Looking at the World with Fresh Eyes: The Perspectival Neutral Monism Model

By

Marta Santuccio

Submitted to

Central European University

Department of Philosophy

In partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Supervisor: Katalin Farkas

Budapest, Hungary

2022

COPYRIGHT NOTICE

I hereby declare that the dissertation contains no material accepted for the completion of any other degrees in any other institutions and no materials previously written and/ or published by another person unless appropriate acknowledgement is made in the form of biographical reference.

Marta Santuccio

Budapest, January 31st, 2022

ABSTRACT

In this thesis I aim to develop an alternative model for thinking about the question of the nature of consciousness and how it may fit with the physical world. I call this *Perspectival Neutral Monism*. Methodologically, I do this by casting a set of *fresh eyes* on the debate, hence questioning the notions and assumptions commonly adopted by popular views, in order to attempt to re-conceptualise and reframe the issue. Some of the ideas I question are the distinction of the phenomenal and the physical and whether these constitute ontological categories, whether we should adopt a physical conceptual scheme as a privileged description of reality, and whether we should assume an ontological hierarchy for investigating consciousness. The heart of this thesis is therefore to exercise our ability to carve a different avenue to approach the question of consciousness, rather than building a bulletproof position. For this reason, I generally refer to perspectival neutral monism as a model.

Perspectival neutral monism is a monist position, it rejects the split ontology of the dualist, and commits to reality that is neutral, thus that extends beyond the material and the mental stuff traditionally posited by material and mental monists. The view, moreover, is aligned with prior versions of neutral monism, such as those defended by Mach (1886), James (1912), Russell (2021), and more recently by Coleman (2017), Nagel (2012) and Heil (2013), in that it commits to a conceptual dualism alongside a uniform neutral ontology. The model I develop, however, departs from prior versions.

First, I expand the notion of neutrality beyond the orthodox *Neither* and *Both* views, adding that the neutral stuff is best seen as an infinite *multitude* that can be wholly captured from a perspective-free position. Second, I argue that the distinction between phenomenal and physical concepts is based on the availability of two distinct and limited perspectives, subjective and objective, and that these concepts pick out properties *as encountered* from the different perspectives. This further supports a commitment to a neutral ontology. Third, I develop a neutral notion of perspective, modelled on the enactive approach, whereby "occupying a perspective" is neutral because it involves the organism as whole (all those processes and states that we can capture with phenomenal and physical vocabularies). I then employ neutral

perspectives to explain how the phenomenal and the physical realms obtain in a neutral reality, following the enactivist idea that an organism narrows-down the environment into her world of significance on the basis of its internal specifications. I thus argue that a neutral subject narrows-down the neutral multitude into experience and the physical world, on the basis of her limitations.

Discussing how perspectival neutral monist may deal with challenges such as the Conceivability Argument or the threat of mental monism, I conclude that the model seems to have the potential to coherently weave together the phenomenal and the physical within a neutral ontology and thus that our discussion shows that thinking about the problem of consciousness with fresh eyes is possible.

ACKNOWLEDGEMENTS

The first person I want to thank is my supervisor Kati Farkas. Your ways of raising challenges, your rigour and especially your encouragement for me to think with my own head and express my own ideas has been an incredible motor for me throughout this process. I doubt my project would reflect me as much as it does, without your support.

A big thanks also goes to my former supervisor Philip Goff. The wonder I felt thinking about the big questions during our meetings and your way of pushing the boundaries of the orthodox have shown me that thinking outside the box is acceptable, and that I can keep myself excited doing philosophy.

I also want to thank Howard Robinson for being so amazing, Tim Crane, Hanoch Ben-Yami and Maria Kronefledner for their continued support during my time at CEU. I am also grateful to Martine Nida-Rumelin and Gergely Ambrus, my external supervisors during my study periods at Fribourg University in Switzerland and at ELTE in Hungary, for accepting and supporting my application and giving me feedback throughout my stay with you.

A special mention goes to 'The Blacks', Damian Aleksiev and Jamie Elliott. Thanks for being you, for all the laughter, through the tears, for all the conversations, for growing up together, for being real friends. You are always in my heart.

Garrett Mindt, the depth of everything that has come around is just impossible to capture in words. You have been a big part of this process, and of everything else that has happened these years. Thank you.

Lastly, a big thanks goes to my family. To my sister for showing me that philosophy can be made in movement. And to my father and to my mother, for believing blindly in my wildest dreams, thank you so much. None of this wold have been possible without you.

Table of Contents

INTRODUCTION

The purpose of this thesis is to look at the question of the nature of consciousness and its place in the physical world, and more importantly, to approach this question with a set of fresh eyes. I do this by developing a model that I call *Perspectival Neutral Monism*.

What is consciousness? Is that rich inner feeling we call consciousness yielded by inert senseless atoms? Or are atoms sensitive entities? Or, rather, could our feelings be a manifestation of a wider cosmic consciousness? Is the mind distinct from the body? Are we hidden behind a veil of appearance or is the world just as it appears to us?

When I raise these sort of questions to people outside the walls of University, they get incredibly excited as they explore their own take on the matter, at times led by logic, at others by intuition, belief, some notions they have studied, or faith. These questions excite them because they prompt them to tap into the mystery of their own life, their existence as human beings. Their eyes wide open as they wander through big ideas and dig deeper into questioning reality, from their everyday to the wondrous world of metaphysics, decked with contradictions, errors and buckets of bewildering wonder, often paired with the realisation that we just don't, or can't, know enough. However, when I pose these questions to my fellow philosophers the reaction is, at best, mixed. Some are bored, some dissatisfied with the debate, some others have lost hope that a coherent answer can be produced, some think the whole project of a metaphysics of mind is just misplaced. And while some of us remain excited, truth be told, the debate is not the most hospitable or encouraging, as can be appreciated by looking at the various strands of the mind-body problem that challenge every view in the debate from the more traditional dualism and physicalism to the more experimental contemporary panpsychism and neutral monism.

In the years I took reflecting on this issue, the excited and unstructured conversations I have had outside of Academia have been just as important as the rigorous ones I had with my fellow

philosophers for they reminded me that, alongside rigour and argumentative structures, philosophical thought requires us to *push forward* what we know and explore unknown territories, to keep asking questions, to re-evaluate those assumption we take for granted and embrace the mysterious nature of reality, thus the fact that there is so much we don't, and maybe cannot, know. Or at least, so it is for me.

Approaching the question of the nature of consciousness and the physical world I have conducted my research in line with this sentiment, thus attempting to cast a set of *fresh eyes* on the issue. The main feature of what I call my *fresh eyes* research strategy is to begin by questioning the basic notions and assumptions common in the debate. For example, some of the questions I raise concern how the notions of *phenomenal* and *physical* have developed through history, whether these should be thought of as ontological categories, whether consciousness should be understood with reference to a hierarchical model of reality and, furthermore, whether our scientific picture of the world should be translated into an ontological picture, as in the case of the physicalist. In this process of asking questions, I will eventually target some of those premisses that constitute central features of (more or less) established views such as dualism and physicalism or panpsychism. It is important to highlight, however, that my questioning attitude is not geared towards a desire to throw away the work done by other traditions, nor to undermine them. Rather, I aim to investigate the problem of consciousness starting from the work that those who came before me have done, reflecting on the strengths and weaknesses of their views and considering them in the light of the issues they face, whereby these issues are often rooted in their initial assumption as to the nature and structure of reality. I thus navigate my research by raising questions before adopting any assumption, other than the reality of experience and of a mind-independent world, so as to allow my thinking to develop within a broader logical space., therefore, letting the premisses and assumptions that underpin perspectival neutral monism naturally emerge from the reasoning and be acquired gradually.

One of the central targets of my questions in this thesis is scientific knowledge and its role and value in our intellectual tradition, especially within the philosophy of mind. This is because, no matter what position one may wish to adopt, the issue generally (and heavily) revolves around the idea that the physical sciences provide us with the most adequate account of reality, paired with our commonly held belief that it can produce a theory of *everything* in the

world, including an adequate explanation of consciousness. Science has certainly proved itself to be an incredibly successful explanatory project, from the scientific revolution to this day, however it seems unable to explain the rich, felt, inner aspect of our lives. In other words, the phenomenal and subjective aspect of our existence as human beings seems to fall between the cracks of objective explanation. The root of the problem, as I examine it in this thesis, is that such overwhelming success, based on the objective thinking that underpins it, has established the physical sciences as a privileged description of reality and has been held to constitute a strong intellectual paradigm. This has then been transposed into philosophy making, and particularly into metaphysics of mind. Farkas (2019) suggests that this was largely due not only to the wide success of the scientific method, but also to the common perception that philosophy does little, if any, progress. A such, Farkas (2009) points out that various methodologies for doing philosophy have either been imported from the sciences or have been developed to offer a degree of objectivity that is comparable to the explanatory power of the scientific method. Whereas this transposition may be useful for certain areas of philosophy, for instance in the philosophy of science, this may not be the case with respect to the project of adequately explaining the subjective and felt realm of consciousness.

Particularly, the issue I am mostly concerned with in this thesis is that the scientific structure of inquiry that has been imported into philosophy was then made into a series of philosophical assumptions that underpin the core postulates in the field of the metaphysics of consciousness, therefore shifting the focus from understanding the nature of consciousness to that of understanding how to make space for the subjective and the felt within an objective casual-functional physical world. Physicalism is the most obvious instance of this sort of approach, as well as the most wide-spread position nowadays. However, as we will see various times in the course of this text, any other view involved in the consciousness debate is centred around this issue, from the Cartesian dualist to the contemporary panpsychist. For this reason, this is an issue I spend much time raising questions about. Specifically, I examine the nature of the objective *view from nowhere* (Nagel's (1974) terminology), in particular the genesis and history of science and physics, and I attempt to understand whether we are right to hold the scientific objective explanatory project as a privileged description of reality, thereby tackling whether it can coherently be reframed without discarding it, and how. I hope to show that the scientific physical approach is just one of the frames of enquiry available to us and that, as such,

it should not be held as the single privileged explanation. Instead I will argue that both the objective physical and the subjective phenomenal are both essential elements of a broader map that can guide us towards a coherent ontology, whereby neither is privileged and both held equally valuable for discovering reality.

The ethos of my enquiry is driven by the idea that the subjective and the objective are both *limited* and *differently oriented* perspectives that we can occupy with respect to reality. This is a central notion in the perspectival neutral monism model I develop, and thus I argue at length towards this (possibly controversial) idea. For now, I want to highlight that this important feature of the view is motivated by a deep reflection on our nature as humans, namely our nature as limited beings, and my desire to embed these limitations as virtues of my theory as opposed to seeing them as vices. In effect, I use the very idea that we are limited beings as an explanatory principle in my ontology making. That is, I put the agent at the centre of the ontology/ theory-making and use these limitations to explain why reality is the way it is (for her) and that there is a large part of it that is beyond our grasp. In CD Broad's (1925) words

> how far the discursive form of cognition by means of general concepts can ever be completely adequate to the concrete Reality which it seeks to describe. (…) I do see clearly that we have only to compare a tune, as heard or an emotion, as felt, with any conceptual description which we can give of them, to recognise how inadequate every conceptual description of Reality must be to Reality itself (…) We cannot be acquainted with Reality as a whole, as we can with a tune or an emotion, and therefore the difficulty is at a maximum in Speculative Philosophy. This limitation of the whole conceptual scheme is one which we must simply recognise once and for all and then ignore.

While I agree with most of Broad's passage, I intend to disregard the suggestion presented at the end of it, namely ignoring these limitations. Rather, I take these features to be incredibly helpful tools, guides even, for us to discover and explain reality and they thus become crucial for justifying why positing a neutral reality is adequate.

Discourse on our limited ability to fully grasp reality is already present in the metaphysics of consciousness. Stoljar (2020) notes that this sort of approach is philosophically and scientifically plausible, as featured in the positions of notable philosophers such as Russell (1927), Nagel (1974) and McGinn (1989). McGinn (1989), for example, argues that we are *cognitively closed* with respects to finding out about consciousness, that our inability to grasp it is a matter of biological constitution, hence it shall forever remain a mystery for us. On the other hand, Stoljar (2006), defends the *Ignorance Thesis* according to which we are ignorant about a set of fundamental (non-standard material) properties that are, in effect, consciousness-relevant and that, therefore, we cannot explain consciousness on the basis of our ignorance of these properties. From this, he defends a position he calls non-standard materialism (2019) that is an ontology that if furnished by a set of unknown non-standard (consciousness-relevant) material properties alongside those that science (currently) reveals. I will come back to a discussion of Stoljar's view at various points in this thesis. For now, I just want to note the way he embeds ignorance about certain facts into his ontology by positing a set of non-standard material properties, where ignorance carries explanatory power and can thus be seen as virtuous. This strategy is relevantly similar to mine in the way I attempt to embed human limitations in my theory making, and thus make them into virtues of my metaphysics.

I feel that the admission of our limitations into ontology is another example of taking a fresh eyes approach to the question of the nature of consciousness, as well as to philosophy in general. As Stoljar (2020a) himself notes, this sort of approach is 'scientifically and philosophically plausible and it presents a way of doing philosophy that is quite different from the ways we have settled into the last 100 years' (Stoljar 2020, p. 496). I, hence, also agree with Griffin (2007, p.250) that 'accordingly, with Chalmers (1995, p. 204) and others (McGinn, 1991, pp. 2, 104; Nagel, 1986, pp. 8, 10; Searle, 1992, pp. 26, 49) that a constructive solution, if possible at all, will require a radically new approach'. Maybe this kind of approach can be radical enough, without falling into weirdness or incredulous stares.

The last guiding principle of my thesis is a desire to obtain a uniform picture of reality. This uniformity concerns two main aspects. The first relates to a uniformity between the mental and the material, that sort of uniformity that any monist seeks so as to defy the fragmentation of substances that underpins Cartesian dualism and that in turn raises the question of the

interaction between the mental and the material. In effect, monism is the only metaphysical assumption I adopt at the outset, one which for a lack of space I could not explore in depth here: I take the insolubility of the mind-body problem for the dualist as a sufficient reason to prefer a monist ontology as opposed to a dualist one. The second type of uniformity I seek concerns the desire to be able to trace a certain *continuity of being* across all living creatures. Although this is not a central theme in my thesis, my discussion touches upon it. The central reason for this concerns the relation of the special sciences, and biology in particular, to physics and their resistance to be adequately reduced to physics. I hope my model can be applied to this issue by offering an alternative framework for investigating organic life itself, alongside consciousness, placing both phenomena on the same continuum (De Jesus 2016; Thompson 2007; Wheeler, 1997, Godfrey-Smith, 1994). This would thus be offering a model for thinking about both consciousness and life, whereby the existence of neither appears radical, brute, aboriginal (James, 1912) or too mysterious. McGinn (1991), for example, tells us that if we are able to attribute consciousness to neutrons then it would be easy to understand how consciousness arises. Of course we know, from contemporary panpsychism, that this kind of intuition does not seem to do the desired work when put into practice; however, McGinn's claim clearly articulates the need for an explanation of the continuity of being in nature, which traditional views such as physicalism fail to produce. Within the perspectival neutral monist framework this has the potential be achieved, or so I hypothesise.

