesthesia, UWS/VS, MCS, etc.). In particular we investigated a group of patients affected in different degree by Amyotrophic Lateral Sclerosis (ALS), in which there are also some EEG recordings from one patient in Classical LIS and two in Complete LIS (datasets from Fomina et al., 2017; Hohmann et al., 2018). The aim of the research was to understand whether there is a difference in brain's spontaneous activity between CLIS, LIS, other ALS patients and healthy people. Indeed, according to TTC a normal intrinsic activity of the brain is a reliable marker that can indicate if the world-brain alignment is intact or impaired, and this can be associated to possible alteration of consciousness. We try to operationalise consciousness through specific markers detectable in the EEG, which can be used as indexes of presence of consciousness (Northoff, 2018b; Northoff & Huang, 2017): temporal duration of the brain's neural activity that is correlated with the ability to integrate different stimuli in the same temporal window, i.e. temporal alignment, the level of organization of the brain's activity in different temporal ranges or scales, i.e. temporal nestedness, and the spatiotemporal complexity of brain activity.

The results (papers in preparation) show that while there are no significant differences between the LIS patient, the other ALS patients and healthy people, instead the two CLIS patients present altered EEG recordings. More specifically, longer temporal windows of integration of stimuli, suggesting that the CLIS patient have difficulties in integrating different stimuli in a normal temporal window, pooling together over longer time period several different stimuli which usually would not be pooled and integrated but differentiated from each other, e.g. in the fully awake state. Also, CLIS patients show a general slowdown of the brain's spontaneous activity to low frequencies (delta-theta) and a decrease of the overall complexity, in a similar way to mild sedation or early sleep. In particular, one of the two patients in CLIS shows also fluctuations in the spontaneous activity of the brain between intra-subjective sessions. This is in line with what presented above regarding Complete LIS and the general worsening of the overall state of the patients compared to Classical LIS. In this sense, a CLIS patient affected by de-afferentation as well as de-efferentation could be subject to a potential loss of the environment-body-brain relation as a particular kind

of world-brain relation. This means that the environment-body-brain relation may be considered a necessary predisposition for the maintenance of consciousness. Of course, this is not a question of all or nothing, because experience is a matter of degree, and it is extremely difficult (and dangerous) to declare without any doubt whether someone is conscious or not in these situations. Nevertheless, if the theory is correct, this may have repercussions not only in the theoretical and scientific research fields, but also in the clinical and neuroethical fields (Fernando Vidal, 2018)⁷¹.

Returning to the neuroscientific theories of consciousness previously analysed, according to IIT it is perfectly possible that the brain of a CLIS patients with no relation to the world nor to the body can produce integrated information, i.e. conscious experience. This goes against these results (complexity in CLIS' brain decrease) and the idea that a person deprived of external stimuli and proprioceptive input cannot elicit primary consciousness (G M Edelman, 1989). Following GNWT, consciousness can be preserved during LIS or CLIS, as the high-order level of neuronal processing are preserved and can still elicit P3b. Even in a severe CLIS diagnosis with complete detachment from the world, the cortically intact brain could keep synthesizing and spreading information all over the workspace network, especially in the prefrontal and parietal cortical regions (Baars et al., 2013). Instead, according to TTC consciousness is not identifiable with integrated information or with the spontaneous activity network as such, rather as the elicited process through the brain's spatiotemporal alignment to the world, i.e. world-brain relation (Northoff, 2018b).

⁷¹ «At present, we cannot find out whether or not a presumed TLIS [Total Locked-in Syndrome] patient is conscious, not because there are no 'necessary and sufficient conditions', but because a fundamental uncertainty about what it means to be a human being in all respects arises as soon as the familiar outward criteria of our ordinary discourse are no longer applicable. Confronted with the TLIS patient, we are uncertain of his state of consciousness and suspect that 'new' neurophysiological signs of consciousness might be unreliable or even deceptive. And this is not because the behavioral criteria we applied before (and in non-TLIS cases) were better or more reliable, but rather because these 'ordinary' criteria were (and are) embedded in that complete (though surely not completely correct) conception of man which is part of the '*Lebensform*' we actually live. [...] If Knowledge of other minds presupposes an attitude towards human beings (as persons, as conscious beings) in general and thus a rich and homogenous image of man, then knowledge of a TLIS patient's consciousness will only emerge within a (utopian) complete scientific image of man, which some of us expect from our dramatically evolving neuroscience. But since this image is dreams of the future, we are left with familiar uncertainty and doubt - and with the old pragmatical advice to treat potential TLIS patients as if they are conscious» (Kurthen et al., 1991, p. 77).

As long as the neuro-ecological continuum is maintained, consciousness is preserved, but this does not mean that if the spatiotemporal alignment between brain and world is disrupted, consciousness will completely disappear at once. This would imply a theory of consciousness grounded exclusively on the stimulus-induced activity, but this is not the case. Indeed, it is important not to confuse the world-brain alignment with the stimulus-induced activity as such. In this respect, the spatiotemporal structure of the world-brain relation, as the enabling condition for consciousness, could be understood like the sun for the photosynthesis process: it is not the single photon/stimulus that determines the presence or not of photosynthesis/consciousness, but whether the leaf/brain, in relation with the plant/body is nested within and aligned with the sunny environment/spatiotemporal structure. Explicitly, this means that the elimination of all external stimuli would not lead immediately to a loss of consciousness, but to the progressive degradation of it.

Insofar, following this metaphor, within the world-brain relation the body is necessary not as such, but as the spatiotemporal context through which the relation itself can be sustained. Therefore, consciousness is not merely stimulus-based, rather also the product of the spatiotemporal alignment, through the resting state's degree of phase locking to the onset of the environmental stimuli (Northoff, 2014c). In cases like LIS (and CLIS in particular), as seen, the de-efferent/de-afferent body may possibly lead to a progressive degradation or clouding of consciousness, initially with the reduction of the experienced world to the body alone (interoception, proprioception, etc.): without any connection to the world, the only remained content might be inner body sensations, memory, mental imagery or the mastery knowledge of lost sensorimotor skills (De Massari et al., 2013; Shapiro, 2011). However, according to TTC we may assume that the pervasive and constant absence of relations with the world would progressively lead to an impairment of consciousness. For these reasons, rehabilitation and multisensory stimulation programs may be useful in order to recover at least a partial connection with the body and the environment, so that the world-brain relation can be re-established (e.g. in the case of a patient in LIS due to a brain tumour that, after years in unconscious state and misdiagnosed as end-stage vegetative state, he regained consciousness thanks to rehabilitation programs; Lukowicz, Matuszak, & Talar, 2010).

6.4.1.5 CLIS as potentially being-out-of-the-world

Phenomenologically speaking, these results may suggest that the patient in LIS that enters into complete locked-in state and progressively loses not only the motor skills but also the receptive ones towards the body and the environment, may find himself in a situation of detachment from the world that can hardly become total (some sensorial access can remain intact, as we have seen), but that nevertheless produces an alteration of the relationship with the world. The stimuli coming from the environment are difficult to discriminate without a solid and constant environmental and bodily context that acts as a background, consequently also the complexity of the content of consciousness and the general attention and waking capacity decrease. In this sense, the CLIS patient may lose her/his embodied-being-in-the-world, hence one may speak of a potential “being-out-of-the-world” as the description of this situation in which a person is not able to engage and be spatially and temporally connected with the world anymore.

In this case, the “being-out-of-the-world” is associated with the impairment of the environment-body-brain structure of consciousness (world-brain relation), due to morbid form, i.e. the physiological consequences of Complete Locked-in Syndrome. However, there may be also cases in which the person, although still in partial relationship with the world (e.g. through preserved hearing and partially preserved sight), may find himself in a “being-out-of-the-world” condition due to a malaise. It is the case of Julia Tavalaro, a woman misdiagnosed as VS for six years, while instead she was in LIS and conscious of the surrounding environment. During those years she was considered as a vegetative state and treated as if she was not a human person, until the doctors realized that she was aware and that she could experience the world around her. In this sense, despite being in relation with the world, i.e. her world-brain relation was intact, from her later written testimony⁷², we can understand that during the years

⁷² «“Can you close your eyes, Mrs. Tavalaro?” With these words, *I’m shocked back into reality*. This is *no dream*. *I’m actually being spoken to*. I close my eyes. I open them and see Arlene’s face. “Can you blink twice?” I do it. Silence fills the space between us. Her face shows shock and grief and happiness at once. In the previous six years, no one had thought to ask me these simple questions. “Okay, Mrs. Tavalaro. I’d like you to respond with eye movements. Can you move your eyes up, like this?” She rolls her eyes toward her forehead. I watch her do this. Then, with a quick movement of my eyes *I feel my*

of misdiagnosis she perceived herself as separated from the others, incapable of communicating her being-in-the-world to the other subjects, who instead looked at her as an object among the objects (see the objectifying power of the Look of the Other exposed by Sartre, 1956). She was in relation with the world neurologically and physiologically, but she was partially out-of-the-world socially and existentially speaking. Thus, this confirms that embodied-being-in-the-world and world-brain relation are associated as diffractions of the same experiential phenomenon, still they describe different aspects of it.

About the possibility of consciousness without a world, as Moran and Cohen pointed out, in Ideas I § 49 Husserl discusses the thought experiment of the ‘annihilation of the world’ (*Weltvernichtung*)⁷³. According to Husserl, «it is possible to imagine the flow or stream of the worldly experience being entirely disrupted to the point where all is chaos, but it is impossible at the same time to think away pure consciousness. If the entire experience of the harmonious flow of the world were disrupted so that it became a meaningless chaos, the experience of the ego would be profoundly modified and altered, but it would still exist, even if its flow of temporal experience was chaotic. On this basis, Husserl concludes that pure consciousness is absolute and independent of all objective being» (Moran & Cohen, 2012, p. 33). As said here above, the account of consciousness as embodied-being-in-the-world and world-brain relation does not imply the constant presence of an actual, solid, physical

mind rise from the ocean depths of pain. For the first time in six years, I feel whole» (Tavalaro & Tayson, 1998, p. 121, emphasis mine; see also Vidal, 2018).

⁷³ «Nevertheless, in that case it could be that, to some extent, crude unity-formations become constituted, transient supports for intuitions which were mere analogues of intuitions of physical things because quite incapable of constituting conservable “realities,” enduring unities “which exist in themselves, whether or not they are perceived.” [...] let us recall the possibility of non-being of everything physically transcendent: it then becomes evident that while the being of consciousness, of any stream of mental processes whatever, would indeed be necessarily modified by an annihilation of the world of physical things its own existence would not be touched. Modified, to be sure. [...] But that does not mean that other mental processes and concatenations of mental processes would be excluded. Consequently no real being, no being which is presented and legitimated in consciousness by appearances, is necessary to the being of consciousness itself. Immanent being is therefore indubitably absolute being in the sense that by essential necessary immanent being *nulla “re” indiget ad existendum*. In contradistinction, the world of transcendent “res” is entirely referred to consciousness and, more particularly, not to some logically conceived consciousness but to actual consciousness. [...] consciousness considered in its “purity” must be held to be a self-contained complex of being, a complex of absolute being into which nothing can penetrate and out of which nothing can slip, to which nothing is spatiotemporally external and which cannot be within any spatiotemporally complex, which cannot be affected by any physical thing and cannot exercise causation upon any physical thing – it being presupposed that causality has the normal sense of causality pertaining to Nature as a relationship of dependence between realities» (Husserl, 1982, § 49).

world to ensure the presence of consciousness. Indeed, it is not the material status of the world that is important for the maintenance of consciousness, rather its spatiotemporal structure, so that one may think about an environment in virtual reality with coherent spatiotemporal structure as a possible substitute for the actual surrounding environment⁷⁴. However, according to Sartre's phenomenology, a total absence of a world (in any sense) would lead not to a pure and absolute presence of consciousness, as Husserl would say, but to the consequent annihilation of consciousness, as consciousness is nothingness *per se*, born supported by something which is not itself. According to the neuro-ecological approach, on the other hand, a total absence of any relation with the world would probably lead the brain to an alteration of its intrinsic and experience-dependent activity.

Now, someone might argue that there already are in literature cases of patients misdiagnosed as UWS which they were actually in CLIS but still conscious, and the physicians had the possibility to communicate with them through neuroimaging. I am referring to the famous cases of brain-imaging communication performed by Adrian Owen and colleagues. During the first pioneering experiment they put a woman diagnosed as being in a vegetative state into a functional magnetic resonance imaging scanner, asking her to imagine she was playing a game of tennis, and then to imagine she was moving in her house. They showed that very different patterns of fMRI activity emerged depending on the kind of the asked activity and both the patterns were also indistinguishable from those seen in healthy participants. This means that the woman was able to understand and follow the task of mental imagery in a very similar way as the healthy controls did, revealing therefore a covert conscious activity despite the absence of any behavioural responsiveness, enough to wrongly diagnose her as vegetative state (UWS/VS) (Owen et al., 2006). Furthermore, in the last 13 years they have refined the method for assessing residual or covert preserved consciousness, in order to minimize and avoid reverse inferences supporting the command-following task with other tests, e.g. multiple repetition of the task, comparison with task-evoked activity in mildly sedated healthy volunteers, and specific tasks related to the

⁷⁴ For example, see the use of Virtual Reality as a tool for intervention in children with autism spectrum disorder (Mesa-Gresa et al., 2018). For a neuro-ecological point of view this can be seen as an attempt to perform new alignments for people that generally present an altered connection with the environment as well as the social context.

autonomous will of the patient, which was explicitly instructed not to follow the command of performing motor imagery (Cruse et al., 2011, 2012; Owen, 2013). This method has been crucial for many patients wrongly diagnosed as UWS/VS, while they were actually cases of minimally conscious state (MCS), cognitive motor dissociation (CMD) or even complete locked-in syndrome, since this people were not even able to move the eyes to communicate with the clinicians (Formisano, D'Ippolito, & Catani, 2013). So, this means that among these cases of misdiagnosis there might be also patients in CLIS that preserve consciousness even in this severe clinical situation.

This fact seems to disprove the hypothesis exposed above, about the fluctuations of consciousness in CLIS or, more generally, in absence of connection with the world. However, this is not the case, for two reasons. The first one is that being able to perform mental imagery might be a case of “converted UWS” with preserved island of consciousness (Marino et al., 2017). If that is the case, there would be no problem for the hypothesis of CLIS patients being subject to fluctuations of consciousness, also because the possibility of fluctuation and clouding of consciousness is different from claiming the complete absence of awareness in those patients. As said before, as it is ethically concerning to diagnose states of unconsciousness only on the basis of behavioural scales, in the same way it is not simple to provide a diagnosis of full recovery of consciousness only on the basis of some completed tasks indirectly seen through neuroimaging. However, let us take in consideration those cases in which the fMRI scanner shows not only preserved mental imagery ability but also intentional communication ability (e.g. using different motor imagery tasks as yes/no answers for specific questions; Monti et al., 2010), it seems that there can be CLIS patients with preserved consciousness; but even in this case the hypothesis remains untouched. Indeed, the fact that clinicians can assess consciousness in CLIS patients following mental imagery and communication task means that such patients are able to receive and process external stimuli (at least the auditory ones) and this is not in contradiction with the hypothesis of the fluctuation of consciousness, on the contrary it indirectly proves that if some channel of connection with the world is intact, even consciousness can remain preserved. Differently, in the case of CLIS because of late stage ALS, not only efferentation but also afferentation could be progressively weakened, while the patients with brain trauma misdiagnosed

as UWS/VS could present complete motor disabilities but preserved sensory pathways. Moreover, I would stress again the fact that having fluctuation of consciousness does not mean total lack of awareness, as in UWS, so that the patient in CLIS tested through the tasks of mental imagery under fMRI could show “spikes” of consciousness, facilitated by the task itself that entertains and reconnects her/him to the world.

Taken together, CLIS is a paradigmatic example of potential impairments and fluctuations of consciousness associated with the disruption of the environment-body-brain relation. This disruption, instead of being caused by a brain damage as in cases like coma or UWS/VS, may be caused by the loss of alignment between the world and the brain, with the brain also presenting an altered spontaneous activity but with no actual brain lesion or functional problems. In this regards, classical neuro-centric approaches find difficulties in explaining how this may happen. Differently, a neuro-ecological approach based on a concept of consciousness as world-related and of the brain as experience-dependent, it may be useful to unravel the doubts about this rare syndrome.

6.4.2 Dreaming: Disconnected Consciousness?

As a final point, I would like to present a question concerning a possible counter argument to consciousness as embodied-being-in-the-world and world-brain alignment. In fact, someone could now reply that, without calling into question rare and critical situations such as Locked-in Syndrome, one can find simpler cases of experience disconnected from the environment and from the body, that is, the dreaming state (Loorits, 2018). During dreams we are somehow conscious of something even if it is not actually in front of us and our bed, and we can feel our body also in different ways than the usual (like flying, etc.). Basically, the question is: when we are fall asleep, where do we go? Generally, the dreaming state is considered a state of “disconnected consciousness”⁷⁵. Hence one might argue that during dream the brain can still produce perceptual conscious experience on its own (Tononi, 2009) as a sort

⁷⁵ «Dreams are a most remarkable experiment in psychology and neuroscience, conducted every night in every sleeping person. They show that our brain, disconnected from the environment, can generate by itself an entire world of conscious experiences» (Nir & Tononi, 2010, p. 88).

of “Locked-in Syndrome of the brain” (Northoff 2014: 372), undermining the idea of the necessity of the relation to the world and arguing for the localization and/or production of consciousness inside the brain (Loorits, 2018)⁷⁶.

First, it should be noted that not all stimulus-interaction is suppressed during sleep and dream, since the rest-stimulus interaction within the body is still present, suggesting that interoception plays the role of exteroception during dream (Northoff, 2011)⁷⁷. But for argument's sake, we assume that during the dream the brain is really isolated from the rest of the body and from the environment. According to the neuro-centric perspective this would be the proof that contradicts any neuro-ecological position or that generally provides for the conditions of embodiment, embedment and enactment. However, I will now try to explain how the dream poses no problems to the neuro-ecological approach and to a description of consciousness according to the worldhood paradigm.

For a neuro-ecological perspective, dream is not an exception because the world-brain relation is not restricted only to the stimulus-induced activity, i.e. the

⁷⁶ «We are conscious of our thoughts, which do not seem to correspond to anything out there; we can also imagine things that are not out there. When we do so, sensory areas can be activated from the inside (Kosslyn et al. 2001), though there are some differences (Amedi et al. 2005). Also, stimulus-independent consciousness is associated with its own patterns of activation within cortex and thalamus (Mason et al. 2007). During dreams, we are virtually disconnected from the environment (Hobson et al. 2000)—hardly anything of what happens around us enters consciousness, and our muscles are paralyzed (except for eye muscles and diaphragm). Nevertheless, we are vividly conscious: all that seems to matter is that the thalamocortical system continues to function more or less as in wakefulness, as shown by unit recording, EEG, and neuroimaging studies performed during rapid eye movement (REM) sleep, when dreams are most intense (Maquet et al. 1996). [...] Neurological evidence also indicates that neither sensory inputs nor motor outputs are needed to generate consciousness. For instance, retinally blind people can both imagine and dream visually if they become blind after 6–7 years of age or so (Hollins 1985; Buchel et al. 1998). Patients with the locked-in syndrome can be almost completely paralyzed, and yet they are just as conscious as healthy subjects (Laureys et al. 2005) and can compose eloquent accounts of their condition (Bauby 1997)» (Tononi & Koch, 2008, p. 240).