With the principles that guide my thesis in hand, we can now turn to introducing *perspectival neutral monism*. As the name suggests, this fits into the monist camp and is thus placed alongside mental and material monisms. These posit that reality is all of one kind and that this kind is either mental- for the mental monist- or material- for the material monist. The mental monist then has to explain the existence of the physical world and the nature of the physical properties which characterise it, whereas the material monist has to provide a story about the nature of experience and thus how mental properties might arise from the material base. Very briefly, perspectival neutral monism presents a commitment to a neutral reality, that is an ontology that extends the conception of the stuff of the world beyond the mental and the material as employed by traditional monists (and dualists). This is in conjunction with the idea that "phenomenal" and "physical" ultimately pick out two distinct *perspectives* that an agent can occupy with respects to reality, and which in turn confer the *bona fide* neutral properties either a

phenomenal or a physical character. My model of neutral monism can thus be said to follow the steps of fellow neutral monists with respects to posting a neutral ontology and the desire to coherently re-conceptualise the mental and the material (Russell, 1927; James, 1912; Mach, 1886; but also more contemporary Coleman, 2017; and Silberstein, 2018). I, however, depart from traditional forms in how I explain the existence of the phenomenal and the physical realms, that is by introducing the notion of (neutral) perspectives, one subjective and inward-directed and the other objective and outward-directed. Specifically, I work around a realism about experience and the mind-independent world, yet hold that the terms "phenomenal" and "physical" do not pick out the existence of phenomenal and physical properties, but merely *bona fide* neutral properties *as encountered* from the distinct perspectives. The crucial point here is that I develop a bespoke understanding of perspectives where these are understood as neutral states, on the grounds that they involve the agent as a whole and which *can* thus adequately be captured in phenomenal as well as in physical terms. This contrast with the traditional understanding of perspectives as mental states. The notion of perspectives I develop is based on a re-adaptation of some principles central to the enactive approach to cognition such as sense-making, interiority and its co-emergence with the external world and the interaction between global states and local processes. The enactive approach to cognition also serves as a model explain how experience and the physical world obtain in neutral context.

I want to note from the outset that my aim in this thesis should not be understood as a desire to develop a bullet-proof view or a solid solution to the mind-body problem, that is the problem of understanding how the mental and the material coherently fit into the same ontological picture. Rather, my aim is to explore whether there may be an alternative avenue to approach the issue, rooted in re-conceptualising and re-framing the main features of the debate into a *model* that has the potential to more adequately bring the phenomenal and the physical together. It is thus the process of enquiry, the mode of thought, the fresh-eyes method that I value the most in this thesis as it captures my attempt to show that there may indeed be other ways to think about the matter at hand. Therefore, my hope is that my perspectival neutral monism, if viable, can contribute to the debate by carving an alternative path for thinking about different solutions and by showing that a re-conceptualisation is possible. This is why I will often refer to the perspectival neutral monism I develop as a model: it is an experiment, a step towards different and possibly new territories that need further testing.

My thesis is structured into two parts: chapters 1-3 present the current debate and chapters 4-7 focus on the development of perspectival neutral monism. I begin by defining the "phenomenal" and the "physical" by tracing the history of these concepts from Aristotle to contemporary literature. I do this in chapter 1. In chapter 2 I outline the mind-body problem. I focus particularly on the problem for the physicalist and their attempts to reduce the phenomenal to the physical. In chapter 3 I outline traditional versions neutral monism. I unpack the definition, focusing especially on how we should characterise neutrality and how the mental and the material have been traditionally constructed within a neutral context. I discuss both older as well as more contemporary examples and outline some of the issues these face. In chapter 4 I begin laying the foundations of perspectival neutral monism. I argue that we have two sets of special concepts, namely phenomenal and physical concepts, that each set captures something different and equally valuable about the world, namely the mind independent world and experience, and that each is obtained in opposite manners, hence by occupying wholly distinct perspectives with respects to reality. However, I argue that having these distinct sets of concepts does not entail that this distinction is reflected on the ontological level. I also introduce the idea that phenomenal and physical concepts pick out properties in the world as encountered from the different perspectives. In chapter 5 I argue towards a neutral reality. First, I focus on analysing the "materialist paradigm", how it has developed historically and its philosophical shortcomings. I then question whether this should be embraced as a privileged description of reality, to the conclusion that it should not. I thus draw a map of the elements we know from the two perspectives, in conjunction with a consideration of the issues that traditional monists face, and find that a commitment to a neutral reality naturally emerges from this. I also discuss the limited nature of perspectives and how embracing limitations can support a commitment to a neutral ontology, thereby updating and complementing the orthodox notion of neutrality as a result. In chapter 6 I question the efficacy of adopting ontological hierarchies for a metaphysics of consciousness, a common strategy in the literature, suggesting that no ontological hierarchy should be adopted at the outset if one is to adequately investigate consciousness. In chapter 7 I focus on fleshing out the details of the perspectival neutral model. I unpack the central features of the enactive approach of the Thomson-Varela (1991) strand, which I then readapt to perspectival neutral monism. I successively model my notion of perspectives as properly neutral (rather than mental) on the enactivist principles, and then

investigate how these can help us produce a coherent explanation of experience and the physical world within a neutral ontology. I discuss that there are no physical or phenomenal properties in the perspectival neutral monist landscape, as well as how perspectival neutral monism relates to the issue of reduction. In the conclusion I highlight the features of the view that I feel may contribute to the broader debate, namely the updated conception of neutrality we develop in chapter 5 together with the understanding of perspectives as neutral states inspired by the enactivist approach in chapter 7. I also discuss how the perspectival neutral monist model has the ability to deal with issues such as the Explanatory Gap, the challenge of mental monism that all previous forms of neutral monists face, how it provides us with an ontology that can support the idea that we are really embodied beings, and how our model relates to biology and the life-mind continuum. In this last section we also show how our perspectival neutral monism makes consciousness not mysterious, but continuous with other phenomena such as life.

CHAPTER 1:

THE PHENOMENAL AND THE PHYSICAL

The project of explaining consciousness is parallel to that of understanding the relationship between the *phenomenal* and the *physical*. This revolves around the sentiment that, at least from the outset, the two seem distinct: the phenomenal refers to the sphere of experience that characterises our rich inner subjective life, while the physical refers to the barren objective realm of cause and effect that physics tells us about. The distinction between the phenomenal and the physical is generally taken for granted and used as the foundation for the development of various ontologies which populate the consciousness debate. Dualists, materialists, idealists, and even more controversial and "upstream" movements such as Russellian monism and panpsychism, all treat the distinction between the mental and the material as some sort of metaphysical starting point, albeit each in its own peculiar way.

Throughout this thesis I will use the term “phenomenal" interchangeably with the term “mental”. This is because the former is, in effect, a specific subset of the latter and thus issues about the phenomenal are deeply connected to issues about the broader set of mental phenomena. Moreover, I feel that the model we develop in chapter 7 could potentially be extended to other mental phenomena such as intentional states, although a discussion of this is not possible within the context of this thesis. Similarly, I will use the terms “physical” and “material” interchangeably. In a sense, the notion of “material” (central to the earlier materialist position) can be seen as the predecessor of the notion of the “physical” (central to the contemporary physicalist position). The two labels are often used interchangeably, as Stoljar (2001b) points out (with reference to “physicalism” and “materialism”). This will allow for a smoother exposition.

The way the phenomenal and the physical are defined is not always consistent across the literature. Once the intuitive conception of what is mental and what is material begins to be substantiated with theory, we obtain a number of (slightly) different ways to cash them out. Moreover, the distinction between mind and body that we currently take as being intuitive is one that belongs to modern philosophy and it has not always been a feature of our philosophical conceptualisation of reality. This, I believe, is something worth thinking about especially in the light of the many, and seemingly unresolvable, variations of the mind-body problem we encounter in the debate (which we discuss this in chapter 2).

The purpose of this chapter is to trace the history of the notions of "phenomenal" and "physical", to pin down how we understand them in the contemporary debate and how we conceive of their distinction. I also aim to discover how the conceptions of mental and material landed into the metaphysical realm, hence how they became substances or properties (and ontological categories more generally), thus forming the grounds for the development of various ontologies such as dualism or mental and material monism. This is essential groundwork for our defence of a neutral ontology in chapter 5, where I raise the question of whether the phenomenal and the physical should be understood as ontological categories. My desire is thus not to discard the notions, but to understand them in the light of their history.

The chapter is divided into four sections. In the first section I look at the Aristotelian tradition of matter and soul, as the temporally closest historical predecessors to our contemporary conceptions of the phenomenal and the physical. In the second section I survey Descartes (1596–1650), father of the distinction between the mental and the material as we know it today, and how his view was developed as a response to the Aristotelian tradition as well as to the advance of physics he was witnessing at the time. In the third section, I turn my eyes to the contemporary debate; I focus on the causal closure of the physical principle and the adoption of the notion of consciousness as a qualitative felt experience. In conjunction with these I also consider the idea that objective and subjective points of view are defining flags of the material and the mental respectively. In the fourth section I discuss the idea that the distinction between mental and material is in effect a category mistake, which I take from Ryle's (1949) seminal G*host in the Machine*.

1.1 Background: Aristotle's Soul and Body

Aristotle (1994) does not discuss the "mental" and the "material", rather he is concerned with understanding the relationship that holds between body and soul. It should be noted at the start that Aristotle's notions of body and soul (in line, more generally, with the ancient tradition) are very different to the ones we have today. For Aristotle the soul is marked by its being whole and immortal, while the body is defined by the fact that it dies and decomposes. This being said, Aristotle's theory of the relationship between soul and body is a good starting point for our enquiry into how the distinction between mind and matter arises for two reasons. First, because Aristotle's view does *not* present the distinction as in present-day discourse. Second, because it is the Aristotelian view, adopted by the Christian tradition, which constitutes the direct predecessor of the Cartesian approach, which Descartes objected to and thus which functioned as an important guide towards the development of his dualism and, more importantly, also to our modern conception of the mental and the material.

Aristotle's theory on body and soul is rooted in his more basic conception of *matter* and *form* and the relation that holds between them. It is therefore useful to briefly get a grasp of the latter in order to fully understand the former. For Aristotle, matter is 'that which in itself is not a this' while form or essence 'is that precisely in virtue of which a thing is called a this'. In this sense, Aristotle continues, 'matter is potentiality, form actuality' (De Anima, 1994, 412a6-412a11). What does this mean? This means that for Aristotle matter is nothing other than unshaped stuff, or material, that pervades the universe and which all things in the world are ultimately made out of. Matter exists only as a potentiality. It exists only as potential, rather than as actual, because it has the ability to be shaped into a thing, any thing, but it has not been shaped yet. This can be thought of as analogous to a piece of clay which holds the potential to be shaped into a plate, a vase, or a statue for example, but it is still just a lump of clay. Only once matter is combined with form does it become actualised. To become actualised means for matter to be shaped into a specific object or body. The clay finally becomes a vase, it becomes a plate, it becomes a statue. Form, from this perspective, constitutes the essence of matter because it confers matter actuality, it allows matter to be transformed into some specific thing. Form, hence, converts the existence of that lump of matter from *that which is in itself not a this* into a thing which can be called *a this*. It is the form that the lump of matter is shaped into,

therefore, that gives the piece of now-actualised matter the purpose or principle of its existence, the actuality of what it is and what it does during its existence.

The analogous reasoning applies to body and soul. For Aristotle, the soul is the form of the body, the principle of its existence. The body, in other words, is that lump of matter in potential, the raw clay, which becomes a specific body once the soul-form is "impressed" on it, so to say. Therefore, while form constitutes the essence of matter, with reference to the soul-form, Aristotle argues that the soul is the principle of being, the principle of life such that 'what has soul in it differs from what has not in that the former displays life' (*De* Anima, 1994, 413a2). This means that a body that has no soul in it is a lifeless body, may it be dead or an inanimate object.

In addition to this, having a soul is not confined to humans for Aristotle, rather it is to be associated with everything that has life including plants and animals. Specifically, having life for Aristotle equates with the body in question having certain powers or abilities; the soul is thus conceived as a 'system of active abilities to perform the vital functions that organisms of its kind naturally perform' (Lorenz, 2009). All beings which are alive, therefore, are similar in that they are all endowed with a soul such that they have some powers and abilities. Yet, there is another sense in which they differ from each other and this is the specific soul-form, thus set of powers, that each holds and which confers upon them a plant-shape, or an animal-shape, or a human-shape.

There obtains, therefore, a tripartite classification of the soul for Aristotle: the nutritive, the sensitive and the rational. 'The nutritive power is the originative power the possession of which leads us to speak of things as *living* at all' (De Anima, 1994, 413b1). This means that *all* living things must posses the power of nutrition, that is plants, animals and humans alike. This is because 'what has been born must grow, reach maturity, and decay—all of which are impossible without nutrition' (De Anima, 1994, 434a27). The power of nutrition here can thus be said to constitute the defining feature of a biological organism in general. The powers of the plant, as Aristotle explains, are wholly subsumed by the nutritive soul, such that the nutritive soul is the very principle of a plant's life: it shapes its body in a plant-form and endows it with those specific nutritive powers. Animals, on the other hand, on top of the basic nutritive abilities are also

endowed with sensitive powers, and in effect it is these sensitive powers that constitute the principle of an animal's life. Sensitive powers are those powers that enable an animal to 'perceive, experience sensations and appetites, feel emotions and imagine' (James, 2000, p113). For Aristotle these are closely related to having organs, because having organs by a certain body just is the embodiment of the sensitive soul; as we can see in contrast, the plant lacks organs and thus cannot feel, because it does not have the sensitive soul but only the nutritive. From this perspective it is clear how the sensitive soul comes to define the existence of an animal, how the sensitive abilities constitute the principle of an animal's life. Finally, human beings are further endowed with a rational soul, on top of the more basic nutritive and sensitive powers. The rational soul confers humans the ability to know, to understand, to execute rational and theoretical reasoning (Farkas, 2008; p8) and thus to make true or false judgements (James, 2000; p.113). Having this intellectual, or rational, soul distinguishes humans from animals, just like having a sensitive soul sets animals apart from plants.

Before moving onto Descartes in the next section, it is important to point out that the objects that the intellectual soul engages with are the objects of the external world as captured by the senses. In other words, the subject-matter of intellectual activity is crucially supplied by the senses from the outside world and therefore it is deeply linked to sensory perception, albeit not restricted to it. For Aristotle, having nutritive and sensitive souls corresponds directly to having certain bodily arrangements or organs- think about the sensitive soul and animal's organs- whereas the rational soul often seems to be disconnected from any specific feature that we find in the body. However, Aristotle 'does not mark off those vital functions that are mental by relating them to the soul in some special way that differs from and goes beyond the way in which vital functions in general are so related' (Lorenz, 2009). This means that even though vital nutritive and sensitive functions are more obviously related to the formed body, those function that are intellectual- and thus purely mental- remain tightly related to the body and its form. As we will see below, this is to become a major point of fracture between the Aristotelian tradition and Cartesian philosophy, which relies on a stricter distinction between the two (the body as mechanical and the mental as felt).

1.2 The Birth of the Mental and the Material: Descartes

The origin of our contemporary understanding of the mental and the material, how we conceive of them and of their distinction, can be arguably pinpointed in the context of the 17th century scientific revolution, and especially with reference to Descartes and his dualism.

Cartesian dualism is, in a nutshell, the metaphysical view that mind and matter are two distinct and independent kinds of stuff. More specifically, Descartes (1996) presents us with an ontology that features two independent substances, which are distinct in virtue of each of being defined by a distinct essential attribute. In a Cartesian context, substances are to be understood as fundamental metaphysical entities that enjoy independent existence, this means substances are self-sustaining entities which depend on nothing other than themselves for their existence[1]. Each of Descartes' substances are set apart from the other in virtue of being characterised by a distinct essential property. In expounding Descartes' view, Swinburne (2018, p.137) explains that a property of a substance should be thought of as essential if and only if it is inconceivable that a substance could exist without having that property. In this sense, he continues, an essential property of a substance constitutes its 'nature and essence' (Swinburne, 2018, p.137). For example, say that the essential property of a knife is "having a cutting blade". A knife may be very large or pocket-sized, have a red handle or a black one, it may be more or less sharp, but all these properties are not essential to the knife being a knife, because it is conceivable that a knife could exist without it being either sharp or red or big, etc. Instead it would be inconceivable to think of a knife not "having a cutting blade." According to Descartes, the two distinct essential properties that define the substances in question are *extension* and *thought*, respectively yielding the extended substance (*res extensa)* which underpins the material, and the thinking substance (*res cogitans*) which underpins the mental.

Understanding what the *res extensa* and the *res cogitans* are enables us to trace the source of our modern day conception of the mental and the material and the distinction we

[1] Note that in Descartes' own original view, the substances in questions presupposed the existence of a third type of substance on which they both depended: God. For the purposes of understanding Descartes' inception of mental and material, however, a consideration as to the God substance is not necessary, as also supported by the fact that the contemporary cartesian legacy takes God out of the equation. I will therefore not discuss the role of this substance in my thesis.

commonly impose between them. I start by introducing the *res extensa*, and in order to do this I consider the context in which Descartes developed his thought.