⁷⁷ In fact, it has been shown that during REM sleep the phase of delta/theta frequencies is locked to the onset of the heartbeat in a way that resembles the awake state, suggesting the relevance of the spatiotemporal body-brain alignment for consciousness (Northoff & Huang, 2017). Also, some stimulations during sleep can infiltrate and appear within the dream, e.g. bodily pains, electric pulses, pressure, etc. (Windt, 2015). Therefore, in spite of a more disrupted or altered alignment with the environment during dreams, we are not completely locked-out of the world. So, differently from coma and UWS/VS, even during dream there is no complete detachment between brain and the world, especially from the body. ERPs reflecting excitatory processes have been observed during different sleep phases, suggesting preservation of some sensory discrimination capabilities (Coenen, 2012; Cote & Campbell, 1999; Muller-Gass & Campbell, 2014). Yet, in REM sleep these external stimuli are not well-encoded as during the awake state, because they may compete with endogenous activity (Andrillon et al., 2016; Colrain & Campbell, 2007), therefore it can be said that there is no concrete rest-(extero)stimulus interaction during sleep/dream (Northoff, 2011).

actual sensory stimulation in particular time and space points. Indeed, during REM sleep the increased brain spontaneous activity can structure its rest-rest interaction (its own intrinsic activity without induced activity from the external environment) in a very similar way to exteroception (Hobson, 2009; Northoff, 2011, 2014c). This suggests that the world-brain relation on which consciousness relies is not purely constituted by presence of actual environmental interaction, e.g. as in the sensorimotor enactivist theory of consciousness (Noë, 2004), but also by the resting-state activity with its spatiotemporal pattern, which can virtually substitute and integrate the missing environmental stimulation, as a simulacrum of the world⁷⁸. However, this does not imply a neuro-centric view, because the resting state's spatiotemporal structure is not independent from the environment, actually it is shaped by it and contained within the larger spatiotemporal scale of the world (Northoff, 2018b).

6.4.2.1 Hallucinations and Dream Worlds

From a phenomenological perspective, I maintain that the relationship between dream and reality is completely different from what is supported by the neuro-centric positions. In chapter four I mentioned that often conscious reality, being a brain production according to neuro-centrism, is considered a “controlled or shared hallucination”, in other words wake state is a particular case of hallucination in which we process the external neutral stimuli and we transform them in a sharable reality. In this sense, the dream would simply be a more subjective and private kind of hallucination. This position is wrong because it is based on the idea that consciousness can be produced independently of the rest of the body and the environment, going as far as to claim that the reality in which we live is a shared hallucination. If this were true, we should first understand how we create an entire reality through the brain and how, at the same time, we render hallucinations stable, shared and homogeneous (the conscious reality), given that, as such, it is an act of imagination that is mistaken for

⁷⁸ For example, as also exposed by Tononi and Koch (2008) retinally blind people can both imagine and dream visually, nevertheless only if they become blind after 6–7 years of age. This is useful for explaining how dreaming consciousness is not a pure production of the brain but comes from the previous alignments and adaptations of the brain to the world, so that congenitally blind people cannot, of course, imagine colours, while instead people that go blind later still can imagine and dream colours.

perception, and how we communicate with others in the same shared world, extremely complicating the problem of the existence of other subjects and other minds.

Thus, the issue needs to be seen from another perspective, applying also the principle of Ockham's razor. It is not that all conscious states are hallucinations, i.e. creations of the brain, on the contrary, our waking states are strongly anchored in the world, a world that offers space-time coherence and stability and therefore allows us to have generally coherent and stable experiences⁷⁹. This does not always happen, and when the relationship with the world is altered there may be cases of hallucination, as with schizophrenia. Indeed, schizophrenia is described as a disturbance of patient's being-in-the-world, in the form of the alteration of self-other relations and of the relation with patient's own body (alienation)⁸⁰ that contemporarily disrupt the normal relation with the world (Bizzari, 2018b; Fuchs & Röhrich, 2017; Grosz, 2015; Noel et al., 2017). This interpretation of schizophrenia has also found confirmation through neuroscientific experimentation, suggesting that hallucination are cases of misalignment from the world. For instance, it has been showed that patients with schizophrenia are not able to properly align and adapt their spontaneous brain activity to the onset of the tones presented in the experiment, suggesting that schizophrenic patients remain unable to properly link their auditory cortical neural activity to the external auditory stimuli (phase entrainment or alignment) (for details about the experiment see (Northoff, 2018a). Also, these results suggest that schizophrenia may imply a sort of "reversed entrainment or phase alignment" to internal rather than external stimuli, leading to an impairment of the world-brain relation⁸¹. Therefore,

⁷⁹ One entailment of the extended model of visual phenomenal character is that the stability of the perceived visual field is a function of the stability of the visual world — the world towards which the exploratory perceptual activity is directed. Without the stable world to hold our perceptual activity together, we should predict that any quasi-perceptual activity that is not directed towards a similarly stable world should lack the consistency, coherency and stability of genuine perception. And this, of course, is precisely what we get in the case of dreams and hallucinations» (Rowlands, 2002, p. 177).

⁸⁰ «A variety of bodily experiences occur at the onset of schizophrenia, consisting in experiencing one's own body as an object lived in the third person. Thus, a patient says: "I am no longer myself (...) I feel strange. I am no longer in my body, it is someone else; (...) I hear my voice when I speak, but the voice seems to originate from some other place." When he does something, he has a feeling of observing his actions as a witness, without being actively involved: "One might think that my person is no longer here (...) I walk like a machine;" (Parnas et al., 2003)» (Pérez-Álvarez et al., 2016)

⁸¹ «Such "reversed entrainment or phase alignment" to internal (rather than external) stimuli entails that there is spatiotemporal detachment (rather than spatiotemporal attachment) between world/environment and brain in the spontaneous activity. This is reflected in the decoupling of experience from the ecological context, e.g., the world and results psychopathologically in auditory hallucinations (and probably delusions)» (Northoff, 2018a, p. 85).

according to this approach, it is not the waking state that is a particular case of hallucination, rather the opposite, it is the hallucination that is a particular case of consciousness of reality (and as we saw in the previous chapter a case of imaginative consciousness, completely different from the perception of world).

What about the dream? Similarly, the dream is a particular case of consciousness in which the relationship with the world is altered, in the sense that the environment-body-brain relation is weakened and even the embodied-being-in-the-world changes considerably. This does not mean that consciousness is possible without a world. First, the dreaming state is a temporary and reversible state of experience that is produced starting from our experiences with the world, as seen above. Even vivid dreams are different in their phenomenological structure compared to perception of the world, insofar as while we are sleeping generally we do not realize we are dreaming and the “reality” in which we live seems absolutely possible, but once we wake up, if we remember the dream, we immediately realize the absurdity of some details (dream bizarreness⁸²). To use the Sartrean terminology, this happens because during the dream the consciousness has lost its openness to the world and becomes a pre-reflexive non-positional consciousness of itself as imprisoned and fascinated by its own productions⁸³. Moreover, during dreams there is no stable world that offers a solid

⁸² In bizarre dreams there can be internally distorted or contextually incongruous element, i.e. «an element which has a feature that does not belong to it in waking reality or which appears in a context in which it would not appear in waking reality. Examples: My room is much bigger than it is in reality [Internally Distorted Place]. There is a swan in my backyard. [Contextually Incongruous Animal]». Exotic element, i.e. «an element which is highly unlikely to occur in the dreamer's waking reality, but the occurrence of which is nevertheless possible in principle. Examples: I was in a tropical jungle [Exotic Place]. I met the Secretary General of the United Nations [Exotic Character]. Impossible elements, i.e. An element or a feature of an element the existence or occurrence of which is not possible in waking reality. Examples: My dead grandfather came to visit us [Impossible Character]. I was flying above the city like Superman [Impossible Action]. We were on an alien planet [Impossible Place]». Vague elements, i.e. «an element or a feature of an element whose identity or precise nature is indeterminate, unknown, or obscure in a way which does not occur in waking life. Examples: I was in a place which seemed to be my room but at the same time it was also somehow a submarine [Indeterminate Identity of Place]. I was in a city but I do not know or remember which one [Vague Identity of Place]». Discontinuous elements, i.e. «an element which is temporally discontinuous; it suddenly and unexpectedly appears or disappears or is transformed in the dream. The Discontinuity score is assigned to the element as it is after the change. Examples: I see a horse [Non-Bizarre Animal] that rises to stand on two legs and simultaneously it changes partly into a human being [Incongruous Person] [Discontinuous Person]. I look at a map full of small islands. Suddenly I am on one of the islands [Non-Bizarre Place] [Discontinuous Place]» (Revonsuo & Salmivalli, 1995, pp. 174, 175).

⁸³ «Certainly, my consciousness, when awake, is characterized by its ‘being-in-the-world’, but precisely because this ‘being-in-the-world’ characterizes the relation between consciousness and reality, it cannot be applied to the consciousness that dreams. A consciousness cannot ‘be-in’ an imaginary world, unless it is itself an imaginary consciousness. But what is an imaginary consciousness if not a certain object for a real consciousness? To tell the truth, a consciousness that dreams is always nonthetic

background for our experience rather every dreamed object of dreaming consciousness is a stand-alone “dream world”⁸⁴. so when we are dreaming we usually recognize places, things and people from single details and not from a rich environmental context which instead we reconstruct after waking up, through reflection (“I remember I was in my bedroom, at least I remember that I was in front of my bed and my desk”).

Taken together, dreams and hallucinations pose no problem either to the neuro-ecological approach or to the description of consciousness according to a worldhood paradigm. From a neuroscientific perspective, it is risky to say that during the dream the brain is completely detached from the rest of the body and from the world, moreover the contents of the dream experience derive from the adaptation and alignment of the brain to the body and to the world. Phenomenologically speaking, unlike the brainhood paradigm, it is easier to explain the dream according to a worldhood perspective, as a particular case of imaginative consciousness, compared instead to considering the perception of the world as a particular case of hallucinated consciousness. In this sense, as Sartre says, dreaming is an exceptional case in which we can understand what it would be like to live in a reality created by us (brain-body-environment) instead of by the world itself (environment-body-brain), i.e. incoherent, unstable, unconstrained, deficient, inaccurate and chaotic: «The dream is not fiction taken for reality, it is the odyssey of a consciousness dedicated by itself and in spite of itself to building only an unreal world. The dream is a privileged experience that can help us to conceive what a consciousness would be like that had lost its ‘being-in-the-world’ and had, at the same time, been deprived of the category of the real» (Sartre, 2004a, p. 175).

consciousness of itself as being fascinated by the dream, but it has lost its being-in-the-world and recovers it only on awakening» (Sartre, 2004a, p. 170).

⁸⁴ «We have seen above that imaginary space and time are given as internal qualities of the imaged thing. It is necessary to make here an analogous observation: the ‘worldliness’ of the dreamed image does not consist in an infinity of relations that it sustains with other images. It is simply a case of an immanent property of the oneiric image; there are as many ‘worlds’ as images, even if the sleeper, passing from one image to another, ‘dreams’ that they remain in the same world. It would therefore be advisable to say: in the dream, each image is surrounded by a worldly atmosphere. But for greater convenience I will use the expression ‘dream world’, since it is in current use, and simply warn not to take it unreservedly» (Sartre, 2004a, p. 167).

Conclusion

A genuine glimpse into what [consciousness] is would be the scientific achievement, before which all past achievements would pale.

James, 1892/1984, p. 401

Let's go back on the first paragraph of the introduction. I was sitting on my desk writing on my laptop and the world constantly manifested itself to me through different ways and appearances. My positional consciousness, mainly directed towards the words of the text on the screen, from time to time was engaged by other objects inserted in the background of the room or it revealed myself as an object of reflection. The colours around me were not genuine artefacts of my neuronal activity, as if they were sitting somewhere in the brain, but they were part of the way in which the bubbles of the lava lamp and the book on the left of my hodological space manifested themselves to my sight, insofar as they trace a relationship that goes from the environment through my body, in harmonic interaction with the activity of my brain. The fingers of my hands moved quickly on the keyboard, but not as if they were a different being from my consciousness, it was unquestionably me writing even though I was not directly focused on my hands but on the words-which-had-to-be-written. Indeed, the movements of the hands towards the affordances offered by the keyboard (every single key reveals a precise meaning to my fingers) remained at the level of pre-reflective consciousness of myself writing those words. Except that, at one point, while my hands were non-positionally moving towards the inviting affordance offered by the handle of the mug, one of the fingers felt the sharpness of the thorns of the succulent plant, and immediately my body was indicated to me as being-in-pain by the world. I was forced to take off the headphones that misaligned me from the rest of the surrounding noises thanks to ambient and whispering sounds by which I was deeply entertained, but after that my consciousness of the world has been almost completely captured by the annoying buzz of a mosquito, leaving everything else to the pre-reflective background.

With this research I wanted to explore and discuss how we are conscious of the world in the way we are every day. The paragraph above is an attempt to read – quite superficially I must admit – the same situation described in the introduction through the embodied-being-in-the-world condition and the environment-body-brain alignment. All the concepts used above for the description come from what was discussed in the previous chapters. It seems now appropriate to conclude with a brief summary of what has been discussed in these chapters.

- **The phenomenon of consciousness is the condition of possibility for the manifestation of other phenomena**

First, consciousness is not merely a complex problem, is a hard problem in the sense that it resists the usual attempts at physicalist reduction of natural sciences. In this regard, I have exposed in the first chapter how experience cannot be treated as a normal natural phenomenon, without necessarily having to invoke spooky dualisms to justify its existence. Several issues have been exposed and discussed, such as the nature of qualia, the phenomenal character of experience and the meta-problem of consciousness (i.e. why it seems so special to us), but all of these are parts of what Rowlands calls “objectual consciousness”. The real theme of this research, instead, is mainly focused on the so-called “actual consciousness”, the very same act of directedness towards the objects of the world (objectual consciousness itself included), i.e. the epistemic access to the world. This is why it seems we have a meta-problem of consciousness, given that experience is not merely a phenomenon among others in the world, but also the condition of possibility of every other phenomenon manifested to us. In this way, if we consider consciousness as the very act of directionality towards the world, the illusionist positions lose strength, since although it is possible that the experience of the world is an illusion, the act of consciousness remains irrefutable even for the Deniers, otherwise they fall into the performative contradiction.

- **“Consciousness” is “being aware of something”: intentionality, phenomenality, self-awareness**

I argued that the distinction between objectual and actual consciousness is not generally considered by philosophy of mind and cognitive science, so instead of starting from a given description of consciousness, I started from the way we use the

concept of consciousness every day, analysing it through Kripke's direct reference theory. From a non-descriptive perspective, the term "consciousness" (as actual consciousness) is a rigid designator, because it necessarily refers to the act of being aware of something in every possible world (even where consciousness is an illusion). Differently, objectual consciousness is made by the set of possible descriptions (both objective and subjective) that can characterize a conscious state, and therefore is contingent because it may be different in some possible world, therefore "consciousness" and "brain activity x" can be at most one case of contingent identity. So, rather than focusing on contingent relation between consciousness and brain I argued that the best way for providing a provisional conception of consciousness was through the analysis of consciousness as "being aware of something". Then, from the term, I passed to the phenomenological description of what it is like to be aware. This first phenomenological analysis outlined three main characteristics that are intrinsic to the concept of consciousness as "being aware of something": intentionality, i.e. consciousness implies something which is not itself, phenomenality, i.e. the fact that there is something it is like to be conscious, and pre-reflective self-awareness, i.e. we understand the meaning of "consciousness", as seen above, not necessarily because the fact that any conscious state of something intrinsically involves a form of awareness of itself as being conscious (not necessarily a reflective, cognitive, representational or high-order self-consciousness).

Someone may argue that these are not necessary features of the conscious act, using cases like the so-called "unconscious experiences", however I discussed some of these examples: the refrigerator light problem, the long-distance truck driver and the case of blindsight. Using Sartre's analysis of the pre-reflexive non-positional consciousness in *The Transcendence of the Ego*, and examples of bodily awareness, it turned out that from the fact that sometimes we do not positionally perceive some stimulus from the environment or the body does not imply that we are not completely aware of it, for example at a pre-reflective non-positional level. Therefore, the examples of unconscious perception cannot be used as cases of intentional consciousness without self-awareness or phenomenality.

- **Neuro-centrism and brainhood paradigm lead to inadequate and incoherent accounts of experience**

One of the main points of this research is represented by the critical analysis of contemporary neuroscience of consciousness, since nowadays consciousness is not just a concept used in phenomenology and philosophy of mind, but as we have seen, neuroscience has become progressively relevant in recent studies on consciousness. Maybe surprisingly, I argued that neuroscience relies principally on the mind-based approach and its consequent radicalization, i.e. brain-reductive approach, because they still start from the categorization of the mind as the model from which the brain can be investigated. In particular, the search for the neural correlates of consciousness relies upon the idea that consciousness is somewhere in the brain or at least is the product of the neural activity. This view usually leads to drastic approaches towards consciousness implying the reduction (Dennett), the transmutation (Churchland) and the elimination (Irvine) of the concept within the ontology of contemporary science. However, I demonstrate how this kind of approaches is inadequate and inconsistent; inadequate because it is focused merely on the objectual consciousness, or better on its empirical conditions, moreover reductive and eliminative approaches towards consciousness often are based on methodological errors and epistemological fallacies: myth of localization, mereological fallacy, fallacy of misplaced concreteness; incoherent because, although almost all neuroscientific theories claim that consciousness is a natural phenomenon completely explicable within a monistic and materialistic domain, the root of the so-called brainhood paradigm actually derives from Cartesian dualism. What differentiates it from Cartesianism is rather the shift of the focus from the mind-body dualism to the brain-body dualism, anyway the structure remains vaguely the same.

- **Neuroscience and philosophy must escape the Appearance-Reality Fallacy**

Nevertheless, neuroscientific findings are extremely interesting and they could be crucial to a deeper understanding of how we experience the world, but I argued that, in order to do this, a change of perspective is also necessary, to get out of a typically Cartesian mind model that was thought to be eliminated but still influences philosophical and scientific studies of consciousness. More than simply Cartesian, this paradigm generally derives from the persistent modern paradigm that has divided the world into primary and secondary qualities (Galilei, Locke, etc.) and which consequently produced an ontological division between reality and appearance. This

ontological division, which I believe is still rooted in contemporary philosophy of mind and neuroscience, is what I called “Appearance-Reality Fallacy”. Indeed, when I speak of modern roots of cognitive sciences, I was not referring merely to the neuroscience and its neuro-centric paradigm, but also to some of the theories of consciousness from the Anglo-American philosophy of mind, that I would call “mentalistic” or “mento-centric”, since they fall into very similar methodological and epistemological problems as for the neuro-centric ones, e.g. mereological fallacy, hypostatization of psychic objects, illusion of immanence, and so forth. In this sense, I critically analysed the fact that both the physicalist and mentalistic perspective argue for the existence of two series of objects in the world, the apparent ones (subjective, erroneous, malleable, unreliable) and the real ones (empirically objective, or necessary a priori, true, unchangeable).