Descartes lived and worked at the beginning of the 17th century, a period recorded in history as the beginning of what came to be known as the scientific revolution sparked by the works of Galileo (1564-1642) and culminated in the physics of Isaac Newton (1642-1726). One of the main innovations that this revolution yields is a conception of the natural world based on the principle that nature operates according to strict *mechanical* laws, which are to be determined empirically. An example of a mechanical understanding of nature can be given in terms of the interaction between a couple of billiard balls: a force is applied to the first billiard ball which rolls over the table and hits the second one, which in turn starts rolling itself, where the velocity of the motion of each ball is inherently determined by the force applied on them by the previous event. From this perspective, nature's operations resemble that of a machine. This is in stark contrast to the Aristotelian view that the movement of objects in nature is caused by the form or soul that potential matter is imparted with.

The (cause-effect) mechanical conception of nature arises in conjunction with the notion of matter as extended. The reasoning is that for bodies to act mechanically, as per the billiard balls, they must necessarily be extended, that is they need to be located in space, have shape and size. Ben-Yami (2015, p.52) explains that for Galileo to conceive of a material or corporeal substance, for instance, means to necessarily conceive of it as 'having shape and place, as being in motion or at rest, and as having other properties of this kind; namely, matter necessarily has *geometrical* properties.' The notion of matter as having geometrical properties is, in effect, very important for our understanding of matter as extended, because it implies that matter is shaped in and of itself; in this sense it no longer only exists only as potential and awaiting a form as per the Aristotelian tradition. Moreover, given that these geometrical properties account for motion as well as shape, they are sufficient for explaining how bodies move in nature, mechanically. This entails that extended matter exists and operates independently from any kind of soul, and as such it is always *per se*, inanimate.

Finally, the Galilean notion of matter also imposes mathematics and the principles of geometry as the proper language by which scientific reasoning is to be executed, writing that

> Philosophy is written in this grand book, the universe, which stands continually open to our gaze. (…) It is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures, without which it is humanly impossible to understand a single word; without these, one wanders about in a dark labyrinth. (Galilei 1957, pp. 237–238)

Matter thus becomes something that can be wholly described in mathematical terms and according to the principles of geometry. A crucial implication of this conception of matter is that qualities such tastes, odours and colours are transferred from being had by objects in the world to being wholly held in consciousness, in the mind (Goff, 2017, p.12). This marks the beginning of the distinction between body and mind that we are so familiar with, as well as a strong departure from the Aristotelian tradition according to which sensory qualities were picked up by the senses from the outside world[2].

It is in this context that Descartes develops his dualism and, undoubtedly, the Galilean notion of matter deeply resembles Descartes' own, providing us with a good starting point to understand Descartes' view. It should be noted that in his metaphysics Descartes pushes the notion of matter as extended even further away from the older Aristotelian paradigm, by coupling it with his development of the *res cogitans*. In order to get a full grasp of Descartes' own view of the *res extensa*, therefore, it is useful at this stage to move on to consider the *res cogitans*. Specifically, as we will see in a moment, what will help us get a better grasp of why extension is an *essential* property of matter, rather than just another property, becomes even clearer once we discuss Descartes' idea that sensory perception is a purely intellectual activity. A striking consequence of this is that sensory perception no longer involves the body, which is then transformed from an animate entity to a purely mechanical portion of *res extensa*. This is the seed of the stark distinction between the conceptions of the mental and the material which resonates to this day. By the end of our discussion of the *res cogitans*, we should thus have

[2] Galileo's view is that the subject produces sensory qualities in her body as stimulated by the outside world. This changes dramatically with Descartes as we will see below.

achieved a conception of the mental and the material and how they are distinct (and independent) for Descartes.

The question we need ask at this stage, is how does Descartes reach the conclusion that the essential attribute of the mind is thought? In the *Mediations on First Philosophy* (1996; work originally published in 1641), Descartes embarks on a process of systematic skeptical doubt through which he calls into question all the certainties he deemed to have, he writes the his aim is 'demolishing all those falsehoods that I accepted as true in my childhood, and the highly doubtful nature of the whole edifice I had subsequently based upon them' (Descartes, 1996 from Patterson p. 71). He states that the purpose of this project is to rebuild the foundations of knowledge based only on those claims that are absolutely certain. Specifically, these certain foundations are the ones that survive the scrutiny of the most radical skepticism. This reflects Descartes' desire to move beyond the limits of the Aristotelian tradition, upheld by Christian philosophers such as Aquinas, which rely on sensory perception as the foundation of knowledge (James, 2000). As a consequence, the first thing that Descartes brings under skeptical scrutiny is the knowledge we acquire from the senses. Specifically, he scrutinises the possibility that sense perception may yield illusory data under unfavourable conditions, and then he moves on to considering the indistinguishability of sense perception while dreaming *vis à vis* while being awake. These two scenarios allow him to cast some doubt on the reliability of sense perception in general, but are not sufficiently pressing to wholly eradicate the certainty of knowledge obtained by the senses. As a result, Descartes introduces the Evil Demon, a vicious creature whose existence is devoted to deceiving the meditator, so that the perceptions of external things amount to nothing other than delusions. The role of the Evil Demon is to solicit perceptions of objects in the external world which are in fact non-existent. Under the spell of the Evil Demon, the doubt that is cast upon sensory perception is unescapable and, as a result, the meditator cannot have any certainty as to the external world.

A crucial aspect of this reasoning is that Descartes' scepticism about the external world, as acquired by the senses, also calls into doubt the existence of his own body. He writes, 'I shall consider myself as not having hands or eyes, or flesh, or blood or senses, but as falsely believing that I have all these things' (Descartes, quoted from Patterson, 2000, p.72) The Evil demon, therefore casts doubt on the meditator's prior certainty that he has a body. With this, Descartes

places the body within the realm of the external world, as opposed to the more common sense view, as well as to the ancient view (Burnyeat, 1982), according to which "external" meant external to the embodied person, thus external to the mind-body composite. We lose the mind-body composite and instead we are left with the illusion of a body.

In the *Second Meditation* (1996), however, Descartes finds certainty in his own existence. He reasons that no matter how much the Evil Demon deceives the meditator, there must nevertheless be an I to be deceived. While the Evil Demon can deceive the meditator as to the correctness of her thoughts, she just cannot doubt that she is thinking. The meditator, at the very least, *seems* to see, to hear, to taste. She cannot doubt that there is an I who doubts, who understands, who desires, who imagines, and it is in virtue of this intellectual ability to think that one's own existence can be stated with certainty. In a world of illusions, we know we exist in virtue of the fact that we think. From this perspective, sensory activity becomes purely intellectual, just as much as holding a belief or a desire, just as much as imagination. Descartes states explicitly that 'what is called "having a sensory perception" is strictly just this, and in this restricted sense of the term is simply thinking' (Descartes from Patterson, 2000, p.73). Descartes, therefore, identifies his own existence with his ability to think, *cogito ergo sum*, an ability that subsumes all the activities that are *internal* to the mind such as perception, imagination, judgements, beliefs, desires, and so on. As a result, he concludes that the substance that he is made out of is essentially a thinking substance. In other words, it is by following this reasoning that we obtain the Cartesian conception of the mind as being essentially defined by thought.

At this stage we can proceed to spell out some features of the mental, namely of the mind and its states. Patterson (2000) suggests that according to the traditional interpretation of Descartes, the *Second Meditation* has been read as marking the arrival of *subjective truth*. We now have a conception of the mind as a realm where the way things appear simply is the way things are; the mind is thus 'transparent to itself, an inner realm which is infallibly knowable' (Patterson, 2000, p.74). As a consequence, a subject's statement about her own states of mind are absolutely true and cannot be the object of deception, they are fully secured from the possibility of error, as opposed to our knowledge of the world external to our minds.
It is here that we start to see the crack, typical of dualism, between the mental and the material as based on the different epistemic access we have with respect to them.

The discussion above brings up two important elements of Descartes' view. First, that sensory qualities exist solely within the bounds of the mind. Second, that the body itself, which as humans we consider a defining aspect of our existence is part of the external world. As James (2000) points out, Descartes' understanding of the body is parallel to his understanding of inanimate objects in the natural world. This is what Descartes called the *homme machine* (James, 2000), literally the machine-man. In effect, Descartes often uses analogies with machines, for instance the hydraulic machine, to describe the physiology of the body. This is a passage of crucial importance in the history of our understanding of body and mind, as it marks the shift from the conception of the mind-body as an embodied compound of organisms that was featured in the Aristotelian tradition to an understanding of body as wholly mechanical. This is a conception of the body that has still echoes in the contemporary debate, as we will see below with reference to physicalism, and which is often contra-posed with the apparent freedom that the mind has in imagination, for instance.

What we have at this stage is thus two distinct *conceptions* of what exists in the world: we categorise the world into extended material things and thinking mental things. The move to metaphysical dualism is based on having these distinct conceptions. According to Descartes, in effect, 'all the things that we clearly and distinctly conceive of as different substances (as in the case of mind and body) are in fact substances which are really distinct from one another' (Descartes *Meditations* quoted from Patterson, 2000, p. 96). Descartes justifies the move from his epistemic conception to a metaphysical view by arguing that the distinction that holds between our conception of mental and that of material is real, rather than merely formal.

> The distinction between the motion and shape of a given body is a formal one. I can very well understand the motion apart from the shape, and vice versa, and I can understand either in abstraction from the body. But I cannot have a complete understanding of the motion apart from the thing in which the motion occurs, or of the shape apart from the thing which has the shape; and I cannot imagine there to be motion in something which is incapable of possessing shape, or shape in something which is incapable of motion (…) By contrast, I have a

> complete understanding of what a body is when I think that it is merely something having extension, shape and motion, and I deny that it has anything which belongs to the nature of a mind. Conversely, I understand the mind to be a complete thing, which doubts, understands, wills, and so on, even though I deny that it has any of the attributes which are contained in the idea of body. This would be quite impossible if there were not a real distinction between the mind and the body (from Descartes *Meditations* quoted by Patterson, 2000, p.97).

Having a complete understanding is based on the notion of essential properties, as explained by Swinburne (2018) above. Descartes' complete understanding, which yields the real distinction between one substance and the other, is based on his conception of the mental and the material as defined by their respective essential properties of thought and extension. We simply cannot conceive of a mind that does not think and a body that is not extended. The fact that each substance can be completely understood by reference to its essential property tells us that they are two distinct substances.

1.3 Physics and Feeling *What it is Like*

The conception of mental and material that we use in the contemporary metaphysics of consciousness debate, albeit strongly influenced by Descartes' set up, differs from his in terms of content. Particularly, the way we conceptualise their distinction reflects Descartes' track in that the conception of the material is deeply influenced by science, while the mental captures an inner realm of experience. This, in the contemporary debate, boils down to the idea that the material is objective reality while the mental is subjective. The objective/ subjective divide is indeed something that was not directly discussed by Descartes (Farkas, 2008), but it is rather the fruit of a more contemporary analysis of the physical and the phenomenal. It is to this conception that I now turn my attention. I start by looking at the material and then move to the mental.

It should be noted at the outset that the evolution of our contemporary conception of the material, as discussed within the consciousness debate at least, is very much due to the

development of physicalism, and the conception of the mental that complements it is developed in parallel to it, often as a result of the challenges os anti-physicalists to these views.

1.3.1 Physics and the Physical

According to Stoljar (2001a) there are two conceptions of what counts as "physical": one is object-based and the other theory-based. The *object-based conception of the physical* ties the notion of the physical to that of an object. It expresses the idea that a property is physical just in case it is the type (he uses “sort”) of property required by a complete description of the intrinsic nature of a paradigmatic physical object and its constituents, or if it is the type/ sort of property that metaphysically or logically supervenes on those types/ sort of properties required by a complete description of the intrinsic nature of a paradigmatic physical object and its constituents (Stoljar, 2001a, p.257). For example, if a rock is a paradigmatic physical object then the property of being a rock counts as a physical property; moreover, if the property of mass is required by a complete description of the intrinsic nature of the rock then having mass is also a physical property. All properties that supervene on these kinds are in turn, physical. The *theory-based conception of the physical*, on the other hand, adopts the data that physics provides us with as the mark of the physical. It expresses the idea that a property is physical if it is the type of property that physics tells us about or if it is the type/ sort of property that metaphysically or logically supervenes on those types/ sorts of properties that physics tells us about (Stoljar, 2001a, p.256). For instance, physics tells us about the mass and charge of an electron and therefore the mass and charge of the electron are physical properties that satisfy the theory based condition for what counts as physical. Moreover, the property of being a table for example, is to be considered a physical property just in case it either metaphysically or logically supervenes on the properties that physics tells us about. The physical, from this perspective, should be understood as the realm that is wholly characterised by all those properties that physics tells us about, and those that supervene on them. Anything that is either not accounted for in physics or does not supervene on it, therefore, must be considered as laying outside the physical realm and it is therefore not material. Mathematics and moral values are good examples of entities existing outside the material realm, as Papineau (2000, p. 7) points out.

The theory-based conception of the physical is the one that is most commonly adopted in the literature, as it is the one that physicalists most commonly employ (for notable examples see Jackson, 1982; Papineau, 1993; Lewis, 1995; Chalmers 1996) and for this reason it is the conception that I am going to work with.

Both Russell (1927) and Whitehead (1925) have discussed the idea that the purpose of physics is to uncover the causal skeleton, thus the structure, of the world. To uncover the causal skeleton of the world means to identify and describe the web of relations that characterise how the entities that physics tells us about interact with each other. An immediate consequence of this is that entities themselves are wholly described in relational terms. Let us make an example. If we ask a physicist what an electron is, she will explain that an electron is that entity that *has* mass and negative charge. And what are mass and charge? Charge is the physical property that defines whether a certain entity attracts or repels other entities according to how it is charged. Given that the electron is an entity that has negative charge it repels other electrons on the basis that they have the same charge, while it will attract protons which are positively charged. In the same key, the scientist will tell us that the mass of an atom, namely the number that describes the value of its mass, defines the degree of its resistance to acceleration when a force is applied to it, while also determining how strongly the atom is attracted to other atoms. The example shows that the physical description of an electron conveys a story about how the electron behaves, how the electron relates to other entities, thereby allowing the physicist to describe the web of relations that hold between our electron and other relevant physical entities. This way the causal structure of the world is uncovered. As such, physics tells us about relational properties only, without being able to account for what things are in and of themselves, intrinsically. Stoljar (2001a) also notes the issue and refers to Blackburn's idea that 'science finds only dispositions [relations] all the way down' (Blackburn, 1992, pp. 62-63). This is an issue that may be problematic.

Nonetheless, the causal skeleton of the world is an essential element of our present-day conception of the material, given the *completeness of physics* principle (Papineau, 1993). Papineau (1993) explains that the completeness of physics is to be understood with reference to the fact that the physical world is causally closed. In Papineau's own words, physics is complete in the sense that 'all physical effects are determined or have their chances determined by prior

physical events according to physical law' (Papineau, 1993, p.16). In other words, that 'physical causes completely suffice for physical effects, or fix their chances: no other causes are required to bring about physical effects' (Crane, 1995, p.216). For a physical effect to obtain, only physical causes need to occur prior to it, and this is what renders physics complete. Let us make a contrast with a different field of enquiry to better grasp this notion. The psychological realm cannot be said to be complete because certain psychological effects could not occur in the absence of other neurological and physiological factors which act as causes for the psychological state in question. Take the case of hallucinations, where a subject misrepresents reality, for example by seeing flashing colours where there is in effect no such event occurring. Hallucinations of this sort can often be caused by lack of sleep or the intake of drugs, that is non-psychological causes. This contrasts with the physical realm whereby physical effects are necessarily and sufficiently brought about by physical causes, thereby "completing" the realm.

Given this conception of the physical, it follows that the body is to be considered a wholly physical entity. This is because, firstly, the body can be said to either logically or metaphysically supervene on the sort of properties that physics tells us about, as according to Stoljar's (2001a) theory-based definition of the physical. Secondly, because the body functions within the bounds of the completeness of physics. The same goes for other biological organisms, chemical processes, physical micro-entities such as electron, cosmic bodies such as planets and so on. The Cartesian conception of the *homme machine*, therefore, resists centuries of scientific research and philosophical contemplation and so do other elements of mechanistic theory that we still find in the guise of the completeness of physics, and which are now applied to all organisms, chemical processes, micro-entities and cosmic bodies *alike*.

But more needs to be said about the nature of the physical, and this is something rooted in the essential purpose of science itself: to produce an objective account of reality. That is, the purpose of science is to tell us about the world from an *objective* perspective, to tell us the story of the world as disentangled from our experience of it, that is external to it and thus mind-independent. An objective perspective is one that can be equally occupied by anyone and from which the same set of information, the same description of reality, is obtained. This knowledge is "public" so to say, because it can be had by anyone who steps into that perspective, and can be wholly accounted for in third-personal terms. As a consequence, we could say that our

contemporary conception of the physical is strongly characterised by the objective nature of science. We will discuss this at length in chapter 4. This objective feature of the physical highlights a feature that is central for distinguishing it from the mental which, in contrast, is defined by its essentially subjective character. It is to the mental that I now turn my attention.

1.3.2 The Phenomenal and W*hat it is Like* to Undergo Experience

The most popular way to construe the mental in present-day discourse originates from Nagel's (1978) notion that mentality is defined by the fact that *there is something it is like* to have it. More specifically, for Nagel (1978) to say that an organism is conscious, or has conscious experience, means that the organism in question has some sort of *phenomenology* such that it *feels* something for that specific organism to be itself and undergo the experience. Phenomenal experience, therefore, is the rich qualitative felt aspect of our lives that we are so familiar with and which in the debate we commonly capture through talk of phenomenal properties, or *qualia*.

Let me describe my phenomenal experience as I write this text to better clarify what phenomenal properties are. As I type on my computer I see a series of words appear, letter after letter, black on white, onto the screen I am looking at here in front of me. Coming from the open window I hear the soft sound of leaves as they are moved by a summer breeze and this blends in with the low sounds of the music coming from my speakers. I feel the hardness of the letter-keys under my finger tips as I move them from key to key, and I also feel extremely hot because the temperature today is over 33 degrees, so I try to cope with the heat by drinking some warmish mint tea that tastes very good. The phenomenal properties that characterise my experience, and which are in turn crucial to our conceptualisation of the mental, are seeing the blackness and whiteness of the letters, feeling the hardness of the letter keys, hearing the softness of the sounds, tasting the sweetness of the mint tea, and so on. They pick out a set of qualitative features that characterises *what it is like* for a subject such as myself to be in a certain mental state, what it is like for me to undergo this experience and, furthermore, allow me distinguish this experience from another. These phenomenal properties are those properties which, as Goff (2017) often puts it, constitute our rich inner life. The set of the phenomenal properties that characterise experience is what Nagel (1978) calls the *subjective character of*

experience. This is the central concept that underpins our contemporary understanding of the mental and its distinction from the material.

The main feature of the subjective character of experience, as Nagel (1978) points out, is that it is essentially connected to a *single point of view*. This should be understood as a subject's own phenomenal perspective on the world, hence how a subject experiences the world through the lens of the phenomenal qualities that characterise it. It is undeniable that an experience of mine, such as the one I undergo as I am typing this paper, is somehow duplicable by other subjects who are constituted the way I am and who engage in similar activities and in similar environmental conditions. However, in a certain sense this token experience singularly captures the way *I* am undergoing it. A phenomenal point of view, therefore, can only ever be occupied by a single experiencer. This can be further understood by looking at James' (1890) description of a personal consciousness, which in effect complements Nagel's. As you read James' (1980) passage below, please substitute the term "thought" with our understanding of "phenomenal experience".

> The only states of consciousness that we naturally deal with are found in personal consciousness, minds, selves, concrete particular I's or you's. Each of these minds keeps its own thoughts [phenomenal experience] to itself. There is no giving or bartering between them. No thought [phenomenal experience] even comes into direct sight of a thought [phenomenal experience] in another personal consciousness than its own. Absolute insulation, irreducible pluralism, is the law. The breaches between these thoughts are the most absolute breaches in nature. (James, 1890, p226)

My phenomenal experience exists within the bounds of my own personal consciousness, it belongs to my own point of view, my mind, and it is by nature impossible to come to share it with other personal consciousness or have somebody else step into my phenomenal perspective. The boundary that encloses it simply cannot be breached by nature, as James puts it.

It follows that occupying a single point of view has strong implications on the *accessibility* of the phenomenal experience that characterises our notion of the mental, in the sense that it can only be accessed by the experiencer alone. The mental is then characterised as a realm that is privately accessible. And this is certainly an important feature of phenomenal experience and what it means to say that it is subjective. This special accessibility of the subjective mental realm stands in stark contrast with the inherent objectivity of the material. To further explain this distinction Nagel (1974, p.443) writes that 'entities in the material realm are observable from the point of view but external to it, hence they can be comprehended from other points of view also.' For example, he continues '[l]ightning has an objective character that is not exhausted by its visual appearance, and this can be investigated by a Martian without vision' (Nagel, 1974, p. 443). On the other hand, the visual experience of the same lightning can only be had by certain phenomenal agents that are constituted like us, that have neurones and eyes for instance, and moreover that the token experience of the lightning is my own and nobody else's. It cannot be "seen" by any other person as it seen in my mind, as James above suggests.

To sum this up, the subjective character of the mental is generally conceptualised with reference to its qualitative nature and to its special accessibility. In contrast, the objective character of the material is understood by reference to principles such as the completeness of physics and by being equally accessible to anyone.

In addition to the above, I would like to add that the way these concepts are acquired also helps to shed better light on how we conceptualise the difference between the mental and the material in contemporary literature. I talk at length about this in chapter 4, but for now let us briefly look at Jackson's (1982) famous Knowledge Argument to get an idea of the issue. It should be noted that the Knowledge Argument is usually employed by anti-physicalists to establish that facts about consciousness cannot be reduced to physical facts. For our purposes, however, the argument helps us get a better grip on how we normally conceive of mental properties, the relationship between these and the experiencing subject who has them and how the difference between these properties and physical ones is usually conceptualised.

Mary, the main character of the knowledge argument, is the most knowledgeable scientist in the world. She has all the material knowledge one can have about colour sight, she knows everything there is to know about wavelengths and how they bounce off surfaces and through

space; she also knows everything about the neurophysiology of vision and how redness for instance comes to be perceived when a subject is exposed to right environmental condition that should produce a red- experience, and how this differs from other colour experiences. Unfortunately for her, however, she was born and raised in a black and white room and she has never seen any other colour in her life. One day Mary is released from her colourless prison and finally sees the blue of the sky, the green of the grass, the redness of roses. For our purposes it is interesting to note that only when she gets exposed to colour can she obtain knowledge of *what it is like* to see colour. Notwithstanding that she had all the objective material facts about colour and colour vision, she lacked the relevant, first personal phenomenal experience that would allow her to form a subjective perspective of the colour experience in question. As a consequence, Mary teaches us that material knowledge is acquired differently than mental knowledge, and in this sense we have very different epistemic position with respects to the mental and the material.

1.4 A Final Reflection: The Mental, the Material and a Category Mistake

We have seen that in present-day discourse we define the mental as the subjective realm, marked by its qualitative phenomenal character and the material as the kingdom of the objective, governed by the principle of causal closure. We have also seen that the notions of subjective and objective have not always been used to mark the distinction between the mental and the material. The distinction was in effect absent in Aristotle's account, where the soul was understood as the principle of life, and identified by certain powers such as nutrition or growth or sense perception, processes that we now label physiological (thus physical). The distinction between mental and material is therefore due to more recent philosophical developments, stemming from the ideas initially put forward by Descartes. Introducing the now famous *ghost in the machine* concept to criticise Descartes' *homme machine*, Ryle (1949) objects that the view wrongly imposes a polar opposition between the mental and the material. More specifically that this distinction is based on a category mistake between things that are in truth, ontologically speaking, not distinct. It is interesting to note that Descartes' dualist project was, in effect, driven by his desire to make space for his new physics (Patterson, 2000), a physics that was anti-Aristotelian and according to which nature operates wholly mechanically. Descartes' development of the distinction between mental and material was thus a necessary step for him

to introduce and develop his physics, while also preventing the mind from perishing under the mechanistic rule and eventually losing its immaterial character and peculiar freedom. In light of this, it is interesting to ask how much the distinction that Descartes introduces genuinely captures a real distinction between the two substances he posits, that is a distinction that effectively goes beyond the conceptual. In any case, the knowledge of Descartes' own motivations helps us to shed light on why sense perception was put under such skeptical scrutiny and why qualities were wholly transferred in the mind under the guise of an intellectual activity, while the body becomes part of the mechanical world the understanding of which belongs to science.

Ryle (1949) also notes that Descartes' conceptual bifurcation very much relies on the contraposition of the internal and external contraries, where everything that is mental is internal to the subject whereas the material is whatever is external. Of course this is a metaphorical use of the terms as the mind, not existing in space, simply cannot be internal to anything, Ryle emphasises. However, this certainly led us to conceive of the mental and of the material in terms of contrary opposition, by which the mind is everything that matter is not and vice versa (Ryle, 1949). The conception of mental as internal and the body as external, together with the rest of the world, is indeed one that has had a strong influence in post-Cartesian research on the mind-body problem and which we find embedded in the ideas of mental-as–subjective and material–as–objective framework we use today. Thinking in these terms imposes an inescapable parallelism between the mind and the body, which lays the ground for the problem of consciousness and the difficulty in solving it that characterises the debate. We delve deeper into this issue in chapter 2 where we discuss the mind-body problem, in chapter 4 where we discuss the distinct nature of phenomenal and physical concepts and their relation to distinct perspectives, and in chapter 5 where we discuss their mutual exclusivity.

Before concluding, it is important to note that in the contemporary debate we have seen a shift in terms of how we characterise the mental: while for Descartes the mental is an entirely intellectual activity, we now focus on the idea that there is s*omething it is like* to be conscious, something *felt,* that more closely relates to the body in a certain sense. The boundary between the mind and the body, however, does not appear to be particularly mitigated by this more

contemporary understanding, although it seems to manifest a desire to move towards a more monistic and uniform understanding of the body and the mind, as we will see in the conclusion.

Lastly, emphasis has to be put on the fact that the distinction between the mental and the material originates from a difference in how we *conceive* of them, thus a conceptual difference, which is rooted in the different epistemic positions we hold with respect to what we identify as mental or material. This conceptual distinction then finds its way into a metaphysical realm with Descartes' argument of a real distinction based on a complete understanding. The move from conceptual to metaphysical we find in Descartes constitutes the seed for the development of the various ontologies that followed Descartes' work, although each view manipulates the metaphysical according to the set of assumptions they want to defend, whether the fundamentality of the casual closure of the physical or the fundamentality of sense perception for material and mental monists respectively, and the consequent need to explain the non-fundamental property, as we will discuss in the next chapter.

CHAPTER 2:

THE MIND-BODY PROBLEM

The mind-body problem is the problem of appropriately understanding how the mental and the material (co-)exist and interact the way we ordinarily know them to. That is generally understood as the difficulty of explaining how our mental lives, specifically the phenomenal (subjective and mind-dependent) aspect of experience relates to the physical body and particularly the brain (objective and mind-independent); thus more broadly how the phenomenal fits into the world that physics tells us about. This is a problem that, in a form or another, affects all views partaking in the consciousness debate to date. As such, the mind-body problem should not be understood as a single problem, but rather as a closely knit family of questions about the relationship of the mental and the material.

Historically, the question of mind-body relations originates as a response to Descartes' dualism, as the problem of understanding how the distinct and independent mental and material substances ultimately interact, specifically how the mind may cause the body to move (and vice versa how the body may affect the mind). A closely related version of the problem arises for traditional forms of mental and material monists, whose attempt to solve the dualist interaction problem by positing the existence of a single substance ultimately leads them into the challenge of explaining the nature and existence of the non-fundamental material and mental properties respectively. As with the rest of this thesis, in this chapter I focus on material monism and how the mind-body problem applies to it, but note that the same types of criticisms can often translate to a mental monist framework (at least in terms of the structure of the argument). Moreover, in the context of monism the challenge can take either an epistemic focus, as per the Explanatory Gap (Levine, 1983), or an ontological stance as expressed in the Knowledge Argument, briefly outlined in the previous chapter, and in the Conceivability

Argument we treat in § 2.3.3, below. The latter is the more modern incarnation of the mind-body problem, and it is widely discussed in the contemporary literature under the label of the Hard-Problem of consciousness (Chalmers, 1995), which we look at in §. 2.3.2.

In my view, it is the assumptions as to the metaphysical nature of the mental and the material that each view adopts which leads them straight into their bespoke version of the mind-body problem. For instance, the physicalist assumes that the world is all of one kind and that this kind is physical, and it is this basic assumption that attracts its own version of the mind-body problem, namely to explain how the mental fits into the physical picture of reality. In contrast to this, I feel the perspectival neutral monist has the potential to overcome this sort of problem, because her methodology is based on questioning the assumptions as to the metaphysical nature of the mental and the material adopted by her predecessors, understanding in depth the problems that their ontological pictures attract, and thus allows her assumptions and ontological commitment to gradually emerge out of her reasoning.

There are three reasons, therefore, why I devote this chapter to expanding on the mind-body problem: (1) fully understanding the consequences of the traditional approaches and of the assumptions they adopt, (2) to understand the depth of the problem that affects the research on consciousness on traditional views such as dualism and physicalism from an ontological standpoint, and (3) to set the ground for understanding how a neutral monist tackles this issues in general, and specifically whether and how the perspectival neutral monist has the potential to solve or avoid these issues. I will come back to point in chapter 7 and in the conclusion.

In this chapter I outline the modern origin of the mind-body problem, looking at the challenge that interaction poses to Descartes' dualism in section 2.1. Then, in section 2.2 I move to an overview of physicalism, its assumptions and the reductive strategy adopted by the physicalist. In section 2.3 I focus on anti-physicalist arguments and I introduce the Explanatory Gap and the Hard Problem. In section 2.4 I discuss how some alternative views attempt to overcome the issues that the physicalist faces by adding an *extra ingredient* to their ontology. Lastly, in section 2.5 I round up by emphasising some elements of the discussion that we build upon later in the thesis.

2.1 Descartes and the Problem of Interaction

As discussed in Chapter 1, Cartesian dualism is the view that there exist two distinct and independent substances, each defined by an essential attribute: extension for the physical substance (*res extensa*) and thought for the mental substance (*res cogitans*). Human beings are thus composed of these two substances, 'joined and united' (Descartes' *Discourse on Method* (1637) from Rozemond, 2003, p.350). Understanding how these substances are ultimately joined and united is the root of the problem of interaction for the Cartesian dualist, as it is unclear how a causing event occurring in one substance may bring about an effect in the other substance. It is unclear, for instance, how a (mental) desire to eat ice-cream or a belief that it is raining may bring about (physical) actions such as buying and eating ice cream or picking up an umbrella before leaving the house. One may re-formulate the problem conversely, by asking how the physical event of my skin burning can cause me to feel pain, or having certain light waves hit my retina can cause me to see the colour red. Richardson (1982) dubs this the Heterogeneity Problem, highlighting that it rests on the Cartesian dualist positing the mental and the material substances as 'utterly diverse in nature, [such that][3] it is unintelligible how mind could act on body or body on mind' (Richardson, 1982, p.20).

Princess Elizabeth of Bohemia (1618–1680) explains the issue with great clarity in a letter she addresses to Descartes.

> Given that the soul of a human being is only a thinking substance, how can it affect the bodily spirits, in order to bring about voluntary actions?
>
> The question arises because it seems that how a thing moves depends solely on (i) how much it is pushed, (ii) the manner in which it is pushed, or (iii) the surface-texture and shape of the thing that pushes it. The first two require contact between the two things, and the third requires that the causally active thing be extended.