- **Consciousness must be approached by an ontological, epistemological and methodological pluralism**

However, consciousness, or our experience, can hardly be adapted to such a scheme, as it is neither a simple set of subjective and qualitative reactions to objective stimuli, nor an underlying reality that is more real than others. Consciousness is not merely a contingent point of view nor of non-physical object with some undetectable features nor only a peculiar way of knowing the world, rather I argued that it is the revealing of a series of phenomena, which reveal themselves as appearing, with no hidden reality behind but through various modes of manifestation (physical, perceptual, psychological, cognitive, physiological, etc.). Thus, if we want to go beyond the problematic project of a pure “objective”, neuro-maniac or mentalistic science of consciousness, I proposed to consider the possibility of alternative ontologies regarding the world and consciousness, i.e. a different way to conceive the relation between different sets of information we can collect from different perspective. Instead, the reduction of these relations into one single ontological sets of entities – be they neuronal, mental, behavioural, and so forth – will always be based on the appearance-reality fallacy. Now, the key question is no longer about how the domain x can explain and incorporate the domain y, but how the x and y domains can be linked together, in order not to conceive consciousness as a specific object of empirical/conceptual analysis, but as lifeworld, i.e. the condition in virtue of which

and with which we experience the world (this is also related to the actual consciousness exposed in the first chapter).

- **Interdisciplinary and non-reductive forms of collaboration between phenomenology, philosophy and neuroscience**

After the critical analysis of the neuro-centric and mento-centric positions, the last two chapters were dedicated to the constructive part, where I presented the project of a pluralist approach to study consciousness. Against any reductionism I proposed and discussed some interdisciplinary approaches to consciousness, like neurophenomenology, neuropragmatism, 4E Cognition and non-reductive neurophenomenology. Then, I inserted the topic of consciousness within a non-reductive context – ontological, epistemological and methodological – where the empirical results are interpreted in the light of a conceptual and phenomenological analysis, and where at the same time the phenomenological descriptions and philosophical conceptualization are informed by data obtained from neuroscience, as well as philosophical a priori concepts are challenged by empirical evidences. This implied that the kind of phenomenology used here could not confine itself to the realm of pure transcendental, as well as the neuroscientific data that will be illustrated was not to be interpreted according to the neuro-centric paradigm. This necessarily implied also a change in the way in which we consider consciousness, given that purely physicalist models of experience (“experience is in the brain”) or mentalists (“consciousness is a mental feature”) cannot be accepted anymore by this approach, which can be seen as a form of non-reductive neurophilosophy or neurophenomenology.

- **Philosophical-phenomenological part – Consciousness as pure intentionality, contentless disclosing-activity of the world**

I started the constructive section of my research by analysing the concept of consciousness from the way we experience the world, from the things of the world themselves. To do this, I took the fundamental characteristics drawn from the initial conceptual analysis (intentionality, phenomenality and self-awareness) and I related them to Sartre’s phenomenological studies. The goal was to provide a non-mentalistic description of consciousness that started from the way in which the world manifests

itself to us in its series of appearances. From this point of view, Sartre's first phenomenological works are fundamental, as they present an idea of experience as total openness to the world, devoid of internal contents and without idealistic subjectivisms. This pure intentionality towards the objects of the world is possible by means of pre-reflective self-awareness as the necessary condition through which every intentional content – with its phenomenal character – is given to us as something which is not us, and thus giving rise to the feature of subjectivity.

In order to be clearer from a conceptual point of view, I compared the Sartrean position with contemporary positions from philosophy of mind, however a perfect matching between theories was not possible. Indeed, Sartre's phenomenological ontology shares with intentionalism the idea that consciousness does not have mental contents rather phenomenal properties are properties of worldly objects, however intentionalism is still tied to a representationalist conception of consciousness (which therefore falls into the appearance-reality fallacy). Instead, like disjunctivism Sartre takes perception and imagination to be completely different modes of experience the world, and not simply more or less precise levels of adherence to the world. Lastly, in order to preserve the Sartrean No Content Thesis, i.e. consciousness has no mental content inside, I used the adverbialist interpretation of the Sartrean account of consciousness made by Rowlands, which argues that for explaining how experience of external objects is given to me as for-me (for-me-ness or mineness) we do not need to hypostatize additional intentional objects like qualia or representations, rather we can conceive the subjective modification in an adverbial sense I see/imagine/perceive things "minely".

- **Phenomenological part – Consciousness as embodied-being-in-the-world**

The last part of the work has been dedicated to a deeper investigation of how consciousness acts in the world. I claimed that intentionality is the condition that connects us with the world, in the way that we are body- and environment-related, and in order to do this, from a phenomenological perspective I proposed to fuse together the phenomenological analysis of bodily consciousness and the condition of being-in-the-world, while from a neuroscientific perspective I exposed and discussed some neuroscientific findings that provide a theory of how consciousness is structured

through an environment-body-brain system, instead of a mere neuro-centric framework.

On the one hand consciousness is not something inserted in or connected to the body, but is manifested through the body, whereas the body is the orientation and the point of view of the conscious being within the world, in such a way that everything I perceive or even imagine is not somewhere without any reference to me. On the other hand, the world is not in front of me like an external and detached set of things, I am within it with a body that gives me the coordinates towards the things of the world, so that the laptop is in front of me, while the door is behind me. This model of bodily consciousness derives from Sartre's *Being and Nothingness*, and I discussed his descriptions by comparing it with the 4E Cognition. Sartre's account of consciousness, using the 4E Cognition's vocabulary, intrinsically implies embodiment and embedment also it shows elements of enactment and of extendedness. In particular, the last points indicate how the fact that our experience is always embodied is not the centre of the Sartrean analysis, because embodiment, embedment, enactment, and extendedness exist as a function of the relation with the world. This idea led me to reflect not on bodily consciousness as such (how we perceive our body, etc.), rather on how the body is the expression of intentionality in the world, i.e. consciousness as embodied-being-in-the-world.

- **Neurophilosophical part I – The worldhood paradigm and the spatiotemporal nature of consciousness**

Neither the brainhood paradigm nor mento-centrism or a potential body-centric perspective are sufficient for the aims of this work. For this reason, I claimed instead that the “worldhood paradigm” is the right model for understanding consciousness and how we experience the world. In this regard, I presented clinical and daily situations in which the “worldhood” interpretation is more efficient than the neuro-centric one. Indeed, it is not possible to conceive a conscious being as a “worldless” subject taken from the outside, since consciousness is transcendence towards a world that is already given and into which we already find ourselves, therefore, the way in which the condition of being-in-the-world is structured intrinsically influences the way in which we experience. And this being-in-the-world is structured in a spatio-temporal way,

given that the nature of our experience is intrinsically temporal and spatial, in terms of what Sartre would call temporalization and spatialization. This is a key point of classical and contemporary phenomenology, which instead is often forgotten by philosophy of mind, 4E Cognition and cognitive sciences.

The spatio-temporal nature of experience allowed the final connection with neuroscience, given that I thought it necessary to offer an empirical explanation to the way in which the intentionality in our embodied-being-in-the-world develops in a spatial and temporal sense. Thus, I examined and discussed three of the main neuroscientific theories of consciousness, in an attempt to identify which was most suited to the worldhood approach. The Integrated Information Theory and the Global Neuronal Workspace Theory have proved to be inadequate for the purpose, because they are based on the neuro-centric model that does not consider the relationship with the body and the environment as necessary.

- **Neurophilosophical part II – Northoff's world-brain relation and the environment-body-brain alignment as predisposition for consciousness**

The Temporo-spatial Theory of Consciousness, proposed by Northoff, conceives the brain as a relational and experience-dependent organ that is part of a systematic and broader perspective over the human being and consciousness. Instead of the neuro-sensory or the neuro-cognitive models, he proposes a neuro-ecological model of the brain as necessarily related with body and environment. Thus, the environment-body-brain alignment is the empirical (therefore measurable) manifestation of a fundamental integration within the world and the brain in order to maintain the basic predisposition for the presence of consciousness, i.e. what Northoff called “world-brain relation”. Just as intentionality towards objects is not phenomenologically possible without the subject being first situated into a pre-given being-in-the-world, so the emergence of consciousness is not empirically given without an adequate environment-body-brain alignment.

- **Neurophilosophical part III – Locked-in syndrome and dreaming state are not counterexamples to experience as inherently world-related**

As I exposed in the introduction, the last part of the research has been dedicated to some experiments aimed at testing the strength of this account of consciousness as embodied-being-in-the-world and environment-body-brain alignment. Therefore, two cases have been examined, in which consciousness seems to remain intact even in the absence of relationship with the body and the environment. However, I demonstrated through both phenomenological and neuroscientific insights that these potential counterexamples miss the point. Indeed, in Locked-in syndrome, even if the body is efferently damaged it is still able to perceive and therefore to be connected with the environment; moreover in some cases where the syndrome gets worse (Complete LIS), it may be possible to see fluctuations and clouding of consciousness, precisely because the relationship with the world is altered. About dreams, the fact that during the dream the brain is completely disconnected from body and environment is far from being definitively demonstrated. Even if this were the case, dreaming states, instead of being an example of how consciousness can be produced without a body or environment, are useful to show that the experience without a world becomes incoherent, illogical, chaotic, spatiotemporally unstable; in a nutshell, a disturbed being-in-the-world.

I started this research based on the idea that experience is a crucial and peculiar phenomenon that is different from other complex issues, and I still believe that its multifaceted characteristics must be considered from many perspectives and approaches. In this sense, I think that the pluralist and interdisciplinary approach with philosophy, phenomenology and neuroscience here presented offers an alternative to reductionist views that are standard in both philosophy of mind and cognitive sciences. This approach heads towards the interpretation of experience as consciousness-in-the-world, strongly suggesting the need to reconsider consciousness not merely in terms of brain- or mind-relation, but intrinsically as world-related. And I am convinced that the results of this work, especially the concepts of embodied-being-in the-world and environment-body-brain alignment, are not only adequate for a purely theoretical purpose, rather they may also be developed towards an ethical and clinical context in future researches.

References

- Aaronson, S. (2014). Why I Am Not An Integrated Information Theorist (or, The Unconscious Expander), Retrieved July 8, 2019, from <https://www.scottaaronson.com/blog/?p=1799>.
- Aleksander, I. (2017). Machine Consciousness. In S. Schneider & M. Velmans (Eds.), *The Blackwell companion to consciousness* (II, pp. 93–105). Wiley-Blackwell.
- Allen, C., & Trestman, M. (2017). Animal Consciousness. In S. Schneider & M. Velmans (Eds.), *The Blackwell companion to consciousness* (II, pp. 63–76). Wiley-Blackwell.
- Alter, T. (2007). The knowledge argument. In M. Velmans & S. Schneider (Eds.), *The Blackwell Companion to Consciousness*. Blackwell.
- Andrillon, T., Poulsen, A. T., Hansen, L. K., Léger, D., & Kouider, S. (2016). Neural Markers of Responsiveness to the Environment in Human Sleep. *The Journal of Neuroscience*, 36(24), 6583 LP – 6596.
- Anscombe, G. E. M. (1981). The first person. In *The collected philosophical papers of G. E. M. Anscombe* (Vol. 2: Metaphy, pp. 21–36). Blackwell.
- Aristotle. (n.d.). *De Anima*. (C. Shields, Ed.) (2016th ed.). Oxford: Clarendon Press.
- Armstrong, D. (1997). What is consciousness? In N. Block, O. J. Flanagan, & G. Güzelde (Eds.), *The nature of consciousness: philosophical debates* (p. 843). MIT Press.
- Armstrong, D. M. (1968). *A Materialist Theory of the Mind*. Routledge.
- Arnaudo, E. (2017). Pain and dualism: Which dualism? *Journal of Evaluation in Clinical Practice*, 23, 1081–1086.
- Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neuroscience and Biobehavioral Reviews*.
- Astuti, R. (2001). Are we all natural dualists? A cognitive developmental approach. *Journal of the Royal Anthropological Institute*, 7(3), 429–447.
- Augustine, S. (2006). The confessions. In T. L. Cooksey (Ed.), *Masterpieces of Philosophical Literature*. Greenwood Press.
- Azzouni, J. (2013). *Knowledge and reference in empirical science. Knowledge and Reference in Empirical Science* (I). Routledge.
- Baars, B. J., Franklin, S., & Ramsay, T. Z. (2013). Global Workspace Dynamics: Cortical “Binding and Propagation” Enables Conscious Contents. *Frontiers in Psychology*, 4, 200.
- Babo-Rebelo, M., Richter, C. G., & Tallon-Baudry, C. (2016). Neural Responses to Heartbeats in the Default Network Encode the Self in Spontaneous Thoughts. *The Journal of*

Neuroscience.

- Baker, L. R. (2013). *Naturalism and the First-Person Perspective*. New York: Oxford University Press.
- Barandiaran, X. E. (2017). Autonomy and Enactivism: Towards a Theory of Sensorimotor Autonomous Agency. *Topoi*, 36(3), 409–430.
- Bargh, J. A., & Morsella, E. (2008). The Unconscious Mind. *Perspectives on Psychological Science*, 3(1), 73–79.
- Bassetti, C., Mathis, J., & Hess, C. W. (1994a). Multimodal electrophysiological studies including motor evoked potentials in patients with locked-in syndrome: report of six patients. *Journal of Neurology, Neurosurgery, and Psychiatry*, 57(11), 1403–1406.
- Bassetti, C., Mathis, J., & Hess, C. W. (1994b). Multimodal electrophysiological studies including motor evoked potentials in patients with locked-in syndrome: Report of six patients. *Journal of Neurology, Neurosurgery and Psychiatry*, 57(11), 1403–1406.
- Bates, D. (2003). The vegetative state: Guidance on diagnosis and management. *Clinical Medicine*.
- Bauby, J.-D. (1997). *The diving bell and the butterfly*. A.A. Knopf.
- Bauer, G., Gerstenbrand, F., & Rimpl, E. (1979). Varieties of the locked-in syndrome. *Journal of Neurology*.
- Bayne, T., Cleeremans, A., & Wilken, P. (Eds.). (2009). *The Oxford companion to consciousness. The Oxford companion to consciousness*. New York, NY, US: Oxford University Press.
- Bender, A., Jox, R. J., Grill, E., Straube, A., & Lulé, D. (2015). Persistent Vegetative State and Minimally Conscious State A systematic review and meta-analysis of diagnostic procedures. *Deutsches Ärzteblatt International*, 112, 235–242.
- Benini, A. (2017). *Neurobiologia del tempo*. Raffaello Cortina.
- Bennett, M. R., Dennett, D., Hacker, P., Searle, J., & Robinson, D. (2007). *Neuroscience and philosophy : brain, mind, and language*. New York: Columbia University Press.
- Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Oxford: Blackwell Pub.
- Berghofer, P. (2018). Why Husserl is a Moderate Foundationalist. *Husserl Studies*, 34(1), 1–23.
- Berkhoff, M., Donati, F., & Bassetti, C. (2000). Postanoxic alpha (theta) coma: A reappraisal of its prognostic significance. *Clinical Neurophysiology*.
- Bickle, J. (2003). *Philosophy and Neuroscience: A Ruthless Reductive Account. The Philosophical Review* (Vol. 57). Springer Netherlands.
- Binswanger, L. (1964). On the manic mode of being-in-the-world. In E. W. Straus (Ed.),

- Phenomenology: Pure and Applied*. Duquesne University Press.
- Bizzari, V. (2018a). A Phenomenological Approach to Psychopathologies: an Embodied Proposal Analyzing the Structures of Subjectivity. A Phenomenological Interview View project. *Special Issue of InterCultural Philosophy – Phenomenological Anthropology, Psychiatry, and Psychotherapy in Theory and Practice*.
- Bizzari, V. (2018b). Schizophrenia and Common Sense: A Phenomenological Perspective. In I. Hipólito, J. Gonçalves, & J. G. Pereira (Eds.), *Schizophrenia and Common Sense*. Springer.
- Blackmore, S., & Troscianko, E. T. (2018). *Consciousness - An Introduction*. Routledge Taylor & Francis.
- Block, N. (n.d.). *The Border between Seeing and Thinking*.
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18(2), 227–247.
- Block, N. (2002). The harder problem of consciousness. *Journal of Philosophy*, 99(8), 391–425.
- Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*.
- Bloom, P. (2004). *Descartes' Baby: How the Science of Child Development Explains What Makes Us Human*. New York: Basic Books.
- Blumenfeld, H. (2009). The neurological examination of consciousness. In *The Neurology of Consciousness*.
- Bockelman, P., Reinerman-Jones, L., & Gallagher, S. (2013). Methodological lessons in neurophenomenology: Review of a baseline study and recommendations for research approaches. *Frontiers in Human Neuroscience*.
- Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are the Neural Correlates of Consciousness in the Front or in the Back of the Cerebral Cortex? Clinical and Neuroimaging Evidence. *The Journal of Neuroscience*, 37(40), 9603–9613.
- Boulé, J. P., & O'Donohoe, B. (2011). *Jean-Paul Sartre: Mind and Body, Word and Deed*. Cambridge Scholars.
- Bourget, D., & Mendelovici, A. (2014). Tracking Representationalism. In A. Bailey (Ed.), *Philosophy of Mind: The Key Thinkers*. Continuum.
- Bouwsma, O. K. (1948). Naturalism. *Journal of Philosophy*, 45(1).
- Boyle, M. (2018). Sartre on Bodily Transparency. *Manuscripto*. scielo.
- Boyle, R. (1991). The origin of forms and qualities according to the corpuscular philosophy. In M. A. Stewart (Ed.), *Selected Philosophical Papers of Robert Boyle* (1666th ed.). Indianapolis/Cambridge: Hackett Publishing Company.
- Breitenfeld, T., Jurasic, M. J., & Breitenfeld, D. (2014). Hippocrates: The forefather of