[3] I added text in square parenthesis for clarity

> Your notion of the soul entirely excludes extension, and it appears to me that an immaterial thing can't possibly touch anything else (Elizabeth from Descartes and Bennett, 2017, p.1).

From the passage above it is clear that the kernel of the problem lays in the brute metaphysical distinction between the mental and the material, specifically on the grounds that the essential attributes which define each of the two substances, namely thinking and extension, are *mutually exclusive* and hence *incompatible* for the purposes of causation. I want to stress Elizabeth's intuition regarding the mutual exclusivity of the mental and material substances, as this is a central issue in our argument for a neutral ontology, in chapter 5.

Elizabeth explains that causation requires contact between two things and that contact requires things to be extended; that is for some object A to push some other object B both must have an extended surface upon which the contact required for causation can occur. Elizabeth's reasoning is rooted in the idea that the causal process is mechanical, in line with the view of nature that was being developed during the time Descartes and Elizabeth lived and worked, as discussed in chapter 1. From this, she then states how difficult it is to understand how a non-extended, immaterial mental substance may bring about effects in a physical substance at all.

Elizabeth's perspective is particularly interesting because it seems, to me, to anticipate the causal closure of the physical principle (introduced in the first chapter and which forms the basis of physicalist reductionism which we will look at in the next section, §2.2). This is because her challenge to Descartes is based on tracing 'the causal ancestry or posterity' (Kim, 1997, p. 282) of a mental or material event, and reasons that effects in extended realm are most coherently brought about by extended causes on the grounds that both cause and effect need to be extended. That is, it seems she reasons that being extended, thus being located in space, having shape and size for instance, is a necessary and sufficient factor to bring about effects in relation to other extended things, thus things which are also spatially located, have shape and size. In other words, that the mechanical causal process of the *res extensa* 'never takes you outside the physical domain' (Kim, 1997, p.282), where we should substitute the term "physical" with "*res extensa.*" Her criticism thus revolves around the idea that supposing that minds can bring about effects on extended things would breach this mechanical cause-effect process,

which characterises the extended (material) substance, and it would thus imply that the extended realm is not "causally closed," so to speak. It is in this sense that I feel that Elizabeth heralds a rudimentary version of the causal closure principle, where we understand physical effects to be brought about by physical causes (and not caused by non-physical events), such that the causal history of the physical 'never takes you outside the physical domain,' as Kim (1997, p.282) claims with reference to the causal closure of the physical.

Nonetheless, the substantial import of Elizabeth's passage is an acute reflection on the two assumptions underlining Cartesian dualism, namely that (i) there is a physical realm "out there" made up of extended particles/ entities that move by obeying to mechanical laws of nature, as shown by science, and that (ii) there is an immaterial realm that seems to fail to be subject to mechanistic processes, as suggested by the mind enjoying free-will and being immortal. What is crucial to Elizabeth's passage is that it highlights the severe tension between the two assumptions, and identifies this tension as the basis for the mind-body problem for the dualist.

Descartes (1996) disagrees with Elizabeth and claims that mental and material interaction occurs notwithstanding the brute distinction between the two substances. He argues that interaction takes place in the pineal gland and that the mechanics of interaction may simply be beyond the reach of our intellect (Richardson, 1982). In a certain sense, therefore, Descartes agrees that mental and material substances are inherently incompatible, while simultaneously rejecting the idea that the physical world enjoys "causal closure" as understood by Elizabeth. He thus favours the causal efficacy of the mental over the physical, thereby accommodating his assumptions about the distinction between the mental and the material substances. Whether Descartes motivations are sufficient to solve the mind-body interaction problem is beyond the scope of this chapter, but it suffices to say that to this day the Cartesian dualist is still in the process of producing an adequate answer to the problem.

2.2. The Mind-Body Problem and Monism

Kim (1979, p.31) tells us that '[t]he fact of psychophysical causal relations has been thought to constitute part of the data that must be explained, or at least explained away, by any

satisfactory theory of mind. Few philosophers have called this datum itself into question'. The value of this datum, together with the impossibility to produce the required sort of explanation once the dualist assumption is accepted, motivates any form of monism. In effect any kind of monism takes this issue head-on by assuming that reality is all of one kind. Therefore, by removing the essential distinction between mental and material substances, they hope avoid the problem of interaction: if reality is all of one kind, then we cannot ask how one substance may causally affect the other, and the question of interaction simply does not (seem to) arise.

The question for the monist, one the other hand, can be summarised as the problem of having to explain how some properties with very specific characteristics, such as the phenomenal character of experience, can arise out of a fundamental substratum that entirely lacks those characteristics, such as the barren world of quarks and electrons. We thus have a shift of focus from interaction to the existence and metaphysical nature of a set of non-fundamental properties and its potential relation to the (different) base.

2.2.1 Physicalism and the Causal Closure Assumption

Let us first define physicalism as follows

Physicalism: everything is physical

The term physical here is understood according to its theory-based conception (Stoljar, 2001a) seen in chapter 1 (§ 1.3.1), where a property is physical if it is the type of property that physics tells us about or if it is the type/ sort of property that metaphysically or logically supervenes on those types/ sort of properties that physics tells us about (p.256). The mass and charge of an electron are two paradigmatic examples of physical properties conceived as theory-based, and also the functional and physiological processes that occur in the brain and throughout the body are understood as physical as they supervene with metaphysical necessity on the more fundamental physical properties such as mass and charge. According to the physicalist, phenomenal properties (if real) are also physical and thus they must supervene with metaphysical necessity on the physical.

The problem for the physicalist essentially revolves around the central assumption that underlies their position: the completeness of physics principle. Let us recall, this is the idea that the physical world is causally closed such that all physical effects have a physical cause. Specifically, as introduced in chapter 1, that 'all physical effects are determined or have their chances determined by prior physical events according to physical law' (Papineau, 1993, p16). It should be noted that the completeness of physics emerged as a doctrine only in the 20th century and thus earlier versions of materialism were not motivated by this. In the contemporary literature, however the doctrine is a central issue in physicalism.

The weight of the completeness assumption can be said to rest on the success of the physical sciences to produce a verifiable account of the world, based on empirical evidence. It seems, in fact, undesirable for a philosopher to deny the causal closure principle (Goff, 2017) in view of scientific progress, as doing so could point towards some sort of denial of the knowledge we obtain from the physical sciences, and thus also constitutes an undesirable move on the part of any philosopher who desires to safeguard the veridicality of the data that science provides us with. Nonetheless, once the causal closure principle (and its fundamentality) is assumed, the physicalist faces the task of explaining (i) the nature of mental/ phenomenal properties and (ii) how these relate to the underlying fundamental physical base. Tackling this question comes with two constraints for the physicalist: (a) safeguarding the causal closure of the physical and (b) showing that mental states are somehow causally efficacious in the world, thus avoiding epiphenomenalism (the view that phenomenal states are caused by physical states, yet bear no causal role towards physical states themselves). From the two points above, it is essential for the physicalist to produce and adequate account of the mind, for failure to do so would entail her failure to account for the way a subject's experience is connected to its physical states. In other words, safeguarding the causal closure of the physical in not enough, for the physicalist must also show that physical states cause phenomenal states and that phenomenal states ultimately have some causal relevance in the physical world.

A physicalist can safeguard the causal closure principle, and attempt to avoid epiphenomenalism, by adopting an eliminativist approach and denying the existence of phenomenal properties altogether (Dennett, 1988). This however *eliminates* the reality of phenomenal experience. Alternatively, she can opt for a non-reductive or emergentist position

and claim that phenomenal properties exist and that they somehow emerge from the underlying physical base. This results in phenomenal properties being 'nomological danglers' (Smart, 1959), or epiphenomenal, and thus not causally efficacious in and of themselves, ultimately pushing phenomenal properties out of the physical world in the sense that they would arise from it but have no effect in within it. This is an unwanted result because, as per Kim (1979) above, it is clear that my feeling thirsty brings about my reaching out for a drink and therefore that mental states seem to have some sort of causal efficacy in the physical world. It seems that the safest route to safeguard the causal closure principle, in conjunction with admitting that mental properties are casual efficacious in the physical world, is to adopt a view that allows us to identify phenomenal properties with the underlying physical base. This is what the reductive physicalist attempts to do.

2.2.2 Reduction

In general, reduction obtains when a set of higher level properties Xs is nothing over and above the more fundamental set of properties Ys (Kim 1993, p 275). The locution *nothing over and above* can be grasped in terms of an identity claim, such that the Xs just are the Ys. For instance, my phenomenal state of feeling pain *just is* the physical state of C-fibres firing in my brain. In a certain sense, we are supposed to understand the *nothing over and above*, or identity, relation that holds between pain and C-fibres firing the same way we understand the relationship that holds between Lewis Carrol and Charles Lutwidge Dodgson or water and H2O (Kim, 1993, p. 333 from Van Riel p.2). This is the strategy adopted by those physicalists that defend the identity theory (for example functionalists such as Lewis, 1966; and Armstrong, 1968). By identifying phenomenal properties with physical properties, they effectively admit the existence of the phenomenal and simultaneously deny its independent and distinct, non-material nature (as conceived by the Cartesian dualist). The job that reduction sets out to do is, therefore, to explain how the mental ultimately exists as a material entity. Specifically, as expressed in simple terms by Crane (2001, p.54), reduction is that procedure that the physicalist employs to demonstrate that, and in what way, mental things are physical. In reduction, Crane (2001, p. 54) continues 'we start off with the "target" entity, X, and find a reason for identifying X with Y. Our reduction tells us something we didn't know about X: that it is Y'.

The locution *nothing over and above* highlights a further feature of reduction, namely that it is a relation of dependence between a set of higher level properties on a set of lower level properties; while identity is a symmetrical relation, reduction rather tells us about an ontological hierarchy that holds between the two sets of properties in question. This is inspired by reduction as seen in the sciences, for example where temperature is taken to be the motion of molecules. Here the motion of molecules is ontologically prior to temperature and, moreover, it is the motion of molecules that renders the phenomena of temperature intelligible or predictable. This example allows us to clarify what ontological reduction is and how it differs from standard identity, but also helps us introduce another widely used type of reduction: explanatory reduction. Explanatory reduction obtains between two sets of theories, a set that explains the Xs and the other set that explains the Ys, and where Xs can be used to explain away (show the truth, make intelligible) the Y theory. Reductive physicalists are generally ontological reductionists. However some arguments, such as the Explanatory Gap we discuss in the next section, show that explanatory reduction of the mental to the material is particularly difficult. Some physicalists thus embrace that explanatory reduction of the mental to the material is not possible, while nonetheless defending ontological reduction. We will discuss this approach in more details in chapter 4.

With this in mind it should be clear how the reductive physicalist hopes to solve the problem of the nature of mental properties, while also attributing them causal efficacy within the context of the causal closure of the physical. On the grounds of supervenience (or identity or realisation or constitution), the idea is that if the mental is nothing over and above the physical, thus the mental simply is physical, then the issue of causation is solved because the causal story is fully accounted by the physical story. My *feeling thirsty* is nothing other than its underlying **thirst physical structure** and thus my action of picking up the bottle in front of me and drinking from it to quench my thirst is caused by the **thirst physical structure**, rather than the fact that I feel thirsty per se (I am simplifying for clarity, but this should illustrate the point). This means that in order to explain how my feeling thirsty causes me to reach out for the glass of water I only need to look further down in the ontological hierarchy and find the physical structure that underlies my feeling thirsty.

2.3 Anti-Physicalist Arguments

The physicalist thesis is undermined by a number of arguments that target the necessary connection posited between the physical and the phenomenal, showing instead that there is an appearance of contingency between the two which should block the inference to the physicalist conclusion about the physical nature of the phenomenal. In this section, I examine Levine's Explanatory Gap (1983) and Chalmers' (1996) Conceivability Argument, the former focuses on epistemic concerns whereas the latter on the ontological picture.

2.3.1 The Explanatory Gap

Levine's *Explanatory Gap* (1983) shows the existence of an epistemic gap between the physical and the phenomenal. According to Levine, the full account of the physical-causal-functional structure, no matter how detailed, leaves us in the dark as to why that particular structure should feel, qualitatively speaking, the way it does. For example, the epistemic gap with reference to a pain experience is captured in the following terms: why should C-fibres firing in our brain feel the way pain ultimately feels when a subject undergoes the experience?

The Explanatory Gap therefore captures the problem that the description of the physical-causal-functional structure that the physicalist posits as the ontological underpinning of the feeling of pain is simply not explanatory of the pain experience itself. Rather, it leaves out something crucial unexplained, namely the qualitative character of the pain experience. According to Levine (1983), in effect, what is most worrying for reductive physicalism is that an explanation of the physical, causal and functional structure of experience does not make the experiential phenomena in itself intelligible, which is exactly what an explanation should do. While the physical account does provide us with information as to how the pain was caused- for instance by touching the hot tray in the oven- and about the functional/ physiological processes- for instance the skin rips open, C-fibres fire and a withdrawal mechanism is activated- it leaves us completely in the dark as to why that particular phenomenology should fit with that physical story. In other words, understanding the physical story does nothing to help us understand the phenomenal story. Levine's (1983) argument, therefore, targets the epistemic difficulty of explaining the phenomenal in terms of the physical (it targets explanatory reduction) and as

such it does not constitute a direct objection to physicalism. It does, however, bring up the central issue that the physicalist faces with great clarity. The force of the issue that the Explanatory Gap expresses has led many philosophers to infer the existence of a parallel ontological gap, see for example Nagel (1978), Kripke's (1980) and Chalmers' (1996) conceivability arguments (we examine Chalmers' in § 2.3.2 and 2.3.3 below), or Jackson's (1982) Knowledge Argument (mentioned in chapter 1).

Reductive physicalists can respond to the challenge posed by the epistemic gap by responding that there is no gap, whether explanatory or ontological, on the grounds that that there is an *a priori* necessary connection between the physical and the phenomenal (for example Lewis, 1966; or Armstrong 1968). Chalmers (2003) labels this kind of response Type-A physicalism. More recently, however, a different sort of response has been given by what Chalmers (2003) dubs Type-B physicalists, which we discuss in chapter 4 as we re-work their Phenomenal Concept Strategy. Briefly, Type-B physicalists admit that there is an epistemic gap between the phenomenal and the physical, however they deny that this entails an ontological gap on the grounds that the distinction between the properties is merely *conceptual* rather than metaphysical. They argue that the connection between the mental and the material is necessary but *a posteriori*, much like the connection between water and H2O (Kripke, 1980). They thus admit the power of the Explanatory Gap and then proceed to explain why and how such gap obtains. This way, they hope to block the inference to the ontological gap between the mental and the material, which was in turn supposed to block the reduction of the former to the latter.

The Type-B physicalist can thus be said to support a form of conceptual dualism which explains the appearance of contingency in terms of a conceptual distinction, in conjunction with a monist ontology of a materialist kind. As we will see, this position is structurally very similar to the position that the perspectival neutral monist seeks to defend with reference to the epistemic dualism plus a monist ontology, but different in that the neutral monist does not land on a materialist metaphysics.

2.3.2 The Hard Problem

The Hard Problem of consciousness raises the question of how there can be subjective experience, hence consciousness, at all. Why and how do brain processes give rise to subjective experience? The question is supported by the idea that all the brain processes we associate with experience could happen in the total absence of any phenomenal character, they could happen "in the dark" so to speak. The Hard Problem, often supported by the zombie version of the Conceivability Argument, therefore pushes the focus of the mind-body problem from explanation, as in Levine (1983), to ontology. That is, it raises the question concerning specifically how the rich inner life that characterises our lives as human beings could exist at all in a world that is fundamentally the way science tells us.

The Hard Problem was initially characterised by David Chalmers (1995), as he isolates it from a series of other "easy" problems. These differ from the Hard Problem because they are susceptible to the methods of science, hence helping us understand why the Hard Problem is so hard by making a contrast with these.

According to Chalmers (1995), good examples of easy problems of consciousness are:

- the ability to discriminate, categorise, and react to environmental stimuli
- the integration of information by a cognitive system
- the ability to report mental states
- the ability of a system to access its own internal states
- the focus of attention
- the deliberate control of behaviour
- the difference between wakefulness and sleep

As one can tell from the above examples the term "easy" is tongue in cheek as the explanation of a person's accesses to her mental states, or of the difference between wakefulness and sleep, are not inherently easily solvable. There are labelled "easy" merely on the grounds that they are tractable by employing scientific methods, and could in principle adequately be reduced to or be explained in terms of computational or neural mechanisms of the brain. The reason for this is that the easy problems concern abilities and functions of the

brain which are directly explicable in terms of the mechanism that underpins them, given that these abilities and functions are nothing other than the mechanisms which underpin them. Specifically, Chalmers (1995, p. 203) tells us that '[t]o explain a cognitive function, we need only specify a mechanism that can perform the function'. Moreover, he adds that 'the methods of cognitive science are well-suited for this sort of explanation, and so are well-suited to the easy problems of consciousness' (Chalmers, 1995, p. 203). In effect, it is precisely to explain these abilities and functions that cognitive science exists in the first place: to develop models and theories of brain processes that can make intelligible the mechanism that underpins functions and abilities, such as the focusing of attention or the integration of information within a cognitive system; this is also based on the assumption that these kinds of functions and abilities are nothing other than the processes which underpin them. Chalmers thus trusts the power of cognitive science to adequately tackle these problems, which in turn renders them easy.

On the other hand, the issue with the Hard Problem is that, when it comes to the nature of consciousness and its place in nature (as characterised by physical science), the reductive methods of physiology, neurology and cognitive science simply do not do the trick: they have failed- and will fail- to explain why there should be subjective experience at all in conjunction with the neutral and physiological processes that occur in the brain. The hard problem is hard particularly because it resists the methods of science, as Chalmers (1995, p.203) writes, 'because it is not a problem about abilities and functions.' If we construe consciousness in terms of there being *something it is like* for a person to feel something, it is the nature of the feeling that we seek to explain, the felt quality that attaches to any experience, which is something more than the processual mechanism that we pair with the physical process of seeing red or feeling pain.

Chalmers (1995) thus starts from the epistemic difficulty to explain consciousness in terms of a physical base, thus embracing that there exists an Explanatory Gap between the mental and the material as per Levine (1983), and proceeds by telling us that no physical explanation can *ever* produce an adequate explanation of why there should be consciousness associated with the physical states. He then concludes by suggesting that an adequate explanation of consciousness will necessitate the addition of an *extra ingredient* to our ontology, hence that the bare physicalist ontology is not sufficient for explaining the nature of consciousness. This is a suggestion that many philosophers embrace in order to overcome the

issue manifest in Chalmers' argument. I will come back to a discussion of the extra ingredient below in § 2.4. First, however, it is useful to turn to the Conceivability Argument which helps us depict the severity of the Hard Problem with greater clarity.

2.3.3 The Conceivability Argument

The first step is to imagine a world populated by zombies. These are not the blood-sucking zombies dressed in ripped clothes that we see in movies, rather they look like regular human beings such as myself or you. They are in fact exact physical duplicates of human beings, where all physical facts that obtain in humans are also facts that obtain in zombies, such as their micro-physical constitution and the chemical and physiological processes they undergo. In other words, they are observationally indistinguishable from us. The zombie version of myself would be writing this thesis and show joy when a concept is finally well expressed, or show frustration when an idea does not seem to shape up; the zombie version of you would manifest enjoyment or boredom while reading this thesis. However, these zombies wholly lack consciousness, namely the zombie version of me would not feel the struggle or the joy in writing this thesis, and the zombie version of you would not be undergoing the phenomenology associated the enjoyment or boredom of reading it. In other words, the me-zombie or the you-zombie will look identical to myself or you from the third-person perspective, but will be lacking the first-person perspective that characterises *what it is like* for us to be ourselves.

Imagining this zombie scenario is an exercise in conceivability and the first step of the Conceivability Argument: it tells us that it is conceivable that there exists a zombie world in which physical systems such as ourselves behave the way we do in the absence of any conscious experience. From this, the proponents of the Conceivability Argument move to metaphysical possibility finally concluding that consciousness cannot be physical.
Following Chalmers (2003, p.106):

(1) It is conceivable that there be zombies (P & ~Q)

(2) If it is conceivable that there be zombies, it is metaphysically possible that there be zombies

(3) If it is metaphysically possible that there be zombies, then consciousness is non-physical

————————-

(4) Consciousness is nonphysical

Premise (1) relies on the zombie scenario to establish the conceivability of metaphysical zombies. More specifically, that the proposition that P & Q is conceivable, where P are all the physical facts and Q are phenomenal experiences, such that P does not entail Q. The first premise therefore introduces the contingency between phenomenal and physical facts, showing that the physical facts do not (a priori) entail phenomenal facts. This is the premise that Type-A physicalist would deny.

Premise (2) is based on the assumption that conceivability entails metaphysical possibility, therefore telling us that if zombies are conceivable then they must be metaphysically possible, hence that it is possible that a zombie world may coherently exist. This premise is the one that Type-B physicalist, which defend the *a posteriori* metaphysical connection between the mental and the material such as the proponent of the Phenomenal Concept Strategy we discuss in chapter 4, would deny.

Premise (3) expresses that given the metaphysical possibility of zombies, consciousness simply cannot be physical: the metaphysical possibility of zombies is not compatible with consciousness (necessarily) being a physical fact. If a physical system like myself could in principle exist in the total absence of a first-person phenomenology, then consciousness simply cannot be physical. This is such because, had consciousness been physical, once all the physical facts were in place consciousness would automatically obtain. But this is not the case, according to the proponent of the Conceivability Argument.

The premisses thus bring us to the conclusion (4) that consciousness cannot not be physical.

2.4 The Extra Ingredient

On the basis of the Hard Problem, Chalmers (1995) advises us that if we are to find an adequate explanation for consciousness we need to look for an *extra ingredient* to "inject", so to speak, into the fundamental furnishing of the world alongside all the entities that physics tells us about. Commonly, the extra ingredient we encounter is consciousness itself, such as in the case of panpsychism defended by Goff (2017) or Hassl-Mørch (2018). According to these positions, consciousness is a fundamental and ubiquitous feature in the world, such that the physical ultimates that physics posits presents a very rudimentary form of consciousness. From this they hope to overcome the Conceivability Argument on the grounds that the existence of this rudimentary consciousness at the fundamental level is responsible for yielding that consciousness we enjoy here at the macro-level. This should answer why there is consciousness at all in our reality. Goff (2017), in defending panpsychism, explains that adopting this sort of strategy has the potential to give an account of consciousness and its relation to the physical world in a way that is consistent with the widely held thesis that the physical world is causally closed, while at the same time avoiding the difficulties encountered by reduction. In this sense, the addition of the extra ingredient should adequately account for consciousness while simultaneously safeguarding the causal closure of the physical.

Others prefer a less extreme approach. For instance, they posit that physical reality at the fundamental level is paired with protoconsciousness (Chalmers, 2013; also Montero, 2015 argues for something much alike protoconsciousness defending a form of Russellian physicalism), that is a set of properties that are not properly consciousness-properties, as in the case of panpsychists, but which are conducive to consciousness. Similarly, relying on his ignorance thesis outlined in the introduction, Stoljar (2019) argues that we should pair the physical description of the world with a set of non-standard physical properties, namely properties that are ontologically physical but which we do not know (enough) about, and which differ in a relevant sense to the ones we know at this time in that they are consciousness-relevant. This way Stoljar (2019) does not risk impeding on causal closure and the truth of physics, yet adds something extra to the fundamental furnishing of the world that supposedly would help us explain consciousness in a way that overcomes the Conceivability Argument (that is if we knew all the physical facts about fundamental physical reality we would see how consciousness arises). Stoljar's (2019) proposal is particularly elegant and parsimonious for it

avoids introducing new types of properties at the fundamental level of reality, keeping our ontology lighter than the other types of views.

Of course this is a very cursory overview of the positions, where the details of these views can be spelled out in many different ways, but for our purposes we only need to note that they try to leave the physical story intact with reference to fundamental reality, while imbuing the world with consciousness or something that is consciousness-relevant or otherwise conducive to it. This, I believe, marks the acceptance of the Hard Problem on the part of many philosophers, but who nonetheless want to leave the scientific story just as it is. Most of these views, however, potentially fall into bespoke versions of the Conceivability Argument, as we will see in chapter 6 in more details. To all we can ask: how does the consciousness we find on the macro-level yield the consciousness we are so familiar with?

2.5 A Final Note

The views outlined above attempt alternative strategies to explain consciousness based on adding an extra ingredient. However, I want to note that all of them start from the assumption that the mental and the material are ontological categories and must therefore be found in our ontology, in line with how we define these concepts. They then proceed by a *re-organisation* of these elements into (hopefully) more coherent ontologies. A similar consideration can be made with reference to the traditional neutral monist who, albeit positing neutral reality, ultimately finds mental and material properties in her metaphysics at a less fundamental level, as we will see in the next chapter. This issue is precisely what the perspectival neutral monist tackles when she posits a neutral reality and introduces the idea that the terms "phenomenal" and "physical" refer to perspectives rather than bona fide properties. Therefore, this is one of the issues where the perspectival neutral monist has the potential to present a set of fresh eyes to the mind body problem: the perspectival neutral monist attempts to revise our conception of mental and material, of their nature to be more precise, and couples it with the idea that mental and material properties are in effect perspective-dependent rather than ontological categories. As such she should avoid falling into the need to recombine and restructure the phenomenal and the physical within the ontological landscape, thus also avoids having to introduce any extra ingredient into the world of physics. We will focus on this in the second part of this thesis.

Before moving onto neutral monism in the next chapter, I would like to draw attention to a further aspect of the views discussed in the second part of this chapter, from reductive physicalism onwards: that the assumption that consciousness, as enjoyed by beings such as ourselves or animals, is something that must somehow be derived from the micro-level of reality. More specifically, the issue that I want to highlight is that there is a presupposition that the fundamental level of reality is that of micro-entities, and that in general the rest of the world is built up by means of these entities combining together and giving rise to extra and dependent levels of existence. This assumption is mostly taken for granted and it imposes a bottom-up ontology, according to which the world we enjoy depends on, and can be made sense of with reference to, micro-physical facts. However, that the world is ordered in a bottom-up fashion is not something that should be taken for granted. For instance, Schaffer's Priority Monism (2010) illustrates that it may be equally efficient, if not more so, to think of the world in terms of it being ordered in a top-down fashion, where the whole may have primacy over its parts. Would a top-down reading make any difference if one embraces the assumptions of physicalism? Probably not. At any rate the relevance of this issue will become clearer later on in the thesis, in chapter 6, where I argue against assuming ontological hierarchies for the purposes of investigating consciousness specifically. In effect, one of the elements I question with fresh eyes in this thesis is whether this sort of hierarchical organisation is effective for explaining consciousness, arguing that it is not. As we will see in more details in chapter 7, leaving behind the ontological hierarchy framework affords the perspectival neutral monist a point of strength, on the grounds that it allows her to begin her analysis of consciousness by focussing on the conscious being itself, rather having to refer to priorly assumed more fundamental levels of reality. This creates a broader and more flexible space for inquiry. Moreover, rejecting a commitment to an ontological hierarchy, together with positing neutral reality and the notion of perspectives as neutral modelled on the enactive approach, allows the perspectival neutral monist to produce an explanation of the reality of consciousness and of the physical world without having to resort to any extra ingredients.

CHAPTER 3:

INTRODUCING NEUTRAL MONISM

The origin of neutral monism is usually attributed to Ernst Mach (1886), William James (1912) and Bertrand Russell (1921,1927) who devised it as a reaction to the phenomenalist and materialist views that were popular at the time. The *big three*, as Stubenberg (2018) labels them, held the desire to commensurate the gap between the mental and the material, while also wishing to avoid positing inscrutable or abstract entities, such as atoms, as the fundamental constituents of the world. The view is also traced back to a larger number of authors such as R. Avenarius, K. Pearson, R. Ardigò, CW Clifford, M. Schlick and A.J Ayer. Some (see Cook (1994) for instance) even interpret Wittgenstein as endorsing neutral monism. Hume is also quoted as a proponent (Wishon, 2021) and Spinoza's metaphysics is often mentioned as an inspiration for neutral monism, with many contemporary philosophers, such as Heil (2013), Stubenberg (2019), Westphal (2016) to name a small sample, whom overtly refer to Spinoza as inspiration.

In spite of its older origin neutral monism has only recently begun to get traction in the debate, both in light of the issues that challenge traditional monist views and as a result of a movement in the philosophy of mind prompted by Chalmers' seminal *The Conscious Mind (*1996), which encourages the philosopher to look beyond the frame of thought held by traditional monists. Neutral monism therefore, can be grouped with a set of more novel views such as contemporary panpsychism (Goff, 2017; Hassl-Mørch, 2018), panprotopsychism (Chalmers, 2013), Russellian physicalism (Montero, 2015), Russellian monism (Howell, 2015), non-standard physicalism (Stoljar, 2019) and dual-aspect monism (Atmanpsacher, 2012; Benovsky, 2018), which are becoming alternatives in the strive to produce an adequate account of the nature of consciousness.

Neutral monism locates itself in the monist camp alongside traditional forms of mental and material monism, hence directly opposing the split ontology of dualism, and is motivated by the difficulties that traditional forms of material and mental monisms encounter as they explain the mental in terms of the more fundamental physical stuff for the material monist and vice versa for the mental monist. Particularly, the hope of the neutral monist is that by positing a fundamental neutral stuff she may produce the unified ontology that any monist seeks by opposing dualism, while also avoiding the issues that traditional monists face. The notion of neutrality in effect seems, at least from the outset, to have the potential to provide a robust ontology that can adequately harmonise the mental and the material. Whether this is the case, however, hinges on how the details of the view are spelled out and specifically on how the neutral monist constructs existence of the mental and the material within her neutral world.

Neutral monism can initially be defined in the words of Russell (1921, p.5) as the view that 'the stuff of the world is neither mental nor material, but a "neutral stuff", out of which both are constructed'. While I am not committed to this precise definition of neutral monism, especially with reference to the neutral stuff being strictly neither mental nor material as we will see in §3.2, I use Russell's quotation to begin fleshing out the three pillars of the view:

(i) monism
(ii) neutrality
(iii) the construction of the mental and the material from, or more generally their relation to, the neutral base

Pillar (i) locates the view within the monist camp alongside the more traditional physicalism, idealism or phenomenalism and opposing dualism; pillar (ii) introduces the notion of neutrality as the new alternative conception of ultimate reality, and pillar (iii) indicates the neutral monist's need to explain how the mental and the material fit within a neutral monist ontology. This is because the neutral monist is generally not an eliminativist with respect to the mental and the material, rather she assumes the reality of both phenomenal experience as well as the existence of a mind-independent reality. The perspectival neutral monist aligns herself with these assumptions as to the reality of the physical and the phenomenal realm, as we will see in chapter 7.

In this chapter I unpack the definition of neutral monism by expanding the three pillars teased out of Russell's quote. In § 3.1 I focus on pillar (i), thus on the monist aspect of the view. In § 3.2 I discuss pillar (ii) to unpack the notion of a *neutral stuff* and I discuss some instances of neutral entities presented in the debate. I devote § 3.3 to pillar (iii), where I begin expanding on the notion that the mental and the material are somehow constructed, or otherwise related to, the common neutral base. The discussion of this section is particularly important for our purposes as this thesis focuses on carving a more adequate avenue for the neutral monist to explain how the mental and the material exist within a neutral ontology. In this chapter, however, I limit myself to presenting the orthodox approaches defended by neutral monists in the literature, while I present the different approach I defend later on in the thesis (a hint in chapter 4 and in depth in chapter 5 and 7). Lastly, in § 3.4 I gesture towards some of the issues that the orthodox versions of neutral monism encounter.