- neurology. *Neurological Sciences*.
- Brentano, F. (1973). *Psychology From an Empirical Standpoint*. London: Routledge.
- Brown, R. (2014). *Consciousness inside and out: Phenomenology, neuroscience, and the nature of experience*. Springer.
- Bruno, M. A., Vanhaudenhuyse, A., Thibaut, A., Moonen, G., & Laureys, S. (2011). From unresponsive wakefulness to minimally conscious PLUS and functional locked-in syndromes: Recent advances in our understanding of disorders of consciousness. *Journal of Neurology*.
- Brüntrup, G., & Jaskolla, L. (Eds.). (2016). *Panpsychism : Contemporary Perspectives* (I). Oxford University Press.
- Bucci, P. (2006). Husserl, Galileo e “la crisi delle scienze europee.” *Galilaeana. Journal of Galilean Studies*, III.
- Buongiorno, F. (2014). Fenomenologia in “prima” e in “terza” persona : Searle e Dennett critici di Husserl. *Rivista Internazionale Di Filosofia e Psicologia*, 5(3), 267–288.
- Burge, T. (2010). *Origins of Objectivity*. Oxford University Press.
- Buyse, F. (2015). The Distinction between Primary Properties and Secondary Qualities in Galileo Galilei’s Natural Philosophy. *Working Papers of the Quebec Seminar in Early Modern Philosophy*, 1, 20–45.
- Byrne, A., & Logue, H. (2009). *Disjunctivism: Contemporary Readings*. MIT Press.
- Cairns, H., Oldfield, R. C., Pennybacker, J. B., & Whitteridge, D. (1941). Akinetic mutism with an epidermoid cyst of the 3rd ventricle. *Brain*.
- Campos, V. (2017). Did Someone Slip the Enactivist’s (Embodied) Mind? Convergences between Jean-Paul Sartre and Alva Noë.
- Canavero, S. (2015). The “Gemini” spinal cord fusion protocol: Reloaded. *Surgical Neurology International*, 6(18).
- Cardwell, M. S. (2013). Locked-in syndrome. *Texas Medicine*.
- Carel, H., & Meacham, D. (2013). *Phenomenology and Naturalism: Examining the Relationship Between Human Experience and Nature*. Cambridge University Press.
- Carr, D. (1970). Husserl’s Problematic Concept of the Life-World. *American Philosophical Quarterly*, 7(4).
- Carroll, W. M., & Mastaglia, F. L. (1979). “Locked-in Coma” in Postinfective Polyneuropathy. *Archives of Neurology*.
- Caruso, G., & Flanagan, O. (Eds.). (2018). *Neuroexistentialism: Meaning, morals, and purpose in the age of neuroscience. Neuroexistentialism: Meaning, Morals, and Purpose in the Age of Neuroscience*. Oxford University Press.
- Caston, V. (2008). Intentionality in Ancient Philosophy. In E. N. Zalta (Ed.), *The Stanford*

- Encyclopedia of Philosophy* (Fall 2008). Metaphysics Research Lab, Stanford University.
- Catalano, J. S. (2010). The Body and the Book: Reading Being and Nothingness. In K. J. Morris (Ed.), *Sartre on the Body* (pp. 25–40). London: Palgrave Macmillan UK.
- Chalmers, D. J. (1987). Consciousness: The first-person and third-person views.
- Chalmers, D. J. (1990). Consciousness and Cognition.
- Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3), 200–219.
- Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Conscious experience*. Book (I). Oxford Scholarship Online.
- Chalmers, D. J. (1998). On the search for the neural correlate of consciousness. In S. R. Hameroff, A. W. Kaszniak, & A. C. Scott (Eds.), *Toward a Science of Consciousness II*. MIT Press.
- Chalmers, D. J. (2000). What is a neural correlate of consciousness? In T. Metzinger (Ed.), *Neural Correlates of Consciousness*. MIT Press.
- Chalmers, D. J. (2002). Consciousness and Its Place in Nature. In *Philosophy of Mind. Classical and Contemporary Readings*. New York: Oxford University Press.
- Chalmers, D. J. (2003). Consciousness and its place in nature. In S. P. Stich & T. A. Warfield (Eds.), *Blackwell Guide to the Philosophy of Mind*. Blackwell.
- Chalmers, D. J. (2007). Phenomenal Concepts and the Explanatory Gap. In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*.
- Chalmers, D. J. (2011). *The Character of Consciousness*. *The Character of Consciousness*.
- Chalmers, D. J. (2018). The Meta-Problem of Consciousness. *Journal of Consciousness Studies*, 25(9–10), 6–61.
- Chapin, H., Bagarinao, E., & Mackey, S. (2012). Real-time fMRI applied to pain management. *Neuroscience Letters*, 520(2), 174–181.
- Charles, E. P., Golonka, S., & Wilson, A. D. (2014). The Most Important Thing Neuropragmatism Can Do : Providing an Alternative to “Cognitive” Neuroscience. In *Pragmatist Neuropsychology: American Philosophy and the Brain*.
- Chaudhary, U., Xia, B., Silvoni, S., Cohen, L. G., & Birbaumer, N. (2017). Brain–Computer Interface–Based Communication in the Completely Locked-In State. *PLoS Biology*.
- Chaudhary, U., Piccione, F., & Birbaumer, N. (2015). Brain-computer interface (BCI) communication in the locked-in. In C. Chatelle & S. Laureys (Eds.), *Assessing Pain and Communication in Disorders of Consciousness* (pp. 104–127). Routledge.
- Churchland, P. (2007). *Neurophilosophy at work*. *Neurophilosophy at Work*.
- Churchland, P. M. (1979). *Scientific realism and the plasticity of mind*. Cambridge: Cambridge

- University Press.
- Churchland, P. M. (1981). Eliminative Materialism and the Propositional Attitudes. *The Journal of Philosophy*, 78(2), 67–90.
- Churchland, P. M. (1989). *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science* (Vol. 101). MIT Press.
- Churchland, P. S. (1983). Consciousness: The Transmutation of a Concept. *Pacific Philosophical Quarterly*, 64(1), 80–95.
- Churchland, P. S. (1986). *Neurophilosophy. Towards a Unified Science of the Mind/Brain*. MIT Press.
- Churchland, P. S. (1990). Is Neuroscience Relevant to Philosophy? *Canadian Journal of Philosophy*, 20(sup1), 323–341.
- Churchland, P. S. (1996). The Hornswoggle Problem. *Journal of Consciousness Studies*, 3(5–6), 402–408.
- Churchland, P. S. (2002). *Brain Wise* (Vol. 66). MIT Press.
- Churchland, P. S. (2013). *Touching a nerve : the self as brain*. W W Norton & Co Inc.
- Churchland, P. S. (2014). *Touching a Nerve: Our Brains, Our Selves*. W. W. Norton & Co.
- Cincotta, M., Tozzi, F., Zaccara, G., Borgheresi, A., Lori, S., Cosottini, M., & Cantello, R. (1999). Motor imagery in a locked-in patient: Evidence from transcranial magnetic stimulation. *Italian Journal of Neurological Sciences*.
- Clark, A. (2001). Consciousness and the meta-hard problem. In *Mindware: An introduction to the philosophy of cognitive science* (Vol. 0). Oxford: Oxford University Press.
- Clark, A. (2008). Embodiment and Explanation. In *Handbook of Cognitive Science*.
- Clark, A. (2009). Spreading the joy? Why the machinery of consciousness is (Probably) still in the head. *Mind*.
- Clark, A. (2019). Perception As Controlled Hallucination. Predictive Processing and the Nature of Conscious Experience. A Conversation with Andy Clark. *Edge*, Retrieved September 3, 2019, from https://www.edge.org/conversation/andy_clark-perception-as-controlled-hallucination.
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19.
- Coenen, A. (2012). Modelling of auditory evoked potentials of human sleep–wake states. *International Journal of Psychophysiology*, 85(1), 37–40.
- Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the Bandwidth of Perceptual Experience? *Trends in Cognitive Sciences*, 20(5), 324–335.
- Cole, J. (2004). *Still lives : narratives of spinal cord injury*. MIT Press.
- Cole, J. D., & Paillard, J. (1995). Living without touch and peripheral information about body position and movement: Studies with deafferented subjects. In J. L. Bermudez, A. J.

- Marcel, & N. M. Eilan (Eds.), *The Body and the Self* (pp. 245–266). MIT Press.
- Colrain, I. M., & Campbell, K. B. (2007). The Use of Evoked Potentials in Sleep Research. *Sleep Medicine Reviews*, 11(4), 277–293.
- Conson, M., Pistoia, F., Sarà, M., Grossi, D., & Trojano, L. (2010). Recognition and mental manipulation of body parts dissociate in locked-in syndrome. *Brain and Cognition*.
- Conson, M., Sacco, S., Sarà, M., Pistoia, F., Grossi, D., & Trojano, L. (2008). Selective motor imagery defect in patients with locked-in syndrome. *Neuropsychologia*.
- Costa, V., & Cesana, L. (2019). *Fenomenologia della cura medica. Corpo, malattia, riabilitazione*. Brescia: Scolè.
- Cote, K. A., & Campbell, K. B. (1999). P300 to high intensity stimuli during REM sleep. *Clinical Neurophysiology*, 110(8), 1345–1350.
- Cox, G. (2006). *Sartre: A Guide for the Perplexed*. Bloomsbury Publishing.
- Crane, T. (2006). Is There a Perceptual Relation? In T. Gendler & J. Hawthorne (Eds.), *Perceptual Experience*. Oxford: Oxford University Press.
- Crane, T. (2009). Intentionalism. In A. Beckermann, B. P. McLaughlin, & S. Walter (Eds.), *The Oxford Handbook of Philosophy of Mind*. Oxford University Press.
- Crane, T. (2017). The Problem of Perception. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Spring 201). Metaphysics Research Lab, Stanford University.
- Crowell, S. (2012). Sartre's existentialism and the nature of consciousness. In S. Crowell (Ed.), *The Cambridge Companion to Existentialism* (pp. 199–226). Cambridge: Cambridge University Press.
- Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T. A., Fernández-Espejo, D., Pickard, J. D., ... Owen, A. M. (2011). Bedside detection of awareness in the vegetative state: a cohort study. *The Lancet*, 378(9809), 2088–2094.
- Cruse, D., Chennu, S., Fernández-Espejo, D., Payne, W. L., Young, G. B., & Owen, A. M. (2012). Detecting Awareness in the Vegetative State: Electroencephalographic Evidence for Attempted Movements to Command. *PLoS ONE*, 7(11).
- Curley, W. H., Forgacs, P. B., Voss, H. U., Conte, M. M., & Schiff, N. D. (2018). Characterization of EEG signals revealing covert cognition in the injured brain. *Brain*.
- Dainton, B. (2018). Temporal Consciousness. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Winter 201). Metaphysics Research Lab, Stanford University.
- De Beauvoir, S. (1965). *The prime of life*. Penguin Books.
- De Caro, M. (2014). Two Forms of Non-Reductive Naturalism. *Phenomenology and Mind*, 7, 54–62.
- De Caro, M. (2017). On galileo's platonism, again. In *Hypotheses and Perspectives in the History and Philosophy of Science: Homage to Alexandre Koyre 1892-1964* (pp. 85–

- 104).
- De Caro, M., & Macarthur, D. (Eds.). (2004). *Naturalism in Question*. Harvard University Press.
- De Caro, M., & Macarthur, D. (Eds.). (2010). *Naturalism and Normativity*. Cambridge University Press.
- De Jong, B. (2013). Complete motor locked-in and consequences for the concept of minimally conscious state. *Journal of Head Trauma Rehabilitation*.
- De Massari, D., Ruf, C. A., Furdea, A., Matuz, T., Van Der Heiden, L., Halder, S., ... Birbaumer, N. (2013). Brain communication in the locked-in state. *Brain*.
- De Monticelli, R., & De Vecchi, F. (2011). Phenomenologizing Cognitive Neuroscience? *Phenomenology and Mind*, 1, 10–16.
- Dehaene, S. (2014). *Consciousness and the brain : deciphering how the brain codes our thoughts*. Viking Press.
- Dehaene, S., Changeux, J. P., & Naccache, L. (2011). The global neuronal workspace model of conscious access: From neuronal architectures to clinical applications. *Research and Perspectives in Neurosciences*.
- Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*.
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79(1), 1–37.
- Dennet, D. (2003). Who's on First? heterophenomenology explained. *Journal of Consciousness Studies*, 21, 366.
- Dennett, D. C. (1988). Quining qualia. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in Modern Science*. Oxford University Press.
- Dennett, D. C. (1991). *Consciousness Explained* (Penguin Bo). Little, Brown and Co.
- Dennett, D. C. (2005). *Sweet Dreams: Philosophical Obstacles To A Science Of Consciousness*. *Sweet Dreams: Philosophical Obstacles To A Science Of Consciousness*.
- Dennett, D. C. (2007). Heterophenomenology reconsidered. *Phenomenology and the Cognitive Sciences*, 6(1–2), 247–270.
- Dennett, D. C. (2018). Facing up to the hard question of consciousness. *Philosophical Transactions of the Royal Society B: Biological Sciences*.
- Dennett, D. C. (Daniel C. (2017). *From bacteria to Bach and back : the evolution of minds*. W Norton & Co Inc.
- Descartes, René. (1644). *Principles of Philosophy*. (V. Rodger Miller & R. P. Miller, Eds.) (1982nd ed.). Dordrecht: Kluwer Academic Publishers.

- Descartes, René. (1649). The Passions of the Soul. In M. Moriarty (Ed.), *The Passions of the Soul and Other Late Philosophical Writings* (2015th ed.). Oxford: Oxford University Press.
- Descartes, René. (1996). *Descartes: Meditations on First Philosophy: With Selections from the Objections and Replies*. (J. Cottingham, Ed.), *Cambridge Texts in the History of Philosophy*. Cambridge: Cambridge University Press.
- Descartes, René. (2003). *Discourse on Method and Meditations*. (E. S. Haldane & G. R. T. Ross, Eds.) (2003rd ed.). Dover Publications.
- Di Paolo, E., & Thompson, E. (2014). The enactive approach. In *The Routledge Handbook of Embodied Cognition*.
- Dietrich, E., & Hardcastle, V. G. (2005). *Sisyphus's Boulder* (Vol. 60). Amsterdam: John Benjamins Publishing Company.
- Dijksterhuis, A. (2010). Automaticity and the Unconscious. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology*. John Wiley & Sons, Inc.
- Dimachkie, M. M., & Barohn, R. J. (2013). Guillain-Barré syndrome and variants. *Neurologic Clinics*.
- Djebbara, Z., Fich, L. B., Petrini, L., & Gramann, K. (2019). Sensorimotor brain dynamics reflect architectural affordances. *Proceedings of the National Academy of Sciences*, 116(29), 14769 LP – 14778.
- Dorsch, F. (2010). Transparency and imagining seeing. *Philosophical Explorations*, 13(3), 173–200.
- Dretske, F. (1996). Phenomenal externalism, or if meanings ain't in the head, where are qualia? *Philosophical Issues*, 7.
- Dretske, F. (2003). Externalism and self-knowledge. In S. Nuccetelli (Ed.), *New Essays on Semantic Externalism and Self-Knowledge*. MIT Press.
- Drummond, J. J. (1990). Husserl and Foundationalism. In *Husserlian Intentionality and Non-Foundational Realism: Noema and Object* (pp. 235–252). Dordrecht: Springer Netherlands.
- Dudzinski, D. (2001). The diving bell meets the butterfly: Identity lost and re-membered. *Theoretical Medicine and Bioethics*.
- Dumas, A. (2003). *The Count of Monte Cristo (Penguin Classics eBook)* (Vol. 1). Barnes & Noble Books.
- Duncan, J. (2007). Sartre and Realism-All-the-Way-Down. *Sartre Studies International*, 11(1–2).
- Eagleman, D. (2015). *The brain*. Pantheon Books.
- Edelman, D. B., & Seth, A. K. (2009). Animal consciousness: a synthetic approach. *Trends in*

- Neurosciences*, 32(9), 476–484.
- Edelman, G.M. (2006). *Second nature : brain science and human knowledge*. Yale University Press.
- Edelman, G M. (1989). *The Remembered Present: A Biological Theory of Consciousness*. Basic Books.
- Edelman, Gerald M. (1992). *Bright air, brilliant fire : on the matter of the mind*. BasicBooks.
- Edelman, Gerald M., Gally, J. A., & Baars, B. J. (2011). Biology of consciousness. *Frontiers in Psychology*.
- Edelman, Gerald M. (2003). Naturalizing consciousness: A theoretical framework. *Proceedings of the National Academy of Sciences*, 100(9), 5520 LP – 5524.
- Edelman, Gerald M, & Tononi, G. (2000). *A Universe of Consciousness: How Matter Becomes Imagination*. Basic Books.
- Engel, A. K., Maye, A., Kurthen, M., & König, P. (2013). Where's the action? The pragmatic turn in cognitive science. *Trends in Cognitive Sciences*.
- Eshleman, M. C. (2016). A sketch of Sartre's error theory of introspection. In S. Miguens, G. Preyer, & C. B. Morando (Eds.), *Pre-reflective Consciousness: Sartre and Contemporary Philosophy of Mind* (p. 520). New York: Routledge.
- Evans, G. (1982). *The Varieties of Reference*. New York: Oxford University Press.
- Evers, K. (2016). Neurotechnological assessment of consciousness disorders: five ethical imperatives. *Dialogues in Clinical Neuroscience*, 18(2), 155–162.
- Facco, E., Caputo, P., Fiore, D., & Giron, G. P. (1989). Sensorimotor and auditory central conduction time in locked-in syndrome. *Electroencephalography and Clinical Neurophysiology*.
- Faivre, N., Mudrik, L., Schwartz, N., & Koch, C. (2014). Multisensory Integration in Complete Unawareness: Evidence From Audiovisual Congruency Priming. *Psychological Science*, 25, 2006–2016.
- Farisco, M., Laureys, S., & Evers, K. (2017). The intrinsic activity of the brain and its relation to levels and disorders of consciousness. *Mind and Matter*.
- Farisco, M., & Petrini, C. (2014). Misdiagnosis as an ethical and scientific challenge. *Annali Dell'Istituto Superiore Di Sanita*, 50(3), 229–233.
- Faw, B. (2009). Cutting “Consciousness” at its Joints. *Journal of Consciousness Studies*, 16(5), 54–67.
- Feinberg, T. (2001). *Altered Egos: How the Brain Creates the Self*.
- Fenton, A., & Alpert, S. (2008). Extending Our View on Using BCIs for Locked-in Syndrome. *Neuroethics*.
- Fenton, M. B., Grinnell, A. D., Popper, A. N., & Fay, R. R. (Eds.). (2016). *Bat Bioacoustics*

- (Vol. 54). New York, NY: Springer New York.
- Finger, S. (2005a). An Ancient Egyptian Physician: The Dawn of Neurology. In *Minds Behind the Brain*. New York: Oxford University Press.
- Finger, S. (2005b). *Minds Behind the Brain: A history of the pioneers and their discoveries*. New York: Oxford University Press.
- Fink, S. B. (2016). A deeper look at the “neural correlate of consciousness.” *Frontiers in Psychology*, 7(JUL).
- Fischer, D. B., Boes, A. D., Demertzi, A., Evrard, H. C., Laureys, S., Edlow, B. L., ... Geerling, J. C. (2016). A human brain network derived from coma-causing brainstem lesions. *Neurology*, 87(23), 2427 LP – 2434.
- Flanagan, O. J. (1991). *The science of the mind* (2nd ed.). MIT Press.
- Flanagan, O. J. (1992). *Consciousness reconsidered*. MIT Press.
- Fodor, J. (2007). Headaches Have Themselves. *London Review of Books*, 29(10), 9–10.
- Fomina, T., Weichwald, S., Synofzik, M., Just, J., Schöls, L., Schölkopf, B., & Grosse-Wentrup, M. (2017). Absence of EEG correlates of self-referential processing depth in ALS. *PLOS ONE*, 12(6), e0180136.
- Formisano, R., D’Ippolito, M., & Catani, S. (2013). Functional locked-in syndrome as recovery phase of vegetative state. *Brain Injury*.
- Forstmann, M., & Burgmer, P. (2015). Adults are intuitive mind-body dualists. *Journal of Experimental Psychology: General*. Forstmann, Matthias: Department Psychologie, Universität zu Köln, Richard-Strauss- Str. 2, Köln, Germany, 50931, matthias.forstmann@uni-koeln.de: American Psychological Association.
- Frankish, K. (2016). Illusionism as a Theory of Consciousness. *Journal of Consciousness Studies*, 23(11–12), 11–39.
- Friedman, Y., Lee, L., Wherrett, J. R., Ashby, P., & Carpenter, S. (2003). Simulation of Brain Death from Fulminant De-efferentation. *Canadian Journal of Neurological Sciences*.
- Frith, C. D. (2007). *Making up the mind : how the brain creates our mental world*. Blackwell Pub.
- Frith, C., & Rees, G. (2007). A Brief History of the Scientific Approach to the Study of Consciousness. In *The Blackwell Companion to Consciousness* (pp. 7–22).
- Fuchs, T. (2013a). Depression, intercorporeality, and interaffectivity. *Journal of Consciousness Studies*, 20(7–8), 219–238.
- Fuchs, T. (2013b). Temporality and psychopathology. *Phenomenology and the Cognitive Sciences*.
- Fuchs, T. (2017). *Ecology of the brain : the phenomenology and biology of the embodied mind*.
- Fuchs, T. (2019). Editorial. More than Just Description: Phenomenology and Psychotherapy.

- Psychopathology*, 52, 1–4.
- Fuchs, T., & Röhrich, F. (2017). Schizophrenia and intersubjectivity. An embodied and enactive approach to psychopathology and psychotherapy. *Philosophy, Psychiatry and Psychology*, 24(2), 127–142.
- Fuchs, T., & Schlimme, J. E. (2009). Embodiment and psychopathology: A phenomenological perspective. *Current Opinion in Psychiatry*, 22(6), 570–575.
- Galilei, G. (1615). *Galileo a Madama Cristina Di Lorena* (1897th ed.). Padova: Tip. Salmin.
- Galilei, G. (1957). The Assayer (1623). In S. Drake (Ed.), *Discoveries and Opinions of Galileo*. New York: Anchor Books.
- Galilei, G. (2017). *Antologia di testi*. (M. Camerota, Ed.). Carocci.
- Gallagher, S. (1986). Body Image and Body Schema: A Conceptual Clarification. *The Journal of Mind and Behavior*, 7(4), 541–554.
- Gallagher, S. (1995). Body schema and intentionality. In J. Bermúdez, A. Marcel, & N. Eilan (Eds.), *The body and the self* (pp. 225–244). Cambridge, MA: MIT Press.
- Gallagher, S. (2003). Phenomenology and Experimental Design. *Journal of Consciousness Studies*, 10(9), 85–99.
- Gallagher, S. (2012). On the Possibility of Naturalizing Phenomenology. In D. Zahavi (Ed.), *The Oxford Handbook of Contemporary Phenomenology*. Oxford University Press.
- Gallagher, S. (2018). *Decentering the brain: Embodied cognition and the critique of neurocentrism and narrow-minded philosophy of mind*. *Constructivist Foundations* (Vol. 14).
- Gallagher, S., & Aguda, B. (2015). The Embodied Phenomenology of phenomenology. *Journal of Consciousness Studies*, 22, 93–107.
- Gallagher, S., & Cole, J. (1995). Body schema and body image in a deafferented subject. *Journal of Mind and Behavior*, 16(4), 369–390.
- Gallagher, S., & Zahavi, D. (2012). *The phenomenological mind*. Routledge.
- Gallagher, S., & Zahavi, D. (2016). Phenomenological approaches to self-Consciousness. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Winter 201). Metaphysics Research Lab, Stanford University.
- Gallese, V. (2011). Neuroscience and Phenomenology. *Phenomenology and Mind*, 1, 31–46.
- Gallese, V. (2017). *Mirroring, a liberated embodied simulation and aesthetic experience*. *Mirror Images. Reflections in Art and Medicine*.
- García-Espesón, C., López-Cancio, E., Martín-Aguilar, L., Millán, M., Castaño, C., Munuera, J., & Dávalos, A. (2016). Fluctuating locked-in syndrome as a presentation of a bilateral pontine infarction. *Neuroradiology Journal*.
- Gazzaniga, M. S. (2005). *The ethical brain*. Dana Press.

- Ghaemi, S. N. (2007). Feeling and time: The phenomenology of mood disorders, depressive realism, and existential psychotherapy. *Schizophrenia Bulletin*, 33(1), 122–130.
- Gibbard, A. (1975). Contingent identity. *Journal of Philosophical Logic*, 4(2), 187–221.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.
- Goebel, H. H., Komatsuzaki, A., Bender, M. B., & Cohen, B. (1971). Lesions of the Pontine Tegmentum and Conjugate Gaze Paralysis. *Archives of Neurology*, 24(5), 431–440.
- Goff, P. (2017). Panpsychism. *The Blackwell Companion to Consciousness*.
- Goff, P. (2019). *Galileo's Error. A new science of consciousness*. Rider.
- Goldman, A. I. (1970). *Theory of Human Action*. Princeton University Press.
- Goodman, L. E. (1992). *Avicenna*. Routledge.
- Gosseries, O., Bruno, M. A., Vanhaudenhuyse, A., Laureys, S., & Schnakers, C. (2009). Consciousness in the locked-in syndrome. In *The Neurology of Consciousness*.
- Gosseries, O., Pistoia, F., Charland-Verville, V., Carolei, A., Sacco, S., & Laureys, S. (2016). The Role of Neuroimaging Techniques in Establishing Diagnosis, Prognosis and Therapy in Disorders of Consciousness. *The Open Neuroimaging Journal*.
- Gosseries, O., Schnakers, C., & Laureys, S. (Eds.). (2019). *Between Theory and Clinic: The Contribution of Neuroimaging in the Field of Consciousness Disorders*. *Frontiers in Neurology* (Vol. 10). Frontiers Media S.A.
- Gosseries, O., Vanhaudenhuyse, A., Bruno, M.-A., Demertzi, A., Schnakers, C., Boly, M. M., ... Laureys, S. (2011). Disorders of Consciousness: Coma, Vegetative and Minimally Conscious States. In D. Cvetkovic & I. Cosic (Eds.), *States of Consciousness: Experimental Insights into Meditation, Waking, Sleep and Dreams* (pp. 29–55). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Green, C. D. (2003). Where did the ventricular localization of mental faculties come from? *Journal of the History of the Behavioral Sciences*, 39(2), 131–142.
- Grosz, E. (2015). Being-in-the-World and Schizophrenia. In D. Goodman & M. Freeman (Eds.), *Psychology and the Other* (pp. 247–269). Oxford University Press.
- Guillot, M. (2017). I Me Mine: on a Confusion Concerning the Subjective Character of Experience. *Review of Philosophy and Psychology*, 8, 23–53.
- Guilmette, T. J. (1997). *Pocket guide to brain injury, cognitive, and neurobehavioral rehabilitation*. Singular Pub. Group.
- Haan, J. (2013). Locked-in. The Syndrome as Depicted in Literature. *Progress in Brain Research*.
- Hales, C. G. (2014). *The Revolutions of Scientific Structure*. World Scientific.
- Hammeroff, S., & Penrose, R. (2014). Consciousness in the universe: A review of the ‘Orch OR’ theory. *Physics of Life Reviews*, 11(1), 39–78.

- Hanoğlu, L., Özkara, Ç., Yalçiner, B., Nani, A., & Cavanna, A. E. (2014). Epileptic qualia and self-awareness: A third dimension for consciousness. *Epilepsy & Behavior*, 30, 62–65.
- Harman, G. (1990). The Intrinsic Quality of Experience. *Philosophical Perspectives*, 4, 31–52.
- Hasan, A., & Fumerton, R. (2018). Foundationalist Theories of Epistemic Justification. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2018). Metaphysics Research Lab, Stanford University.
- Hassan, T., & Mumford, C. (1991). Guillain-Barre syndrome mistaken for brain stem death. *Postgraduate Medical Journal*.
- Haun, A. M., Tononi, G., Koch, C., & Tsuchiya, N. (2017). Are we underestimating the richness of visual experience? *Neuroscience of Consciousness*, 3(1).
- Hayashi, H., & Kato, S. (1989). Total manifestations of amyotrophic lateral sclerosis. ALS in the totally locked-in state. *Journal of the Neurological Sciences*.
- Hayashi, K., Mochizuki, Y., Takeuchi, R., Shimizu, T., Nagao, M., Watabe, K., ... Isozaki, E. (2016). Clinicopathological characteristics of patients with amyotrophic lateral sclerosis resulting in a totally locked-in state (communication Stage V). *Acta Neuropathologica Communications*.
- Heavey, C. L., Hurlburt, R. T., & Lefforge, N. L. (2010). Descriptive Experience Sampling: A Method for Exploring Momentary Inner Experience. *Qualitative Research in Psychology*, 7(4), 345–368.
- Heidegger, M. (1996). *Being and Time*. (J. Stambaugh, Ed.). Albany: State University of New York.
- Hobbes, T. (1811). Human Nature, or the Fundamental Elements of Policy. In W. Molesworth (Ed.), *The English Works of Thomas Hobbes of Malmesbury. Vol IV*. London: John Bohn.
- Hobson, J. A. (2009). REM sleep and dreaming: Towards a theory of protoconsciousness. *Nature Reviews Neuroscience*.
- Hobson, J. A., Pace-Schott, E. F., & Stickgold, R. (2000). Dreaming and the brain: Toward a cognitive neuroscience of conscious states. *Behavioral and Brain Sciences*, 23(6), 793–842.
- Høffding, S., & Martiny, K. (2016). Framing a phenomenological interview: what, why and how. *Phenomenology and the Cognitive Sciences*.
- Hoffman, D. (2010). Human Vision as a Reality Engine.
- Hoffman, D. D. (2019). *Case Against Reality : how evolution hid the truth from our eyes*. Allen Lane.
- Hoffman, D. D., & Ashbach, H. (2019). Interview: The case against reality.
- Hoffman, P. (2006). The Body. In H. L. Dreyfus & M. A. Wrathall (Eds.), *A Companion to*

- Phenomenology and Existentialism*. Wiley-Blackwell.
- Hogan, T., Hinrichs, U., & Hornecker, E. (2016). The Elicitation Interview Technique: Capturing People's Experiences of Data Representations. *IEEE Transactions on Visualization and Computer Graphics*, XX(XX).
- Hohmann, M. R., Fomina, T., Jayaram, V., Emde, T., Just, J., Synofzik, M., ... Grosse-Wentrup, M. (2018). Case series: Slowing alpha rhythm in late-stage ALS patients. *Clinical Neurophysiology*, 129(2), 406–408.
- Hohwy, J. (2007). The Search for Neural Correlates of Consciousness. *Philosophy Compass*, 2(3), 461–474.
- Hohwy, J. (2009). The neural correlates of consciousness: New experimental approaches needed? *Consciousness and Cognition*, 18(2), 428–438.
- Honderich, T. (2006). Radical externalism. *Journal of Consciousness Studies*, 13(7–8).
- Hood, B. M. (2012). *The self illusion : how the social brain creates identity*. Oxford University Press.
- Hopp, W. (2008). Husserl, phenomenology, and foundationalism. *Inquiry*, 51(2), 194–216.
- Horgan, T. (1989). Jackson on Physical Information and Qualia. *The Philosophical Quarterly*, 34(April), 147–183.
- Howell, R. J., & Thompson, B. (2017). Phenomenally Mine: In Search of the Subjective Character of Consciousness. *Review of Philosophy and Psychology*, 8, 103–127.
- Huang, Z., Vlisides, P. E., Tarnal, V. C., Janke, E. L., Keefe, K. M., Collins, M. M., ... Hudetz, A. G. (2018). Brain imaging reveals covert consciousness during behavioral unresponsiveness induced by propofol. *Scientific Reports*, 8(1).
- Hubbard, E. M., & Ramachandran, V. S. (2005). Neurocognitive Mechanisms of Synesthesia. *Neuron*, 48(3), 509–520.
- Hume, D. (1738). *A Treatise of Human Nature : Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects* (2000th ed.). London: Oxford University Press.
- Hume, D. (2007). *A Treatise of Human Nature. A Critical Edition. Volume 1: Texts*. (D. Fate Norton & M. J. Norton, Eds.). Oxford: Clarendon Press.
- Hurlburt, R., & Heavey, C. (2004). To beep or not to beep: Obtaining accurate reports about awareness. *Journal of Consciousness Studies*, 11(7–8), 113–128.
- Hurlburt, R. T., & Akhter, S. A. (2006). The Descriptive Experience Sampling method. *Phenomenology and the Cognitive Sciences*, 5(3), 271–301.
- Husserl, E. (1964). *Phenomenology of Internal Time Consciousness* (1st ed.). Indiana University Press.
- Husserl, E. (1970a). *Logical investigations*. (J. N. Findlay, Ed.). New York: Humanities Press.

- Husserl, E. (1970b). *The crisis of European sciences and transcendental phenomenology : an introduction to phenomenological philosophy*. Northwestern University Press.
- Husserl, E. (1977). *Cartesian meditations : an introduction to phenomenology*. M. Nijhoff.
- Husserl, E. (1982). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy*. (F. Kersten, Ed.). Springer Netherlands.
- Husserl, E. (1991). Rovesciamento della dottrina copernicana nell'interpretazione della corrente visione dei mondo. L'Arca originaria Terra non si muove. Ricerche fondamentali circa l'origine fenomenologica della corporeità, della spazialità, della natura nel senso primario . *Aut Aut*, 245.
- Husserl, E. (1999). *The idea of phenomenology: a translation of Die Idee der Phänomenologie, Husserliana II*. (L. Hardy, Ed.). Kluwer Academic Publishers.
- Husserl, E. (2006). *The Basic Problems of Phenomenology. From the Lectures, Winter Semester, 1910–1911*. Berlin/Heidelberg: Springer-Verlag.
- Hutto, D. D. (2017). REC: Revolution Effected by Clarification. *Topoi*, 36(3).
- Hutto, D. D., & Myin, E. (2012). *Radicalizing enactivism: Basic minds without content*. *Radicalizing Enactivism: Basic Minds without Content*. MIT Press.
- Hutto, D. D., & Myin, E. (2013). *Radicalizing Enactivism: Basic Minds Without Content*. MIT Press.
- Hutto, D. D., & Myin, E. (2017). *Evolving enactivism: Basic minds meet content*. *Evolving Enactivism: Basic Minds Meet Content*. MIT Press.
- Hutzler, F. (2014). Reverse inference is not a fallacy per se: Cognitive processes can be inferred from functional imaging data. *NeuroImage*.
- Irvine, E. (2013a). *Consciousness as a scientific concept: A philosophy of science perspective*. Springer.
- Irvine, E. (2013b). Measures of Consciousness. *Philosophy Compass*, 8(3), 285–297.
- Israel, G. (1996). *La mathématisation du réel. Essai sur la modélisation mathématique*. Paris: Seuil.
- Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, 32(April).
- Jackson, F. (1986). What Mary Didn't Know. *The Journal of Philosophy*, 83(5), 291.
- Jacob, P. (2019). Intentionality. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Spring 201). Metaphysics Research Lab, Stanford University.
- Jaffe, A. (2018). The illusion of time. *Nature*, 556, 304–305.
- James, W. (1984). *Psychology, Briefer Course*. Harvard University Press.
- Janz, B., Reinerman-Jones, L., Bockelman, P., Gallagher, S., & Trempler, J. (2016). *A Neurophenomenology of Awe and Wonder. A Neurophenomenology of Awe and Wonder*.
- Jaynes, J. (1976). *The Origin of Consciousness in the Breakdown of the Bicameral Mind*

- (2000th ed.). Boston - New York: Mariner Book.
- Jensen, M. S., Yao, R., Street, W. N., & Simons, D. J. (2011). Change blindness and inattention blindness. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(5), 529–546.
- Johnson, L. S. M., & Lazaridis, C. (2018). The Sources of Uncertainty in Disorders of Consciousness. *AJOB Neuroscience*, 9(2), 76–82.
- Joshi, N. J., Tononi, G., & Koch, C. (2013). The Minimal Complexity of Adapting Agents Increases with Fitness. *PLOS Computational Biology*, 9(7), e1003111.
- Joubert, C. (2014). Medicine and Mind-Body Dualism: A Reply to Mehta's Critique. *Mens Sana Monographs*, 12(1), 104–126.
- Jylkkä, J., & Railo, H. (2019). *Consciousness as a concrete physical phenomenon. Consciousness and Cognition* (Vol. 74).
- Kant, I. (1787). *Critique of pure reason*. (P. Guyer & A. W. Wood, Eds.) (1991st ed.). Cambridge University Press.
- Kaur, A., & Guan, Y. (2018). Phantom limb pain: A literature review. *Chinese Journal of Traumatology = Zhonghua Chuang Shang Za Zhi*, 21(6), 366–368.
- Kenny, A. (1971). The Homunculus Fallacy. In M. G. Grene & I. Prigogine (Eds.), *Interpretations of Life and Mind*. Humanities Press.
- Kihlstrom, J. F. (1987). The cognitive unconscious. *Science*, 237(4821), 1445–1452.
- Kind, A. (2018). *Philosophy of Mind in the Twentieth and Twenty-First Centuries: The History of the Philosophy of Mind, Volume 6*. Routledge.
- Kirk, R. (2019). Zombies. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Spring 201). Metaphysics Research Lab, Stanford University.
- Klausen, S. H. (2004). What is Really Wrong with Representationalism? *Res Cogitans*, 1(1).
- Knotts, J., Odegaard, B., Lau, H., & Rosenthal, D. (2019). Subjective inflation: phenomenology's get-rich-quick scheme. *Current Opinion in Psychology*, 29, 49–55.
- Koch, C. (2004). *The quest for consciousness : a neurobiological approach*. Roberts and Company Publishers.
- Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016a). Neural correlates of consciousness: progress and problems. *Nature Reviews. Neuroscience*.
- Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016b). Posterior and anterior cortex — where is the difference that makes the difference? *Nature Reviews Neuroscience*, 17(10), 666–666.
- Kondziella, D. (2017). Roald Dahl and the complete locked-in syndrome: “Cold dead body, living brain.” *Journal of the Neurological Sciences*.
- Kotchoubey, B., & Lotze, M. (2013). Instrumental methods in the diagnostics of locked-in