3.1 Pillar (i): Monism

The notion of monism, pillar (i), is one that brings up various questions connected with what it means for an ontology to be monistic. For instance, one may ask whether the monism in question is one that posits the existence of a single entity, such as existence monism (Horgan, 2012) or Schaffer's priority monism (2010). Or, rather, whether it is a form of monism regarding what *kinds* of things exist, thus admitting the existence of a plurality of entities of the same kind. Questions like these apply to all monist views and thus they are most appropriately discussed elsewhere. However, it is important to point out that all versions of neutral monism so far seem to adopt the latter version of monism, thus admitting of a plurality of neutral entities (Stubenberg, 2018). Clear examples of this can be seen in both Mach (1886) and Russell (1921, 1927) for instance, as both speak of arrangements of neutral entities as the ground for mental and material properties, but also in more contemporary versions of the view such as Coleman's (2017) and Nagel's (2012). In addition to this, it should be noted that neutral monists often admit the existence of non-basic (derivative), non-neutral entities such as mental and material properties within their monist ontology, although these are strictly derivative and can be reduced or otherwise analysed as neutral, thus producing a layered picture of reality.

For the purposes of this dissertation, it is not necessary for us, at least not at this stage, to commit to either form of monism as our reasoning works equally on either version.

3.2 Pillar (ii): Neutrality

Spelling out the notion of neutrality can be a thorny task primarily because it requires us to think beyond the mental and material dichotomy, at least as traditionally conceived, which is so deeply ingrained in our contemporary metaphysical theories of consciousness. Emphasis needs to be put on the fact that it is the dichotomy *as traditionally conceived* that the neutral monist targets, given that her purpose is not to replace the mental and the material by positing the neutral stuff. Rather, she hopes to adequately re-conceptualise the terms of the dichotomy in light of her neutral ontology. It is therefore a constructive process of reconceptualisation that drives the neutral monist to posit a neutral stuff, rather than that of eliminating or denying the existence of the dichotomy in question. To conceive of a neutral ontology is therefore an intellectual exercise which requires some creative effort on the part of the reader, for we need to envisage a "stuff" that is fairly different from those that traditional monists and dualists depict. This neutral stuff might therefore appear strange or mysterious at first, but this weirdness is worth exploring if we are to seriously investigate the nature of consciousness. I therefore ask the reader to approach the question of neutrality with an open mind.

There is currently not much systematic work done on neutral monism and therefore I rely mostly on Stubenberg's (2018) classification. He lists five different approaches to define neutrality: two spelled out in terms of intrinsic nature, another two in terms of possible constituents and the last one in terms of laws of nature. In this paper I focus on the two intrinsic nature approaches to defining neutrality:

(a) The *Neither View*: a basic entity is neutral just in case it is intrinsically neither mental nor physical

(b) The *Both View*: a basic entity is neutral just in case it is intrinsically both mental and physical

Note how the two views above define neutrality in opposite terms. *The Neither View*, embraced by Russell (1921) in the quote at the beginning of this chapter, characterises neutrality negatively by conceiving of something that presents neither mental nor material attributes intrinsically. *The Both View*, on the other hand, prompts us to positively conceive of a fundamental reality constituted of entities that are intrinsically both mental and material. Within the current context, I interpret the *Both View* as the view that mental(-ity) and material(-ality) co-exist at the ultimate level of reality as undistinguished (or indistinguishable) and unseparated (or inseparable). This is a necessary requirement if we are to operate as monists otherwise, were mental and material attributes somehow distinguishable or ontologically separable, the view would fall under the umbrella of dualism. A discussion on why this form of neutral monism is not dualism is certainly an interesting one, but not the issues we are currently concerned with.

The *Neither View* is the position that neutral monists usually expressly adopt, while the *Both View* is more often expressly adopted by philosophers defending dual-aspect monism, though ambiguously used by neutral monism (there are often blurry boundaries between the two views). I use the words "expressly" and "ambiguously" because, as Stubenberg (2018) himself notes, neutral monists may fluctuate between the application of one type of classification and the other within the same text, expressly adopting the *Neither* definition, but relying on a covert use of the *Both View* in the application. He provides Russell (1921, 1927) and Nagel (2012) as examples. I feel that this lack of clarity can be justified by the fact that we are still in the early days of a serious exploration of the view and thus still investigating what neutral reality might, or should, ultimately be like.

More importantly, however, I believe that it is precisely the two opposing definitions *taken together* that can help us truly grasp the notion of neutrality. This is because the neutral monist assumes a monist ontology that does not favour either mental or material entities as being fundamental, or more fundamental than the other, hence avoiding having to set up the metaphysical hierarchy between the mental and the material that characterises traditional mental and material monism. Instead, by introducing a fundamental neutral stuff our monist proposes an ontology whereby neither mental nor material properties can be said to arise from, or be reduced to, the other. By introducing a neutral stuff, whether positively or negatively defined, the neutral monist thus aims to directly counteract the traditional monist approaches by

denying the sole fundamentality of one type of property over the other (as well as their distinction as per dualists), thus building on the intuition that the mental and the material must equally arise or otherwise be derived from the same ultimate neutral reality, such that neither is ontologically prior to the other. In a way, therefore, the neutral monist can be said to side with the dualist in supporting the claim that there is no hierarchy between the mental and the material; yet, without sharing the same ontological landscape as the dualist, she attempts to paint a picture of reality that features the strong continuity between the mental and the material that the dualist struggles to achieve. If successful, the notion of neutrality could afford us an explanation of the existence and relationship between mental and material properties that avoids having to reduce one to the other, as in traditional forms of reductive monism, or eliminating one in favour of the other as per eliminativist monists.

With this in mind, we can understand how the two views really help us understand the notion of neutrality taken conjunctively, namely by pushing us to think beyond the bounds traditionally set in the metaphysics of consciousness. Specifically, we can take the *Neither View* as an invitation on the part of the neutral monist to conceive of a stuff of the world that is essentially different from what we usually imagine, thus where the terms "mental" and "material" do no explanatory job in capturing fundamental reality inherently. This guides us to leave behind our traditional conception of what the stuff of the world may be like and presents the vacuum produced by the negative explanation. On the other hand, the *Both View* targets the difficulty of having to extract the material and the mental realms from a fundamental reality that is inherently neither, by inviting us to consider that the neutral stuff is somehow inclusive of - or otherwise conducive to- materiality and mentality in its nature, specifically by being mentality *and* materiality. Both the *Neither View* and the *Both View* tell us that the stuff of the world is not captured by any term of the mental and material dichotomy, while also holding the potential to produce both. Together, then, the two definitions can allow us to conceive the neutral stuff as a stuff that is in nature completely different from how it was previously imagined, yet wherefrom the mental and the material can be teased out. Nagel (2012) uses the terms *transphysical* and *transmental* to highlight the potential of this fundamental stuff that is neither mental nor material (neither is present) yet gives rise to our experience and the physical world (both are present) (Nagel 2012, p. 57). Similarly, Russell (1921, pp. 10-11) captures this idea when he writes that the stuff of the world

> is in my belief, neither mind nor matter, but something more *primitive [my emphasis]* than either. Both mind and matter seem to be composite, and the stuff of which they are compounded lies in a sense between the two, in a sense above them both, like a common ancestor.

Neutrality, as such, at this stage is therefore best understood as a concept conceived functionally to serve as the basis for a working hypothesis that, as mentioned in the introduction, is currently in the early stages of a more serious exploration. From this, I see the *Neither View* and the *Both View* as equivalent for thinking about neutrality, especially for the purposes of this dissertation and at this stage of the development of the view, as each merely intends to capture the idea that ontologically speaking the stuff of the world is not as traditional monists (or dualists) have conceived of it. For the purposes of this chapter, however, I run the discussion with reference to the *The Neither View* following the majority position expressed in the literature (Russell (1927), Nagel (2012), Coleman (2017), Silberstein (2018), etc.).

3.2.1 Notable Candidates for Neutral Entities

Neutral entities are surrounded by a veil of mystery. This might be taken to be a pitfall of the view, but also its power if properly understood. We will come back to the potential value of this mystery time and again in this thesis, especially as we argue for a neutral ontology in chapter 5. In this section, however, we look at various attempts by philosophers to flesh out neutral entities. It must be noted from the start that I neither commit to, nor agree with, these approaches, but feel it is useful to introduce them in order to begin exercising our capacity to conceive of the neural stuff, which is so different from the mental and material stuffs that we are so used to invoking in our ontologies.

Historical proponents of neutral monism such as Ernst Mach (1886) and Bertrand Russell (1927) devote considerable attention to describing what neutral entities are like in their respective views and construe them as *elements*, *pure experience* and *percepts* respectively. In what follows, I give brief definitions of these entities that do no justice to the sophisticated theories that back them. The survey nonetheless enables us to better grasp the historical course

and scope of neutral monism as well as Coleman's (2017) contemporary approach, which I unpack in more details at the end of in this section.

Mach (1886) is guided by the desire to develop an *economy of thought* that would allow him to conduct inquiry in physics and psychology within a coherent and unified framework (Stubenberg, 2018). His *elements,* that he also often refers to as *sensations*, are essentially entities much akin to those we find in our sensory realm such as colour, shape, light and sounds, but also comprise features such as pressures, intensities, times and spaces that we find in physical theory and which are observable through the senses. According to Mach (1886, p.6) 'things, body, matter are nothing apart from the combination of these elements' and that it is in fact these sensations / elements that make up bodies, rather than bodies making up sensations (Wishon, 2021). Elements therefore become constitutive of experience as much as constitutive of the structures we find in physics, thus of the physical world, and this is what makes them properly neutral. Banks (2004, p.41) clarifies that 'an element was a sensation when it varied with the human nervous system, but when it occurred in a physical variation independent of the human sensory apparatus it was called a "physical object".' We will look at this in more details in §3.3. For the time being, however, this should clarify why elements are supposed to be neutral: they are inherently neutral as they can equally give rise to mental and material properties, thus they obtain mental and physical characters according to their relations and interactions with other neutral elements.

James (1912), as opposed to Russell and Mach, is not motivated by a desire to bring together the world of science and that of perceptions, but his concerns lay mainly within the psychological field. Specifically, he is concerned with the mistaken distinction between the object perceived and the perception of the object, thus the often supposed real distinction between the phenomenon and the experience (Stubenberg, 2018). He defines his neutral stuff as *pure experience*, where pure experience is something that is 'only virtually or potentially either object or subject as yet (…) it is plain, unqualified actuality, or existence, a simple that' (James 1912, from Stubenberg 2018). Once again, pure experience is neutral because it obtains as both subject and object, thus experience and the experiencer.

Lastly, Russell's (1921) *percepts* are combinations of those qualities we find in sensory perception and the relations they bear to one another (Wishon, 2021, p.140). Russell (1927) is

driven by the desire to combine psychology and physics, thus he aims to tie together what we know through perception with what we know about matter and its structure from science. Russell (1921, p.142) writes that 'the sensation that we have when we see a patch of colour simply is that patch of colour, an actual constituent of the physical world, and part of what physics is concerned with'. That is, percepts are neutral entities because, although we find them in perception, they are also constituents of the world we discover through physics.

In Russell's quote we can see clearly how a naive realist approach motivates his desire to posit qualities as his neutral entities, which is in turn justified by his desire not to posit inscrutables such as those that science finds and rather to adopt known entities on grounds of parsimony, as he explores in depth in the *Analysis of Matter* (1927). The same can be said of Mach's reasoning, who thought that the scientist should avoid positing inscrutable entities such as atoms. Mach and Russell were both were deeply involved in the scientific debate and were aware of the dangers that could arise from taking scientific theory as the basis for ontology-making. To preserve the scientific project while harmonising the existence of mind with it, was thus a strong motivating force for them as they attempted to reconsider the scope of science and its relationship with the ontology of mind (in this sense our project here greatly resonates with theirs, as we will see in chapter 5). This, on the other hand, contrasts with James' more psychology-based attitude where finding a middle route in the fashion of his radical empiricism (1912), in this case between the mental and the material understood in terms of the subject/object divide, was a stronger motivation than that of making sense of the mind within a physical science framework.

Besides these slightly differing motivations, these three instances of neutral entities all seem to have much in common: they are nothing other than very basic and simplified forms of those qualities we find in experience. As such, sensations, percepts and pure experience can be read as paradigmatically mental. Stubenberg (2018) justifies that this is due to the empiricist tendencies of these philosophers. The reason why they treat elements, percepts and pure experience as neutral, rather than paradigmatically mental, however, is because they find these lend themselves to being captured by objective scientific enquiry as much as being constitutive of subjective experience. The *big three* often attempt to explain that these entities are not inherently mental, as they exist in the world independent of experiencing subjects. They exist as

unexperienced, that is as not had by any experiencing subject. Experiences without a mind, so to say. And this is what makes them truly neutral according to them: that they can equally constitute experience and the physical world depending on their relations and interactions with other neutral entities. How this is supposed to work will become clear in § 3.3.1. No matter how much the *big three* have attempted to show that their neutral entities are truly neutral, rather than mental, the critic has not been convinced, meaning that traditional neutral monists have been challenged with the challenge that their view may be collapsed into a form of mental monism, which we look at below in § 3.4.1.

Contemporary neutral monists (Silberstein (2017), Banks (2010), Coleman (2017), etc..) support and revise this conception of neutral entities. I focus on Coleman's position here. Coleman (2017) brings the approach of the forefathers of neutral monism into the contemporary literature by adopting unexperienced qualities- that is redness, sweetness, loudness- as they exist prior to being experienced as his neutral entities. Of course, whether such qualities may truly exist in the absence of an experiencing subject is a question that attracts its own debate and it is not my purpose to adjudicate this dispute here. Instead, I simply ask you to follow this reasoning in order to see what neutral entities could potentially be like. We can take this as an exercise in stepping into a fresh eyed perspective. In effect, Coleman (2017, p.10) himself notes that qualities may 'ultimately offer us only a conception of the *kind* of properties we seek' and that they 'demonstrate that there *could be* such natures.' It is in this sense that Coleman's qualities should be taken as a model rather than theory of neutral entities, at least here in this thesis.

Coleman (2017) suggests that qualities are good candidates for filling the role of neutral entities because they resist classification as either inherently phenomenal or physical. Let us follow him by taking colours as an example. On the common sense view colour is most often depicted as an objective and mind-independent property existing out there in the physical world. From this perspective, the rose is red such that red is a property of the rose in the sense that the colour red permeates the rose; the rose is red and remains red whether or not someone is looking at it. In other words, on the common sense view of colour for colour to exist in the world there is no requirement that a mind is cast upon it, colours are objective features of the physical world, namely physical properties. On the other hand, many philosophers construe

colour as a mental property, that is a property whose existence depends on a subject's experience. The claim is rooted in the fact that certain colour experiences can be had in dreams or hallucinations, whereby colours are effectively fabricated in the mind of a subject and in the absence of an external stimulus. For this (extensive) group of philosophers, this fact is evidence that qualities like colours exist in the mind and characterise our experience, as opposed to being something that permeates the object. On the philosophical account, therefore, colour is a mental property and there would be no colour in the absence of experiencers such as ourselves. For Coleman it is this ambiguity in defining qualities, such as colour, that makes them the perfect candidates for filling the role of neutral entities.

In effect, I feel it is this very ambiguity that enables qualities like colours to satisfy the criteria of a neutral entity as expressed on the *Neither View* as much as on the *Both View*. The ambiguity of colour suggests that colours can be understood as intrinsically *neither* mental nor material in and of themselves, while at the same time also being the type of entity that can equally give rise to *both* material and mental property. Similarly, the ambiguity of colour may suggest that we construe colour as being intrinsically *both* mental and physical, such that those mental and material properties that we find in our day-to-day life may equally be constructed out of it. The issue with such *qualities*, however, is that their existence is too intimately tied to our experience and therefore to the realm of the mental, as briefly mentioned above. They are paradigmatically mental entities, may they be "out there" in the world or "up here" in the mind. In my opinion, this intimacy with human experience de facto makes qualities inappropriate candidate for neutral entities, as we will see later in § 3.4.1 where we discuss the challenge that neutral monism can be absorbed into a form of mental monism. In this sense Coleman's qualities make no real improvement on those that the *big three* posited. Even though Coleman hopes avoid this criticism by stressing that qualities are to be construed as non-experienced, this may not be the most efficient strategy for designing some intelligible and credible neutral entities, possibly hiding rather than enhancing the potential of neutral monism.