- syndrome. *Restorative Neurology and Neuroscience*.
- Kriegel, U. (2006). The same-order monitoring theory of consciousness. In U. Kriegel & K. Williford (Eds.), *Self-representational approaches to consciousness* (pp. 143–170). Cambridge: MIT Press.
- Kriegel, U. (2009). *Subjective Consciousness: A Self-Representational Theory*. Oxford: Oxford University Press.
- Kripke, S. (1980). *Naming and Necessity* (Vol. 217). Harvard University Press.
- Kurthen, M., Moskopp, D., Linke, D. B., & Reuter, B. M. (1991). The locked-in syndrome and the behaviorist epistemology of other minds. *Theoretical Medicine*.
- Kyselo, M., & Di Paolo, E. (2015). Locked-in syndrome: a challenge for embodied cognitive science. *Phenomenology and the Cognitive Sciences*.
- Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10(11), 494–501.
- Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience*, 1(3), 204–220.
- Landgrebe, L. (1981). Regions of Being And Regional Ontologies in Husserl's Phenomenology. In W. McKenna, R. M. Harlan, & L. E. Winters (Eds.), *Apriori and World: European Contributions to Husserlian Phenomenology* (pp. 132–151). Dordrecht: Springer Netherlands.
- Landi, A., Fornezza, U., De Luca, G., Marchi, M., & Colombo, F. (1994). Brain stem and motor evoked responses in “locked-in” syndrome. *Journal of Neurosurgical Sciences*.
- Langsam, H. (2018). Why Intentionalism Cannot Explain Phenomenal Character. *Erkenntnis*.
- Larmande, P., Hénin, D., Jan, M., Elie, A., & Gouazé, A. (1982). Abnormal vertical eye movements in the locked-in syndrome. *Annals of Neurology*, 11(1), 100–102.
- Latour, B. (2004). How to Talk About the Body? the Normative Dimension of Science Studies. *Body & Society*, 10(2–3), 205–229.
- Laudisa, F. (2014). *Naturalismo. Filosofia, scienza, mitologia*. Roma-Bari: Laterza.
- Laureys, S., Celesia, G. G., Cohadon, F., Lavrijsen, J., León-Carrión, J., Sannita, W. G., ... Dolce, G. (2010). Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. *BMC Medicine*, 8, 68.
- Laureys, S., & Schiff, N. D. (2012). Coma and consciousness: Paradigms (re)framed by neuroimaging. *NeuroImage*.
- Laureys, S., & Tononi, G. (2009a). The Neurology of Consciousness: An Overview. In S. Laureys & G. Tononi (Eds.), *The Neurology of Consciousness*. Elsevier.
- Laureys, S., & Tononi, G. (2009b). *The Neurology of Consciousness*. (S. Laureys & G. Tononi, Eds.), *The Neurology of Consciousness* (1st ed.). Academic Press.

- Lavazza, A. (2016). Free will and neuroscience: From explaining freedom away to new ways of operationalizing and measuring it. *Frontiers in Human Neuroscience*, 10, 262.
- Lavazza, A. (2019). Why cognitive sciences do not prove that free will is an epiphenomenon. *Frontiers in Psychology*, 10(FEB).
- Lavazza, A., & De Caro, M. (2010). Not so fast. on some bold neuroscientific claims concerning human agency. *Neuroethics*.
- Le Neindre, P., Bernard, E., Boissy, A., Boivin, X., Calandreau, L., Delon, N., ... Terlouw, C. (2017). Animal Consciousness. *EFSA Supporting Publications*, 14(4).
- Leder, D. (1990). The Absent Body. *The Absent Body*.
- Leder, D. (2018). Inside Insights: A Phenomenology of Interoception. In *Mind and Interoception (provisional title)*.
- Legrand, D. (2010). Myself with No Body? Body, Bodily-Consciousness and Self-consciousness. In D. Schmicking & S. Gallagher (Eds.), *Handbook of Phenomenology and Cognitive Science* (pp. 180–200). Dordrecht: Springer.
- Lenggenhager, B., Pazzaglia, M., Scivoletto, G., Molinari, M., & Aglioti, S. M. (2012). The Sense of the Body in Individuals with Spinal Cord Injury. *PLoS ONE*.
- León-Carrión, J., Eeckhout, P. van, & Domínguez-Morales, M. del R. (2002). The locked-in syndrome: a syndrome looking for a therapy. *Brain Injury*, 16(7), 555–569.
- León-Carrión, J., Van Eeckhout, P., Domínguez-Morales, M. D. R., & Pérez-Santamaría, F. J. (2002). Survey: The locked-in syndrome: A syndrome looking for a therapy. *Brain Injury*.
- Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*.
- Levine, J. (1999). Conceivability, Identity, and the Explanatory Gap. In S. R. Hameroff, A. W. Kaszniak, & D. J. Chalmers (Eds.), *Toward a science of consciousness III : the third Tucson discussions and debates* (pp. 3–12). MIT Press.
- Levine, J. (2001). *Purple Haze: The Puzzle of Consciousness*. Oxford: Oxford University Press.
- Levine, J. (2016). A Quasi-Sartrean Theory of Subjective Awareness. In S. Miguens, C. B. Morando, & G. Preyer (Eds.), *Pre-Reflective Consciousness: Sartre and Contemporary Philosophy of Mind*. Routledge.
- Levy, N. (2007). *Neuroethics: Challenges for the 21st Century*. Cambridge: Cambridge University Press.
- Libet, B. (2004). *Mind time: the temporal factor in consciousness*. *Mind time: the temporal factor in consciousness*. Cambridge, MA, US: Harvard University Press.
- Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention

- to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act. *Brain*, 106(3), 623–642.
- Locke, J. (1838). *An Essay concerning Human Understanding*. London: Tegg and Company.
- Locke, J. (1860). *An Essay concerning Human Understanding and a Treatise on the Conduct of the Understanding*. Philadelphia: Hayes & Zell.
- Loorits, K. (2018). The Location and Boundaries of Consciousness: a Structural Realist Approach. *Review of Philosophy and Psychology*, 9(3), 523–537.
- Loughlin, V. (2014). Sensorimotor Knowledge and the Radical Alternative. In J. Bishop & A. Martin (Eds.), *Contemporary Sensorimotor Theory* (pp. 105–116). Springer, Cham.
- Lucci, G., & Pazzaglia, M. (2015). Towards multiple interactions of inner and outer sensations in corporeal awareness. *Frontiers in Human Neuroscience*.
- Lukowicz, M., Matuszak, K., & Talar, A. (2010). A misdiagnosed patient: 16 Years of locked-in syndrome, the influence of rehabilitation. *Medical Science Monitor*.
- Lutz, A., & Thompson, E. (2003a). Neurophenomenology: Toward a neurophenomenology as an account of generative passages: a first empirical case study. *Phenomenology and the Cognitive Sciences*, 10(9–10), 31–52.
- Lutz, A., & Thompson, E. (2003b). Neurophenomenology - integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies*, 10(9–10).
- Lycan, W. G. (1998). In Defense of the Representational Theory of Qualia (Replies to Neander, Rey, and Tye). *Philosophical Perspectives*, 12, 479–487.
- Lycan, W. G. (2019). Representational Theories of Consciousness. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Fall 2019). Metaphysics Research Lab, Stanford University.
- Machamer, P. (1998). Galileo's machines, his mathematics, and his experiments. In P. Machamer (Ed.), *The Cambridge Companion to Galileo* (pp. 53–79). Cambridge: Cambridge University Press.
- Madhusoodanan, S., Ting, M. B., Farah, T., & Ugur, U. (2015). Psychiatric aspects of brain tumors: A review. *World Journal of Psychiatry*, 5(3), 273–285.
- Malafouris, L. (2013). *How things shape the mind : a theory of material engagement*. The MIT Press.
- Marino, S., Bonanno, L., Ciurleo, R., Baglieri, A., Morabito, R., Guerrera, S., ... De Stefano, N. (2017). Functional Evaluation of Awareness in Vegetative and Minimally Conscious State. *The Open Neuroimaging Journal*, 11, 17–25.
- Marion, J.-L. (2018). *On Descartes' Passive Thought, The Myth of Cartesian Dualism*. Chicago: University of Chicago Press.

- Martin, W., Gergel, T., & Owen, G. S. (2018). Manic temporality. *Philosophical Psychology*, 32(1), 72–97.
- Marzano, C., Ferrara, M., Moroni, F., & De Gennaro, L. (2011). Electroencephalographic sleep inertia of the awakening brain. *Neuroscience*, 176, 308–317.
- Mashour, G. A. (2018). The controversial correlates of consciousness. *Science*, 360(6388), 493–494.
- Maurel, M. (2009). The Explicitation Interview: Examples and Applications. *Journal of Consciousness Studies*, 16(10–12), 58–89.
- Mazzi, C., Bagattini, C., & Savazzi, S. (2016). Blind-sight vs. degraded-sight: Different measures tell a different story. *Frontiers in Psychology*, 7(901).
- McCullagh, P. J. (Peter J. (2005). *Conscious in a vegetative state? : a critique of the PVS concept* (1st ed.). Springer Netherlands.
- McCulloch, G. (1997). *Using Sartre: An Analytical Introduction to Early Sartrean Themes* (Vol. 47). Routledge.
- McGinn, C. (1991). *The Problem of Consciousness: Essays Toward a Resolution*. Blackwell.
- McKinsey, M. (2018). Skepticism and Content Externalism. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Summer 201). Metaphysics Research Lab, Stanford University.
- McLaughlin, B., & Bennett, K. (2018). Supervenience. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Winter 201). Metaphysics Research Lab, Stanford University.
- Medici, C., Gonzalez, G., Cerisola, A., & Scavone, C. (2011). Locked-in syndrome in three children with guillain-barré syndrome. *Pediatric Neurology*.
- Medline trend for ‘Consciousness.’ (n.d.).
- Menary, R. (2006). Attacking the bounds of cognition. *Philosophical Psychology*, 19(3), 329–344.
- Menary, R. (2010a). Introduction to the special issue on 4E cognition. *Phenomenology and the Cognitive Sciences*.
- Menary, R. (2010b). *The Extended Mind*. MIT Press.
- Merleau-Ponty, M. (1964). *The Primacy of Perception: And Other Essays on Phenomenological Psychology, the Philosophy of Art, History, and Politics*. (J. M. Edie, Ed.). Northwestern University Press.
- Merleau-Ponty, M. (1995). *La nature: cours du Collège de France*. Édition du Seuil.
- Merleau-Ponty, M. (2013). *Phenomenology of perception*.
- Mesa-Gresa, P., Gil-Gómez, H., Lozano-Quilis, J.-A., & Gil-Gómez, J.-A. (2018). Effectiveness of Virtual Reality for Children and Adolescents with Autism Spectrum

- Disorder: An Evidence-Based Systematic Review. *Sensors (Basel, Switzerland)*, 18(8), 2486.
- Metzinger, T. (2000). Neural correlates of consciousness: Empirical and conceptual questions. *MIT Press*.
- Metzinger, T. (2003). *Being no one : the self-model theory of subjectivity*. MIT Press.
- Metzinger, T. (2009). *The Ego Tunnel*. Basic Books.
- Michel, M., Fleming, S. M., Lau, H., Lee, A. L. F., Martinez-Conde, S., Passingham, R. E., ... Liu, K. (2018). An informal internet survey on the current state of consciousness science. *Frontiers in Psychology*.
- Mindt, G. (2017). The Problem with the “Information” in Integrated Information Theory. *Journal of Consciousness Studies*, 24(7–8).
- Monti, M. M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J. D., Tshibanda, L., ... Laureys, S. (2010). Willful modulation of brain activity in disorders of consciousness. *The New England Journal of Medicine*.
- Moore, G. E. (1903). *Principia Ethica* (1st ed.). Cambridge: Cambridge University Press.
- Morales, J., & Lau, H. (2019). The neural correlates of consciousness. In U. Kriegel (Ed.), *The Oxford Handbook of the Philosophy of Consciousness*. Oxford University Press.
- Moran, D. (2010a). Husserl, Sartre and Merleau-Ponty on Embodiment, Touch and the ‘Double Sensation’ BT - Sartre on the Body. In K. J. Morris (Ed.) (pp. 41–66). London: Palgrave Macmillan UK.
- Moran, D. (2010b). Sartre on embodiment, touch, and the “Double Sensation.” In *Philosophy Today*.
- Moran, D. (2011). Revisiting Sartre’s Ontology of Embodiment in Being and Nothingness. In P. V. (Ed.), *Ontological Landscapes: Recent Thought on Conceptual Interfaces Between Science and Philosophy*. Ontos.
- Moran, D. (2012). *Husserl’s crisis of the European sciences and transcendental phenomenology : an introduction*. Cambridge University Press.
- Moran, D. (2018a). Intentionality: Lived Experience, Bodily Comportment, and the Horizon of the World. In D. Zahavi (Ed.), *The Oxford Handbook of the History of Phenomenology*. Oxford University Press.
- Moran, D. (2018b). What is the Phenomenological Approach? Revisiting Intentional Explication. *Phenomenology and Mind*, (15), 72–90.
- Moran, D., & Cohen, J. D. (2012). *The Husserl dictionary*. Continuum.
- Morris, K. J. (2010). Introduction: Sartre on the Body. In K. J. Morris (Ed.), *Sartre on the Body* (pp. 1–22). London: Palgrave Macmillan UK.
- Morris, K. (2009). *Sartre on the Body*. Palgrave Macmillan UK.

- Morris, Katherine. (2010). Sartre on Consciousness: Remembering the Gestalt Psychology Context. In A. Mirvish & A. van den Hoven (Eds.), *New Perspectives on Sartre*. Cambridge Scholar Publishing.
- Moskalewicz, M., & Schwartz, M. A. (2018). Temporal experience in mania. *Phenomenology and the Cognitive Sciences*, pp. 1–14.
- Muller-Gass, A., & Campbell, K. (2014). Event-related potential measures of gap detection threshold during natural sleep. *Clinical Neurophysiology*, 125(8), 1647–1652.
- Murguialday, A. R., Hill, J., Bensch, M., Martens, S., Halder, S., Nijboer, F., ... Gharabaghi, A. (2011). Transition from the locked in to the completely locked-in state: A physiological analysis. *Clinical Neurophysiology*.
- Naccache, L. (2009). Visual Consciousness: An Updated Neurological Tour. In S. Laureys & G. Tononi (Eds.), *The Neurology of Consciousness* (1st ed., pp. 271–281). Academic Press.
- Naccache, L., Marti, S., Sitt, J. D., Trübtschek, D., & Berkovitch, L. (2016). Why the P3b is still a plausible correlate of conscious access? A commentary on Silverstein et al., 2015. *Cortex*, 85, 126–128.
- Nagel, E. (1961). *The structure of science: problems in the logic of scientific explanation*. Harcourt, Brace & World.
- Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review*, 83(4), 435.
- Nagel, T. (1986). *The view from nowhere* (I). Oxford University Press.
- Nagel, T. (1987). *What does it all mean? : a very short introduction to philosophy*. Oxford University Press.
- Nani, A. (2018). Modularità ed esperienza cosciente in una prospettiva neurocognitiva. *Rivista Internazionale Di Filosofia e Psicologia*, 9(3), 213–231.
- Nani, A., Seri, S., & Cavanna, A. E. (2013). Consciousness and neuroscience. In *Neuroimaging of Consciousness* (pp. 3–21).
- Nathan, M. J., & Del Pinal, G. (2017). The Future of Cognitive Neuroscience? Reverse Inference in Focus. *Philosophy Compass*, 12(7).
- Newen, A., De Bruin, L., & Gallagher, S. (2018). *The Oxford handbook of 4E cognition*. (L. de Bruin, A. Newen, & S. Gallagher, Eds.) (1st ed.). Oxford University Press.
- Newton, I. (1704). *Opticks: or a Treatise of the Reflections, Refractions, Inflections and Colours of Light* (4th, 1730th ed.). London: W. Innys.
- Nida-Rümelin, M. (2014). Basic intentionality, primitive awareness and awareness of oneself. In *Mind, Values, and Metaphysics: Philosophical Essays in Honor of Kevin Mulligan - Volume 2*.
- Nida-Rümelin, M. (2015). Qualia: The Knowledge Argument. In E. N. Zalta (Ed.), *The*

- Stanford Encyclopedia of Philosophy* (Summer 201). Metaphysics Research Lab, Stanford University.
- Nir, Y., & Tononi, G. (2010). Dreaming and the brain: from phenomenology to neurophysiology. *Trends in Cognitive Sciences*, 14(2), 88.
- Nizzi, M. C., Demertzi, A., Gosseries, O., Bruno, M. A., Jouen, F., & Laureys, S. (2012). From armchair to wheelchair: How patients with a locked-in syndrome integrate bodily changes in experienced identity. *Consciousness and Cognition*.
- Noë, A. (2004). *Action in Perception*. MIT Press.
- Noë, A. (2009). *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness. Minds and Machines* (Vol. 20).
- Noë, A., & Thompson, E. (2004). Are there Neural Correlates of Consciousness?'. *Journal of Consciousness Studies*, 11(1), 29–86.
- Noel, J.-P., Cascio, C. J., Wallace, M. T., & Park, S. (2017). The spatial self in schizophrenia and autism spectrum disorder. *Schizophrenia Research*, 179, 8–12.
- Northoff, G. (2004a). *Philosophy of the Brain: The Brain Problem*. John Benjamins.
- Northoff, G. (2004b). What is neurophilosophy? A methodological account. *Journal for General Philosophy of Science*.
- Northoff, G. (2011). *Neuropsychanalysis in Practice: Brain, Self, and Objects*. New York: Oxford University Press.
- Northoff, G. (2014a). Localization versus holism and intrinsic versus extrinsic views of the brain: A neurophilosophical approach. *Minerva Psichiatrica*, 55(1), 1–15.
- Northoff, G. (2014b). *Minding the brain : a guide to philosophy and neuroscience*. Red Globe Press.
- Northoff, G. (2014c). *Unlocking the Brain. Volume II: Consciousness*. New York: Oxford University Press.
- Northoff, G. (2016a). Does task-evoked activity entail consciousness in vegetative state? “Neuronal-phenomenal inference” versus “neuronal-phenomenal dissociation.” In M. Farisco & K. Evers (Eds.), *Neurotechnology and Direct Brain Communication*. New York: Routledge.
- Northoff, G. (2016b). *Neuro-philosophy and the healthy mind : learning from the unwell brain*. W W Norton & Co.
- Northoff, G. (2016c). Neuroscience and Whitehead I: Neuro-ecological Model of Brain. *Axiomathes*, 26(3), 219–252.
- Northoff, G. (2016d). Neuroscience and Whitehead II: Process-Based Ontology of Brain. *Axiomathes*, 26(3), 253–277.
- Northoff, G. (2016e). Spatiotemporal psychopathology I: No rest for the brain’s resting state

- activity in depression? Spatiotemporal psychopathology of depressive symptoms. *Journal of Affective Disorders*, 190, 854–866.
- Northoff, G. (2016f). Spatiotemporal Psychopathology II: How does a psychopathology of the brain's resting state look like? Spatiotemporal approach and the history of psychopathology. *Journal of Affective Disorders*, 190, 867–879.
- Northoff, G. (2018a). The brain's spontaneous activity and its psychopathological symptoms – “Spatiotemporal binding and integration.” *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 80, 81–90.
- Northoff, G. (2018b). *The Spontaneous Brain: From the Mind–Body to the World–Brain Problem*. Cambridge, MA: The MIT Press.
- Northoff, G. (2019). Lessons From Astronomy and Biology for the Mind—Copernican Revolution in Neuroscience. *Frontiers in Human Neuroscience*, 13.
- Northoff, G., & Heiss, W. D. (2015). Why Is the Distinction between Neural Predispositions, Prerequisites, and Correlates of the Level of Consciousness Clinically Relevant?: Functional Brain Imaging in Coma and Vegetative State. *Stroke*.
- Northoff, G., & Huang, Z. (2017). How do the brain's time and space mediate consciousness and its different dimensions? Temporo-spatial theory of consciousness (TTC). *Neuroscience & Biobehavioral Reviews*, 80, 630–645.
- Northoff, G., Magioncalda, P., Martino, M., Lee, H. C., Tseng, Y. C., & Lane, T. (2018). Too Fast or Too Slow? Time and Neuronal Variability in Bipolar Disorder - A Combined Theoretical and Empirical Investigation. *Schizophrenia Bulletin*.
- Northoff, G., & Stanghellini, G. (2016). How to Link Brain and Experience? Spatiotemporal Psychopathology of the Lived Body. *Frontiers in Human Neuroscience*.
- Northoff, G., & Tumati, S. (2019). “Average is good, extremes are bad” – Non-linear inverted U-shaped relationship between neural mechanisms and functionality of mental features. *Neuroscience and Biobehavioral Reviews*.
- Northoff, G., Wainio-Theberge, S., & Evers, K. (2019). Is temporo-spatial dynamics the “common currency” of brain and mind? In Quest of “Spatiotemporal Neuroscience.” *Physics of Life Reviews*.
- Northoff, G., Wiebking, C., Feinberg, T., & Panksepp, J. (2011). The “resting-state hypothesis” of major depressive disorder-A translational subcortical-cortical framework for a system disorder. *Neuroscience and Biobehavioral Reviews*.
- Nunziante, A. M. (2012). *Lo spirito naturalizzato. La stagione pre-analitica del naturalismo americano*. Trento: Verifiche.
- O'Hara, K., & Scutt, T. (1996). There is no hard problem of consciousness. *Journal of Consciousness Studies*, 3(4).

- O'Regan, J. K. (2011). *Why Red Doesn't Sound Like A Bell: Understanding the feel of consciousness. Why Red Doesn't Sound Like A Bell: Understanding the feel of Consciousness*. New York: Oxford University Press.
- O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5), 939–973.
- Odegaard, B., Chang, M. Y., Lau, H., & Cheung, S. H. (2018). Inflation versus filling-in: Why we feel we see more than we actually do in peripheral vision. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1755).
- Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. *PLOS Computational Biology*, 10(5), 1–25.
- Olivares, F. A., Vargas, E., Fuentes, C., Martínez-Pernía, D., & Canales-Johnson, A. (2015). Neurophenomenology revisited: second-person methods for the study of human consciousness. *Frontiers in Psychology*, 6, 673.
- Olivier, A. (2003). When pains are mental objects. *Philosophical Studies*, 115(1).
- Olsson, E. (2017). Coherentist Theories of Epistemic Justification. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 201). Metaphysics Research Lab, Stanford University.
- Ortín Nadal, A. (2019). Descartes on the distinction between primary and secondary qualities. *British Journal for the History of Philosophy*, 1–22.
- Ostrum, A. E. (1994). The “locked-in” syndrome - Comments from a survivor. *Brain Injury*, 8(1), 95–98.
- Overgaard, M. (2012). Blindsight: Recent and historical controversies on the blindness of blindsight. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3, 607–614.
- Overgaard, M. (Ed.). (2015). *Behavioral Methods in Consciousness Research*. Oxford University Press.
- Overgaard, M. (2017). The Status and Future of Consciousness Research. *Frontiers in Psychology*, 8(1719).
- Owen, A. M. (2013). *Detecting Consciousness: A Unique Role for Neuroimaging*. *Ssrn*.
- Owen, A. M. (2019). The Search for Consciousness. *Neuron*, 102, 526–528.
- Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., & Pickard, J. D. (2006). Detecting Awareness in the Vegetative State. *Science*, 313(5792), 1402 LP – 1402.
- Owen, M. (2019). Neural Correlates of Consciousness & the Nature of the Mind. In M. Guter (Ed.), *Consciousness and the Ontology of Properties*. New York: Routledge.
- Pagallo, U. (2018). Vital, Sophia, and Co.-The quest for the legal personhood of robots. *Information (Switzerland)*, 9(230).

- Papineau, D. (2016). Naturalism. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Winter 201). Metaphysics Research Lab, Stanford University.
- Park, H.-D., Bernasconi, F., Salomon, R., Tallon-Baudry, C., Spinelli, L., Seeck, M., ... Blanke, O. (2017). Neural Sources and Underlying Mechanisms of Neural Responses to Heartbeats, and their Role in Bodily Self-consciousness: An Intracranial EEG Study. *Cerebral Cortex*.
- Park, H. D., Correia, S., Ducorps, A., & Tallon-Baudry, C. (2014). Spontaneous fluctuations in neural responses to heartbeats predict visual detection. *Nature Neuroscience*.
- Park, H. D., & Tallon-Baudry, C. (2014). The neural subjective frame: From bodily signals to perceptual consciousness. *Philosophical Transactions of the Royal Society B: Biological Sciences*.
- Parnas, J., Handest, P., Sæbye, D., & Jansson, L. (2003). Anomalies of subjective experience in schizophrenia and psychotic bipolar illness. *Acta Psychiatrica Scandinavica*, 108(2), 126–133.
- Pasnau, R. (2007). Democritus and secondary qualities. *Archiv Fur Geschichte Der Philosophie*.
- Patterson, J. R., & Grabois, M. (1986). Locked-in syndrome: A review of 139 cases. *Stroke*.
- Pauen, M. (2012). The Second-Person Perspective. *Inquiry*, 55(1), 33–49.
- Pecere, P. (2018). La coscienza fenomenica tra neuroscienze e metafisica. L' "Ignorabimus" di du Bois-Reymond e il futuro della "Science of Consciousness." *Rivista di filosofia*, 2(agosto), 215–244.
- Peckitt, M. G. (2010). Resisting Sartrean Pain: Henry, Sartre and Biranism. In Katherine Morris (Ed.), *Sartre on the Body* (pp. 120–129). Palgrave-Macmillan.
- Pelczar, M. W. (2015). *Sensorama: A Phenomenalist Analysis of Spacetime and its Contents*. Oxford University Press.
- Pennartz, C. M. A. (2015, September 18). The Brain's Representational Power: On Consciousness and the Integration of Modalities. The MIT Press.
- Pennartz, C. M. A. (2018). Consciousness, Representation, Action: The Importance of Being Goal-Directed. *Trends in Cognitive Sciences*, 22(2), 137–153.
- Pepper, K. (2014). The Phenomenology of Sensorimotor Understanding. In J. M. Bishop & A. O. Martin (Eds.), *Contemporary Sensorimotor Theory*. Springer.
- Perconti, P. (2011). *Coscienza*. Bologna: Il Mulino.
- Pérez-Álvarez, M., García-Montes, J. M., Vallina-Fernández, O., & Perona-Garcelán, S. (2016). Rethinking Schizophrenia in the Context of the Person and Their Circumstances: Seven Reasons. *Frontiers in Psychology*, 7, 1650.
- Persuh, M. (2018). The Fata Morgana of Unconscious Perception. *Frontiers in Human*

- Neuroscience*, 12(120).
- Peters, M. A. K., Kentridge, R. W., Phillips, I., & Block, N. (2017). Does unconscious perception really exist? Continuing the ASSC20 debate. *Neuroscience of Consciousness*, 3(1).
- Peters, M. A. K., & Lau, H. (2015). Human observers have optimal introspective access to perceptual processes even for visually masked stimuli. *ELife*.
- Petitot, J., Varela, F., Pachoud, B., & Roy, J.-M. (Eds.). (1999). *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*. Stanford University Press.
- Phillips, I. (2016). Consciousness and Criterion: On Block's Case for Unconscious Seeing. *Philosophy and Phenomenological Research*, XCIII(2).
- Phillips, I., & Block, N. (2017). Debate on unconscious perception. In B. Nanay (Ed.), *Current Controversies in Philosophy of Perception*. Routledge.
- Place, U. T. (1956). Is Consciousness a Brain Process? *British Journal of Psychology*, 47(1), 44–50.
- Plato. (n.d.). *Timaeus*. (D. J. Zeyl, Ed.) (2000th ed.). Indianapolis/Cambridge: Hackett Publishing Company.
- Plum, F., & Posner, J. B. (1966). *The Diagnosis of Stupor and Coma* (I). Philadelphia: F.A. Davis Company.
- Polák, M., & Marvan, T. (2018). Neural correlates of consciousness meet the theory of identity. *Frontiers in Psychology*, 9(JUL).
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*.
- Poldrack, R. A. (2011). Inferring mental states from neuroimaging data: From reverse inference to large-scale decoding. *Neuron*.
- Prinz, J. (2015). Unconscious Perception. In M. Matthen (Ed.), *The Oxford Handbook of Philosophy of Perception*. Oxford University Press.
- Putnam, H. (1981). *Reason, Truth and History*. Cambridge: Cambridge University Press.
- Pyasik, M., Furlanetto, T., & Pia, L. (2019). The Role of Body-Related Afferent Signals in Human Sense of Agency. *Journal of Experimental Neuroscience*, 13, 1179069519849907.
- Quine, W. V. (1951). Main Trends in Recent Philosophy: Two Dogmas of Empiricism. *The Philosophical Review*, 60(1), 20–43.
- Ragazzoni, A., Grippo, A., Tozzi, F., & Zaccara, G. (2000). Event-related potentials in patients with total locked-in state due to fulminant Guillain-Barre syndrome. In *International Journal of Psychophysiology*.

- Ramsøy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences*, 3, 1–23.
- Rao, A. (2018). Our Experience of Reality Is a Bunch of Hallucinations We Collectively Agree On. *Motherboard*, Retrieved September 3, 2019, from https://www.vice.com/en_us/article/8xjbn3/consciousness-is-just-a-bunch-of-hallucinations-we-collectively-agree-on.
- Ratcliffe, M. (2006). Phenomenology, Neuroscience, and Intersubjectivity. In H. L. Dreyfus & M. A. Wrathall (Eds.), *A Companion to Phenomenology and Existentialism* (pp. 329–345). Blackwell.
- Ratcliffe, M. (2007). The Problem with the Problem of Consciousness. *Synthesis Philosophica*, 22(2), 483–494.
- Revonsuo, A. (2010). *Consciousness : the science of subjectivity*. Psychology Press.
- Revonsuo, A., & Salmivalli, C. (1995). A content analysis of bizarre elements in dreams. *Dreaming*, 5, 169–187.
- Reynolds, J., & Sebold, R. (Eds.). (2016). *Phenomenology and science : confrontations and convergences*. Palrave Macmillian US.
- Robins Wahlin, T.-B., & Byrne, G. J. (2011). Personality changes in Alzheimer’s disease: a systematic review. *International Journal of Geriatric Psychiatry*, 26(10), 1019–1029.
- Robinson, H. (2009). Why phenomenal content is not intentional. *European Journal of Analytic Philosophy*, 5(2), 79–93.
- Rohaut, B., Raimondo, F., Galanaud, D., Valente, M., Sitt, J. D., & Naccache, L. (2017). Probing consciousness in a sensory-disconnected paralyzed patient. *Brain Injury*.
- Rosanova, M., Gosseries, O., Casarotto, S., Boly, M., Casali, A. G., Bruno, M. A., ... Massimini, M. (2012). Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. *Brain*.
- Rosenthal, D. M. (1993). State Consciousness and Transitive Consciousness. *Consciousness and Cognition*.
- Roskies, A. (2002). Neuroethics for the new millenium. *Neuron*.
- Roskies, A. (2016). Neuroethics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 201). Metaphysics Research Lab, Stanford University.
- Rougé, A., Lemarié, J., Gibot, S., & Bollaert, P. E. (2016). Long-term impact after fulminant Guillain-Barré syndrome, case report and literature review. *International Medical Case Reports Journal*.
- Rovelli, C. (2019). *The order of time*. (E. Segre & S. Carnell, Eds.). Penguin.
- Rowlands, M. (2001). *The Nature of Consciousness*. Cambridge: Cambridge University Press.
- Rowlands, M. (2002). Two dogmas of consciousness. *Journal of Consciousness Studies*, 9(5–

6).

- Rowlands, M. (2003a). Consciousness: The transcendentalist manifesto. *Phenomenology and the Cognitive Sciences*, 2(3), 205–221.
- Rowlands, M. (2003b). *Externalism: Putting Mind and World Back Together Again*. Routledge.
- Rowlands, M. (2010a). Consciousness broadly construed. In R. Menary (Ed.), *The Extended Mind*. MIT Press.
- Rowlands, M. (2010b). *The New Science of the Mind: From Extended Mind to Embodied Phenomenology*. Bradford.
- Rowlands, M. (2013a). Enactivism, intentionality, and content. *American Philosophical Quarterly*, 50(3), 303–316.
- Rowlands, M. (2013b). Sartre, consciousness, and intentionality. *Phenomenology and the Cognitive Sciences*, 12(3), 521–536.
- Rowlands, M. (2015). Consciousness Unbound. *Journal of Consciousness Studies*, 22(3–4).
- Rowlands, M. (2016). Sartre on pre-reflective consciousness: The adverbial interpretation. In S. Miguens, G. Preyer, & C. B. Morando (Eds.), *Pre-Reflective Consciousness: Sartre and Contemporary Philosophy of Mind*. Taylor & Francis.
- Rowlands, M. (2018). Disclosing the World: Intentionality and 4E Cognition. In Albert Newen, L. De Bruin, & S. Gallagher (Eds.), *The Oxford Handbook of 4E Cognition*. Oxford University Press.
- Rupert, R. D. (2015). Embodiment, consciousness, and neurophenomenology: Embodied cognitive science puts the (First) person in its place. *Journal of Consciousness Studies*, 22(3–4), 148–180.
- Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson's University Library.
- Sacks, O. (1985). *The man who mistook his wife for a hat and other clinical tales* (1st ed.). New York: Summit Books.
- Salmon, N. (2005). Are General Terms Rigid? *Linguistics and Philosophy*, 28(1), 117–134.
- Sandberg, K., Frässle, S., & Pitts, M. (2016). Future directions for identifying the neural correlates of consciousness. *Nature Reviews Neuroscience*, 17, 666.
- Sarasso, S., Rosanova, M., Casali, A. G., Casarotto, S., Fecchio, M., Boly, M., ... Massimini, M. (2014). Quantifying Cortical EEG Responses to TMS in (Un)consciousness. *Clinical EEG and Neuroscience*, 45(1), 40–49.
- Sartre, J.-P. (1939). Une idée fondamentale de la phénoménologie de Husserl: l'intentionnalité. *La Nouvelle Revue Française*, (304), 129–131.
- Sartre, J.-P. (1948). Conscience de soi et connaissance de soi. *Bulletin de La Société Française de Philosophie*, 42, 49–91.

- Sartre, J.-P. (1956). *Being and nothingness*. (H. E. Barnes, Ed.), *Philosophical Library*, New York. New York: Philosophical Library.
- Sartre, J.-P. (1970). Intentionality: A fundamental idea of Husserl's phenomenology. *Journal of the British Society for Phenomenology*, 1(2).
- Sartre, J.-P. (1977). Sartre visto da Sartre. In F. Fagnani & P. A. Rovatti (Eds.), *Materialismo e Rivoluzione*. Il Saggiatore.
- Sartre, J.-P. (1981). An Interview with Jean-Paul Sartre: Interview by M. Rybalka, O. Pucciani, and S. Gruenheck (Paris, May 12 and 19, 1975). In P. A. Schilpp (Ed.), *The Philosophy of Jean-Paul Sartre* (pp. 3–51). Open Court.
- Sartre, J.-P. (2004a). *The imaginary: A phenomenological psychology of the imagination. The Imaginary: A Phenomenological Psychology of the Imagination*. London: Routledge Taylor & Francis.
- Sartre, J.-P. (2004b). *The transcendence of the ego: A sketch for a phenomenological description*. London and New York: Routledge.
- Sartre, J.-P. (2012). *The Imagination*. Routledge.
- Sartre, J.-P. (2015). *Sketch for a Theory of the Emotions* (Vol. 13). Routledge.
- Sartre, J.-P., & Contat, M. (1975). Sartre at Seventy: An Interview. *The New York Reviewer of Books*.
- Scheer, J. K. (2009). Experience and self-consciousness. *Philosophical Studies*, 144(1), 95–105.
- Schiff, N. D. (2015). Cognitive motor dissociation following severe brain injuries. *JAMA Neurology*.
- Schiff, N. D., & Fins, J. J. (2016). Brain death and disorders of consciousness. *Current Biology*.
- Schlicht, T. (2012). Phenomenal consciousness, attention and accessibility. *Phenomenology and the Cognitive Sciences*, 11(3), 309–334.
- Schlicht, T. (2018). A methodological dilemma for investigating consciousness empirically. *Consciousness and Cognition*, 66, 91–100.
- Schmicking, Daniel, & Gallagher, S. (Eds.). (2010). *Handbook of Phenomenology and Cognitive Science*. Dordrecht: Springer Netherlands.
- Schnakers, C., Perrin, F., Schabus, M., Hustinx, R., Majerus, S., Moonen, G., ... Laureys, S. (2009). Detecting consciousness in a total locked-in syndrome: An active event-related paradigm. *Neurocase*.
- Schnakers, C., Vanhaudenhuyse, A., Giacino, J., Ventura, M., Boly, M., Majerus, S., ... Laureys, S. (2009). Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. *BMC Neurology*, 9,

- Schnuerch, R., Kreitz, C., Gibbons, H., & Memmert, D. (2016). Not quite so blind: Semantic processing despite inattention blindness. *Journal of Experimental Psychology: Human Perception and Performance*, 42(4), 459–463.
- Schnürer, A., Espinosa, A., & Guger, C. (2015). mindBEAGLE: A BCI for communication and assessment of consciousness for patients with disorders of consciousness. *Clinical Neurophysiology*.
- Schörnich, S., Wallmeier, L., Gessele, N., Nagy, A., Schraner, M., Kish, D., & Wiegerebe, L. (2013). Psychophysics of Human Echolocation BT - Basic Aspects of Hearing. In B. C. J. Moore, R. D. Patterson, I. M. Winter, R. P. Carlyon, & H. E. Gockel (Eds.) (pp. 311–319). New York, NY: Springer New York.
- Schwarz, W. (2013). Contingent Identity. *Philosophy Compass*, 8(5), 486–495.
- Schwitzgebel, E. (2007). Do you have constant tactile experience of your feet in your shoes? Or is experience limited to what's in attention? *Journal of Consciousness Studies*, 14(3), 5–35.
- Seager, W. (2016). *Theories of consciousness: An introduction and assessment. Theories of Consciousness: An Introduction and Assessment*.
- Seager, W., & Bourget, D. (2007). Representationalism about Consciousness. In *The Blackwell Companion to Consciousness* (pp. 261–276). Blackwell.
- Searle, J. R. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Searle, J. R. (1992). *The Rediscovery of the Mind*. MIT Press.
- Searle, J. R. (1998). How to study consciousness scientifically. In S. R. Hameroff, A. W. Kaszniak, & A. C. Scott (Eds.), *Toward a Science of Consciousness II*. MIT Press.
- Searle, J. R. (1999). *Mind, language and society: philosophy in the real world*. London: Weidenfeld & Nicolson.
- Searle, J. R. (2019). *Il mistero della realtà (I)*. Raffaello Cortina.
- Seel, R. T., Sherer, M., Whyte, J., Katz, D. I., Giacino, J. T., Rosenbaum, A. M., ... Zasler, N. (2010). Assessment Scales for Disorders of Consciousness: Evidence-Based Recommendations for Clinical Practice and Research. *Archives of Physical Medicine and Rehabilitation*, 91(12), 1795–1813.
- Seth, A. (2009). Explanatory correlates of consciousness: Theoretical and computational challenges. *Cognitive Computation*, 1(1), 50–63.
- Seth, A. (2019). The Neuroscience of Reality - Reality is constructed by the brain, and no two brains are exactly alike. *Scientific American*, Retrieved September 3, 2019, from <https://www.scientificamerican.com/article/the-neuroscience-of-reality/>.

- Seth, A. K., Izhikevich, E., Reeke, G. N., & Edelman, G. M. (2006). Theories and measures of consciousness: An extended framework. *Proceedings of the National Academy of Sciences*, 103(28), 10799–10804.
- Seth, Anil K. (2016). The real problem. It looks like scientists and philosophers might have made consciousness far more mysterious than it needs to be.
- Seth, Anil K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends in Cognitive Sciences*, 12(8), 314–321.
- Shapiro, L. A. (2011). *Embodied cognition*. Routledge.
- Shoemaker, S. (1963). *Self-Knowledge and Self-Identity* (Vol. 15). Cornell University Press.
- Silverstein, B. H., Snodgrass, M., Shevrin, H., & Kushwaha, R. (2015). P3b, consciousness, and complex unconscious processing. *Cortex*.
- Silverstein, B. H., Snodgrass, M., Shevrin, H., & Kushwaha, R. (2016). Unconscious P3b and complex unconscious processing: Reply to Naccache et al., 2016. *Cortex*, 85, 129–132.
- Slowik, E. (2017). Descartes' Physics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2017). Metaphysics Research Lab, Stanford University.
- Smart, J. J. C. (1959). Sensations and Brain Processes. *The Philosophical Review*, 68(2), 141–156.
- Smit, H., & Hacker, P. M. S. (2014). Seven Misconceptions About the Mereological Fallacy: A Compilation for the Perplexed. *Erkenntnis*, 79(5), 1077–1097.
- Smith, E., & Delargy, M. (2005). Locked-in syndrome. *British Medical Journal*.
- Soddu, A., Boly, M., Nir, Y., Noirhomme, Q., Vanhaudenhuyse, A., Demertzi, A., ... Malach, R. (2009). Reaching across the abyss: recent advances in functional magnetic resonance imaging and their potential relevance to disorders of consciousness. *Progress in Brain Research*.
- Soekadar, S. R., Born, J., Birbaumer, N., Bensch, M., Halder, S., Murguialday, A. R., ... Martens, S. (2013). Fragmentation of slow wave sleep after onset of complete locked-in state. *Journal of Clinical Sleep Medicine*.
- Solymosi, T. (2011). Neuropragmatism, old and new. *Phenomenology and the Cognitive Sciences*.
- Solymosi, T. (2013). Neuropragmatism on the Origins of Conscious Minding.
- Solymosi, T., & Shook, J. (2013). Neuropragmatism: A Neurophilosophical Manifesto. *European Journal of Pragmatism and American Philosophy*, V(1).
- Solymosi, T., & Shook, J. (2014). *Neuroscience, Neurophilosophy and Pragmatism: Understanding Brains at Work in the World*. Palgrave-Macmillan.
- Soon, C. S., Brass, M., Heinze, H.-J., & Haynes, J.-D. (2008). Unconscious determinants of

- free decisions in the human brain. *Nature Neuroscience*, 11(5), 543–545.
- Soon, C. S., He, A. H., Bode, S., & Haynes, J.-D. (2013). Predicting free choices for abstract intentions. *Proceedings of the National Academy of Sciences*, 110(15), 6217 LP – 6222.
- Soteriou, M. (2016). The Disjunctive Theory of Perception. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Winter 201). Metaphysics Research Lab, Stanford University.
- Sprevak, M., & Irvine, E. (n.d.). Eliminativism about consciousness. In U. Kriegel (Ed.), *Oxford Handbook of the Philosophy of Consciousness*. Oxford University Press.
- Squire, L. R., Berg, D., Bloom, F. E., Du Lac, S., Ghosh, A., & Spitzer, N. C. (2012). *Fundamental Neuroscience: Fourth Edition*. *Fundamental Neuroscience: Fourth Edition*.
- Stanghellini, G., Ballerini, M., Presenza, S., Mancini, M., Northoff, G., & Cutting, J. (2017). Abnormal Time Experiences in Major Depression: An Empirical Qualitative Study. *Psychopathology*, 50(2), 125–140.
- Stoljar, D. (2006). *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness*. *Philosophy of Mind Series*. New York: Oxford University Press.
- Strawson, G. (2017). Consciousness never left. In K. Almquist & A. Haag (Eds.), *The return of consciousness: A new science on old questions* (pp. 89–103). Stockholm: Axel and Margaret Ax:son Johnson Foundation.
- Strawson, G. (2019). A hundred years of consciousness: “a long training in absurdity.” *Estudios de Filosofia*, 59, 9–43.
- Surlykke, A., Nachtigall, P. E., Fay, R. R., & Popper, A. N. (Eds.). (2014). *Biosonar* (Vol. 51). New York, NY: Springer New York.
- Svenaesus, F. (2006, January 1). Medicine. *A Companion to Phenomenology and Existentialism*.
- Taieb, H. (2018). *Relational Intentionality: Brentano and the Aristotelian Tradition* (1st ed.). Cham: Springer International Publishing.
- Tallon-Baudry, C., Campana, F., Park, H. D., & Babo-Rebelo, M. (2018). The neural monitoring of visceral inputs, rather than attention, accounts for first-person perspective in conscious vision. *Cortex*.
- Tavalaro, J., & Tayson, R. (1998). *Look up for yes*. Penguin Books.
- Thagard, P. (2010). *The brain and the meaning of life*. Princeton University Press.
- Thaler, L., Reich, G. M., Zhang, X., Wang, D., Smith, G. E., Tao, Z., ... Antoniou, M. (2017). Mouth-clicks used by blind expert human echolocators - signal description and model based signal synthesis. *PLoS Computational Biology*, 13(8), e1005670–e1005670.
- Theiner, G. (2017). The Extended Mind. In B. Turner (Ed.), *The Wiley-Blackwell*

- Encyclopedia of Social Theory*. Wiley-Blackwell.
- Thibault, R. T., MacPherson, A., Lifshitz, M., Roth, R. R., & Raz, A. (2018). Neurofeedback with fMRI: A critical systematic review. *NeuroImage*, 172, 786–807.
- Thompson, E. (1994). *Colour Vision: A Study in Cognitive Science and Philosophy of Science* (Vol. 47). Routledge.
- Thompson, E. (2007). Mind in Life. *Mind in Life*.
- Thompson, E. (2008). Representationalism and the phenomenology of mental imagery. *Synthese*, 160(3), 397–415.
- Thompson III, H. E. (1997). The fallacy of misplaced concreteness: Its importance for critical and creative inquiry. *Interchange*, 28(2–3), 219–230.
- Tononi, G. (2003). *Galileo e il fotodiodo: cervello, complessità e coscienza*. GLF editori Laterza.
- Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, 5(1), 42.
- Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *Biological Bulletin*.
- Tononi, G. (2009). Sleep and dreaming. In S. Laureys & Tononi (Eds.), *The Neurology of Consciousness* (pp. 89–107). Elsevier Ltd.
- Tononi, G. (2012). Integrated information theory of consciousness: an updated account. *Archives Italiennes de Biologie*, 150(4), 293–329.
- Tononi, G. (2017a). Integrated Information Theory of Consciousness. Some Ontological Considerations. In S. Schneider & M. Velmans (Eds.), *The Blackwell companion to consciousness* (Second, pp. 621–633). Oxford: John Wiley & Sons Ltd.
- Tononi, G. (2017b). The Integrated Information Theory of Consciousness: An Outline. In M. Velmans & S. Schneider (Eds.), *The Blackwell companion to consciousness* (II, pp. 243–256). Malden, MA: Blackwell Publisher.
- Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nature Reviews Neuroscience*, 17(7), 450–461.
- Tononi, G., & Koch, C. (2008). The neural correlates of consciousness: An update. *Annals of the New York Academy of Sciences*.
- Toombs, S. K. (1992). *The meaning of illness : a phenomenological account of the different perspectives of physician and patient*. Springer Netherlands.
- Tsuchiya, N. (2017). “What is it like to be a bat?”—a pathway to the answer from the integrated information theory. *Philosophy Compass*, 12(3), e12407.
- Tsuchiya, N., Wilke, M., Frässle, S., & Lamme, V. A. F. (2015). No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. *Trends in Cognitive Sciences*.

- Tye, M. (2002). Representationalism and the Transparency of Experience. *Noûs*, 36(1), 137–151.
- Tye, M. (2014). Transparency, qualia realism and representationalism. *Philosophical Studies*, 170(1), 39–57.
- Tye, M. (2018). Qualia. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 201). Metaphysics Research Lab, Stanford University.
- van Dijk, L., & Myin, E. (2019). Ecological Neuroscience: From Reduction to Proliferation of Our Resources. *Ecological Psychology*, 31(3), 254–268.
- Van Gulick, R. (2018). Consciousness. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring2018 ed.). Metaphysics Research Lab, Stanford University.
- Varela, F. (1996a). Neurophenomenology : A Methodological Remedy for the Hard Problem in. *Journal of Consciousness Studies*, (June).
- Varela, F. (1996b). Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies*, 3(4), 330–349.
- Varela, F. J. (1997). Patterns of life: Intertwining identity and cognition. *Brain and Cognition*.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind : cognitive science and human experience*. MIT Press.
- Velmans, M. (2009). How to Define Consciousness and how Not to Define Consciousness. *Journal of Consciousness Studies*, 16(5), 139–156.
- Ventriglio, A., & Bhugra, D. (2015). Descartes’ dogma and damage to Western psychiatry. *Epidemiology and Psychiatric Sciences*, 24, 368–370.
- Vidal, F., & Ortega, F. (2017). *Being brains: Making the cerebral subject. Being Brains: Making the Cerebral Subject*. Fordham University Press.
- Vidal, Fernando. (2009). Brainhood, anthropological figure of modernity. *History of the Human Sciences*, 22(1), 5–36.
- Vidal, Fernando. (2018). Phenomenology of the Locked-In Syndrome: an Overview and Some Suggestions. *Neuroethics*.
- Vimal, R. (2009). Meanings Attributed to the Term “Consciousness”: An Overview. *Journal of Consciousness Studies*.
- Voltoini, A. (2013). The Mark of the Mental. *Phenomenology and Mind*, 4, 124–136.
- Wade, D. T. (2018). How often is the diagnosis of the permanent vegetative state incorrect? A review of the evidence. *European Journal of Neurology*.
- Wagner, N.-F., & Northoff, G. (2017). Personal Identity and Brain Identity. In L. S. M. Johnson & K. Rommelfanger (Eds.), *The Routledge Handbook of Neuroethics*. Routledge.
- Walker, S. I., & Davies, P. C. W. (2017). The “hard problem” of life. In *From Matter to Life*:

Information and Causality.

- Wallace, B. A. (2000). *The taboo of subjectivity: toward a new science of consciousness*. Oxford University Press.
- Walshe III, T. M., & Walshe, T. M. (2016). On the Sacred Disease: Hippocrates. In *Neurological Concepts in Ancient Greek Medicine* (pp. 43–60). Oxford, UK: Oxford University Press.
- Ward, D., Silverman, D., & Villalobos, M. (2017). Introduction: The Varieties of Enactivism. *Topoi*, 36(3), 365–375.
- Ward, J. (2013). Synesthesia. *Annual Review of Psychology*, 64(1), 49–75.
- Watson, J. B. (1925). *Behaviorism*. London: Kegan Paul, Trench, trubner & Co.
- Webber, J. M. (2001). *Consciousness-in-the-world. An analytical Investigation of the Theory of Phenomenal Consciousness in the Early Works of Jean-Paul Sartre*. University College London.
- Weber, M., & Weekes, A. (2009). *Process approaches to consciousness in psychology, neuroscience, and philosophy of mind*. SUNY Press.
- Wehrle, M. (2019). Being a body and having a body. The twofold temporality of embodied intentionality. *Phenomenology and the Cognitive Sciences*.
- Wheeler, M. (2005). *Reconstructing the Cognitive World: The Next Step*. Cambridge, MA: MIT Press.
- Whitehead, A. N. (1920). *The Concept of Nature: The Tarner Lectures Delivered in Trinity College, November 1919*. Dover Publications.
- Whitehead, A. N. (1970). *Science and the modern world*. Free Press.
- Wider, K. (1993). Sartre and the Long Distance Truck Driver: The Reflexivity of Consciousness. *Journal of the British Society for Phenomenology*, 24(3), 232–249.
- Wider, K. (1997). *The Bodily Nature of Consciousness: Sartre and Contemporary Philosophy of Mind*. Cornell University Press.
- Wider, K. (2016). Sartre, enactivism, and the bodily nature of pre-reflective consciousness. In S. Miguens, G. Preyer, & C. B. Morando (Eds.), *Pre-Reflective Consciousness. Sartre and contemporary Philosophy of mind*. Routledge.
- Wilkes, K. V. (1988). *Real People: Personal Identity Without Thought Experiments* (Vol. 39). Oxford University Press.
- Williams, R. (2019). Eliminative Materialism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 201). Metaphysics Research Lab, Stanford University.
- Wilson, R. A. (2017). Embodied Cognition. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 201). Metaphysics Research Lab, Stanford University.
- Wilson, R. A., & Foglia, L. (2017). Embodied Cognition. In E. N. Zalta (Ed.), *The Stanford*

- Encyclopedia of Philosophy* (Spring 201). Metaphysics Research Lab, Stanford University.
- Windt, J. M. (2015). Dreaming: A conceptual framework for philosophy of mind and empirical research. *Dreaming: A Conceptual Framework for Philosophy of Mind and Empirical Research*.
- Winkler, R. (2017). *Phenomenology and Naturalism*. Taylor and Francis.
- Wu, W. (2018). The Neuroscience of Consciousness. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 201). Metaphysics Research Lab, Stanford University.
- Yousif, N., Cole, J., Rothwell, J., & Diedrichsen, J. (2015). Proprioception in motor learning: lessons from a deafferented subject. *Experimental Brain Research*.
- Zahavi, D. (2003). Intentionality and Phenomenality: Phenomenological Take on the Hard Problem. *Canadian Journal of Philosophy*, 33(sup1), 63–92.
- Zahavi, D. (2004). Phenomenology and the project of naturalization. *Phenomenology and the Cognitive Sciences*, 3(4), 331–347.
- Zahavi, D. (2007). Killing the straw man: Dennett and phenomenology. *Phenomenology and the Cognitive Sciences*.
- Zahavi, D. (2008a). Internalism, externalism, and transcendental idealism. *Synthese*, 160(3), 355–374.
- Zahavi, D. (2008b). *Subjectivity and Selfhood: Investigating the First-Person Perspective* (II). Cambridge: Bradford Book/MIT Press.
- Zahavi, D. (2010). Naturalized Phenomenology. In Daniel Schmicking & S. Gallagher (Eds.), *Handbook of Phenomenology and Cognitive Science* (pp. 2–19). Springer.
- Zahavi, D. (2014). *Self and Other: Exploring Subjectivity, Empathy, and Shame*. Oxford University Press. Oxford Scholarship Online.
- Zahavi, D. (2018). Getting It Quite Wrong: Van Manen and Smith on Phenomenology. *Qualitative Health Research*, 1049732318817547.
- Zahavi, D. (2019a). Consciousness and (minimal) selfhood: Getting clearer on for-me-ness and mineness. In U. Kriegel (Ed.), *The Oxford Handbook of the Philosophy of Consciousness*. Oxford University Press.
- Zahavi, D. (2019b). Varieties of Phenomenology (pp. 102–127).
- Zahavi, D., & Kriegel, U. (2015). For-me-ness: What it is and what it is not. In D. Dahlstrom, A. Elpidorou, & W. Hopp (Eds.), *Philosophy of Mind and Phenomenology* (pp. 36–53). Routledge.
- Zanin, F. (2011). *La scomparsa della Filosofia Naturale. Alle origini della Fisica Matematica*. Padova: CLEUP.

- Zelić, T. (2008). On the Phenomenology of the Life-World. *Synthesis Philosophica*, 23(2).
- Zhang, J., Huang, Z., Chen, Y., Zhang, J., Ghinda, D., Nikolova, Y., ... Northoff, G. (2018). Breakdown in the temporal and spatial organization of spontaneous brain activity during general anesthesia. *Human Brain Mapping*, 39(5), 2035–2046.
- Zilio, F. (2018a). Il cervello nel corpo-macchina. Meccanicismo e dualismo nel naturalismo contemporaneo. *Orbis Idearum*, 6(1), 117–138.
- Zilio, F. (2018b). Un approccio pragmatico al “problema reale” della coscienza. *Rivista Internazionale Di Filosofia e Psicologia*, 9(3), 232–238.
- Zilio, F. (2019). (Never)Minding the gap? Integrated Information Theory and philosophy of consciousness. In M. Curado & S. Gouveia (Eds.), *Automata's Inner Movie: Science and Philosophy of Mind* (1st ed., pp. 103–123). Vernon Press.
- Zola, E. (2001). *Thérèse Raquin*. Gallimard.