It is useful to note that other recent attempts at conceptualising the neutral stuff have been made by Seager (2019), Pereira et al (2018) and also certain versions of IIT capture the ethos of neutral monism and can thus be seen as providing a possible route for fleshing out what the neutral stuff may be like (Mindt, 2019).

The model of neutral entities discussed should have been useful for us to illustrate how neutral entities are most often thought of, but most importantly to exercise our creative skill in imagining a world that is so different to the one we usually imagine. I thus want to remind the reader that I am not committed to this model and, thus, I ask the reader to leave it aside and let the neutral stuff remain as mysterious as possible for now, with the only *proviso* that it is neither mental nor material and simultaneously both at the same time.

3.3 Pillar (iii): Constructing the Mental and the Material

The issue of constructing the mental and the material from the neutral stuff begins with asking the question as to how the mental and the material may arise or otherwise relate to the common neutral ancestor. There are two avenues for the neutral monist to explain the existence of the mental and the material and their relation to the neutral base: the first is in terms of *arrangements* of neutral entities and the second in terms of *ways of thinking* about the neutral stuff. In this thesis I develop a framework for a third approach, according to which "mental" and "material" are perspectives, which differs from both prior approaches and yet bears more resemblances in structure to the latter, as we will see briefly below and in detail in chapter 7 and in the conclusion. Here, I focus on the prior approaches to illustrate the various ways neutral monism can be built, the issues associated with them and, in turn, understand the potential of the perspectival neutral monism view I develop in this thesis.

3.3.1 Arrangements of Neutral Entities

> The whole duality to mind and matter (…) is a mistake; there is only one kind of stuff out of which the world is made, and this stuff is called mental in one arrangement, physical in the other (…) therefore the things commonly regarded as mental and the things commonly regarded as physical do not differ in respect to any intrinsic property possessed by the one set and not by the other, but differ only in respect

> of arrangement and contexts (Tully, 2005, p.211, quoting Russell's Collected Papers, 1913, pp. 7-15).

Similarly, Mach (1886, p.16) tells us that it is only depending on their functional relations of dependence that they [neutral entities] are sensations or physical objects. The picture we have is, hence, that of a world made up of intrinsically neutral entities that are arranged or related in differing ways, and it is the differing ways neutral entities are arranged or related to one another that gives rise to either mental or material properties. I will refer to this as the *arrangement approach*. The approach was pioneered by Russell (1927) and Mach (1886), and to a certain extent also by James (1912), and echoed in contemporary literature (Coleman, 2017; Banks, 2010; Nagel 2012; and Silberstein 2017 are good examples).

It follows that in the arrangement approach we can discover the existence of mental and material properties by investigating the organisation of neutral entities directly at the fundamental level of reality. There are two main elements at play here: a set of neutral entities plus the *relations* that hold between them. There will be some number of neutral entities that are "mentally" related and which give rise to material properties, and some other groups of neutral entities that are "physically" related and thus give rise to physical properties. I want to emphasise that individual neutral entities might, and indeed should, be part of arrangements of both mental and material properties. This very well captures the elegant essence of neutral monism whereby reality is made of entities that are fundamentally intrinsically neutral, yet give rise to a reality that can be both physical and mental depending on how the same individual entities are organised. Coleman (2017, p. 9) puts it clearly as he explains that intrinsically neutral entities are 'capable of producing mental and physical properties through their interactions'; then explaining that neutral entities such as colours

> are mental when they play mind-related roles, like being present in sensation, being remembered, or being dreamt, and physical when they play physical roles – like obeying the equations of physics. But these are just roles, and colours, as neutral properties, are no more defined by them than an actress is defined by the roles she takes on (Coleman, 2017, p. 10)

For instance, the neutral entity embodied by the colour quality relates physically to another neutral entity in so far as it is a light-wave bouncing off a surface, hence realising the material properties of magnitude and amplitude. On the other hand, the same colour quality would relate phenomenally to another neutral entity when it interacts as colour experience, for example, producing a colour sensation such as the redness we experience when we see a red rose, thus realising a mental property. Banks (2014) also defends the arrangement approach and holds that mental and material properties simply correspond to functional relations had by the arrangement of neutral entities in question. For Banks (2014) mental and material properties are thus individuated by what they do. It follows that the difference between mental and material properties which we find in our daily life would boil down to a difference in the behaviour of the relevant neutral entities. Colman (2018) interprets Nagel (2012) as having a similar view.

On the arrangement view, therefore, phenomenal and physical properties are real because they are just two ways the neutral stuff is organised. This in turn produces a layered picture of reality, where mental and material properties depend on and are reducible to their common base (Silberstein, 2017). It follows that the neutral monist adopts a sort of *reductionist* framework that is not dissimilar from her physicalist adversary. This is a result that Silberstein (2017) happily embraces, but also one which is also seldom, if ever, discussed by authors adopting the approach.

3.3.2 Ways of Thinking about Neutral Reality

Another possible approach to explain the existence of mental and material properties in a neutral world is based on the different ways we can think about neutral reality. This means that reality is intrinsically neutral, but that mental and material properties depend on how the neutral stuff is considered by a subject. I call this the *ways of thinking* approach. In this case, as opposed to the arrangements approach, looking at the fundamental level of reality will not reveal the existence of mental and material properties *per se*, rather it is the way we investigate neutral reality that allows us to discern mental or material properties. As such it is the direction of investigation that changes, rather than the subject matter (Mach, 1886, p.17).

This approach is the least explored by neutral monists, although this kind of strategy finds echoes in Davidson's (1970) anomalous monism, in the works of the *a posteriori* physicalist (we will look at this in more details in the next chapter) and in dual-aspect monism (Benovsky, 2018, is a good example). A few words need to be spent on dual aspect monism because the view shares blurry boundaries with the type of neutral monism built on the *ways of thinking* strategy, and can thus help us clarify the approach. Dual-aspect monism is the view that the mental and the material are two 'basic and irreducible *aspects* [my emphasis]' (Atmanspacher, 2012) of a fundamental reality which 'manifests itself generally and equally under both aspects, tied together inseparably' (Atmanspacher, 2012). This should produce a uniform reality wherein the split between metal and material domains is not *a priori* but occurs on the epistemic level. Dual-aspect monism is strikingly similar to the *ways of thinking* approach as both defend the idea that the distinction between the mental and the material exists only on the epistemic level. Neutral monists who adopt the *ways of thinking* approach, such as Heil (2013), however, expressly deny the idea that the mental and the material are *aspects* of the stuff of the world, a basic postulate of the dual- aspect view. The purpose of the *ways of thinking* approach is to deprive the ontological level of any categorical distinction between the mental and the material, while at the same time to justify why we talk about mental and material properties and why such talk is true, and this is why they reject the idea that the neutral stuff presents two irreducible aspects. This, moreover, contrasts with the *arrangement* approach which ultimately introduces some sort of ontological, albeit non-fundamental, distinction between the mental and the material based on the different ways the entities are related.

Heil's *powerful qualities* view (2010, 2013), which he develops as a form of neutral monism, is a great example of this approach. Heil (2013) argues that the fact that we can characterise the world qualitatively (in terms of sensations and according to experience) or dispositionally (in terms of physical structures and according to science) does not entail that these pick out (metaphysically) different properties. Heil's idea is that properties have a dual nature, simultaneously qualitative and dispositional, where 'the property contributes in a distinct way to the qualities and dispositionalities of the object possessing it (Heil, 2013, p.214). He brings the example of a vase and asks us to consider the structure of the vase and its fragility. He continues by pointing out that the fragility of the vase is not inherently separate or distinct from its structure. We can nonetheless distinguish the vase's structure and its fragility by a process of

abstraction. Heil (2013) explains abstraction with reference to Locke's (1690) notion of *partial consideration*, whereby when we talk about the vase's structure or about its fragility we are ultimately referring to the same property, albeit differently or partially considered.

Perspectival neutral monism, the view I develop in this thesis, is akin to the ways of thinking approach on the grounds that it places agents at the root of the distinction between the mental and the material: mental and material properties are *conceptually distinct*, yet no distinction can be found in the ontological realm. The *ways of thinking* approach also resembles my perspectival neutral monism in that it does not posit any sort hierarchical structure in its explanation, thus it avoids introducing mental and material properties which depend on the more fundamental neutral base as per the *arrangements* approach. The *ways of thinking* approach, however, lays much emphasis on the relation between the existence and distinction of mental and material properties and the inherently mental act of thinking or abstracting. That is, the mental process of thinking or abstracting is responsible for teasing out mental and material properties from the underlying neutral reality. However, thinking and abstracting are paradigmatically mental processes, and as such the view could be criticised on the grounds that it may be nothing other than mental monism in disguise. My perspectival neutral monism approach, on the other hand, differs greatly from this in that I introduce a notion of perspectives as truly neutral, rather than mental, which thus starkly contrasts with Heil's process of partial consideration and mental abstraction. Very briefly, the perspectival neutral monist develops an account of perspectives inspired by the notion of sense-making that is central to the enactivist approach to cognition. Sense-making for the enactivist is cashed out as a form of rudimentary cognition on the part of the organism, where this is defined by a number of processes that include, but are not limited to, chemical, physiological, metabolic processes as subsumed under the unity given to an organism by means of the boundary and, moreover, where these processes are not seen as passive mechanical but, rather, as involving an active sort of intelligent adaptive reaction on the part of the organism that actively confer meaning onto their environment (a rudimentary form of cognition). From this model, I develop the idea that "occupying a perspective" is a state that involves the organism as a whole, namely all those processes and states that we can capture by mean of a phenomenal or physical vocabulary, but that are ultimately neutral. These neutral perspectives are in turn responsible for the obtaining of both consciousness and the physical world, by means of a process we call the *narrowing-down* of the

neutral stuff. The notion of perspectives as neutral is thus the main point of rupture between my view and Heil's (2013) position, and has the potential to rescue the perspectival neutral monism from the mental monism challenge I outline below.

3.4 Some Problems for Neutral Monism

Neutral monism is usually challenged on two main grounds:

(i) that the view can be collapsed into a form of mental monism
(ii) the problem of understanding how experience (or the physical world) may arise out of the fundamental (neutral stuff)

The first is specific to neutral monist whereas the second can be seen as a version of the Hard Problem for the neutral monist.

3.4.1 The Threat of Mental Monism

One common criticism against neutral monism is that the view can be collapsed into some more or less extreme form of mental monism, namely idealism or phenomenalism and, in more recent times, also panpsychism. (Feigl, 1975; Chalmers, 1996; Tully 2005; Stubenberg, 2018). The criticism targets the explanatory potential of the notion of neutrality and is traditionally directed at the fathers of neutral monism and those who support the *arrangement* approach, thus generally posit neutral entities such as elements, pure experiences and percepts or otherwise qualities such as Coleman's (2017). The issue with these, the critic contends, is that they are paradigmatically mental entities and therefore the notion of a neutral stuff does no explanatory work. In addressing Russell's view, Feigl (1975) writes that 'the data upon which the construction [of the mental and the material][4] is based turn out to be items of immediate experience (sentience) and thus mental after all'. If neutral stuff is made up of entities that belong to the realm of the mental, then what is the reason one should posit a mysterious neutral stuff, rather than the better known mental quality? In other words, what is the explanatory

[4] I added the square parenthesis for clarity

strength of introducing an obscure neutral stuff when the explanatory job is done by mental entities? The explanatory work is done by the sensation or the quality, thus by the paradigmatically mental entity, and thus the neutral stuff is better collapsed into a version of mental monism on the grounds that the latter is more elegant and parsimonious. The neutral monist will answer that the qualities she posits are prior to experience and as such are not intrinsically mental (as traditionally understood) given they can be characterised in both mental and material terms. It must be noted that Mach (1886), James (1912) and Russell (1927) explained at length how the neutral entities they posited are not mental but neutral. Recall the example of a colour in Coleman. However, this has not served to convince the critic yet, hence, as Stubenberg (2018) suggests, whether this is the case is to be adjudged on a case by case basis. Nonetheless, I do think that while there is much to be appreciated from the reasoning behind positing neutral entities such as qualities, the confusion and opposing reaction that the terminology itself raises is sufficient for the philosopher who believes in the potential of a neutral ontology to look for alternatives.

I believe that the threat of mental monism can also be extended to the *ways of thinking* approach. In fact, if mental and material properties are the result of a process of abstraction on the part of the observer/ experiencer, then isn't the material dependent on the mental process of the observer? In other words, the challenge is the following: the mental monist posits that material properties depend on the mental. The critic would thus challenge the neutral monist that her view runs into the same conclusion on the grounds that it is the way the agent looks at the world, and therefore the mental fact, that is ultimately responsible for the existence of material properties. It would follow that the mental must either be fundamental or at least more fundamental than the material for the neutral monist, just like for her mental monist adversary, and therefore that the physical realm is dependent on the mental, again just like for the mental monist. If the argument is as correct as it seems with this cursory exposition, once again the notion of neutrality would do no explanatory work, it would merely crowd our ontological landscape with unwanted obscure entities and, as such, it is best discarded.

3.4.2 The Problem of Experience

The Problem of Experience is the problem of understanding how the rich qualitative experience that characterises our life as conscious agents can be born out of some neutral stuff. The critic can ask: how does the combination of the neutral entities ultimately yield experiencing subjects such as ourselves? This problem is analogous to the Hard Problem for the physicalist (which we looked at in §2.3.2) and it can be loosely stated as the problem of understanding how experience arises from the combination of non-experiential neutral ultimates.

The problem of experience for the neutral monist also resembles the combination problem for the panpsychist, that is the problem of understanding how the combination of a number of sentient physical ultimates may yield macro-experiencing subjects (Seager, 2010; Chalmers, 2013; Goff, 2017). For the neutral monist the combination problem could translate as the question of understanding how neutral ultimates such as qualities or pure experience, that are not had by any subject, may yield experiencing subjects. More specifically, the issue revolves around the fact that even if the fundamental level of reality is populated by the same stuff that makes up our experience, it is not clear how this can be conducive to an experiencing subject with a unified consciousness (James, 1890; Goff, 2009). This is because experience as we know it is not simply the conjunction of various qualities but it involves a subject, hence some awareness of the experiences she is having such that there is *something it is like* for her to undergo the experience. Therefore, we can rightly ask: how does a set of non-experienced fundamental qualities become experience?

3.4.3 Final Remarks

One may say that the threat of mental monism and the problem of experience are based on wrongful assumptions on the part of the critic. Were the critic sufficiently open minded, or maybe just charitable, she may be prone to overcome the issues that stem from the terminology in the case of percepts, elements and pure experience and, rather, understand how these entities truly differ from the mental entities that a mental monist posits as the fabric of her universe. Mach (1886) expressly recognises this close tie may be misleading, and asks the reader to overcome 'our stereotyped thinking' (from Stubenberg, 2018) so as to really capture

the significance of the elements a neutral monist like him may posit. This has not been effective, however, thus resulting in a major challenge for the traditional neutral monist, namely the mental monism challenge explored in § 3.4.1 above. Similarly, I feel the charitable critic could in principle see how a series of non-experienced qualities could come together to form an experiencing subject if paired with the relevant physical theories about subject constitution, for instance. Sure, this may be a stretch on the part of the critic, but I believe that it is intellectually possible for the broader project of tackling the question of consciousness.

Nonetheless, the potential of a view cannot be assessed this way and any confusion or point of opposition that arises must serve us as an eye-opener and a drive to look further, to explore further, to understand further, to further refine our theories. In the rest of the thesis, therefore, I embrace the above criticism and begin developing a view that takes on the difficulties the critic challenges my predecessors with. We start doing this in the next chapter.

CHAPTER 4:

PHENOMENAL AND PHYSICAL CONCEPTS, SUBJECTIVE AND OBJECTIVE PERSPECTIVES

In this chapter I explore phenomenal and physical concepts and the associated subjective and objective perspectives, which are central elements for our argument toward a neutral ontology (in chapter 5), as well as for understanding how the perspectival neutral monist can make sense of our true talk of experience and the physical world (in chapter 7). I thus begin introducing some of the building blocks that make up the perspectival neutral monist position. Specifically, in this chapter, I discuss that the distinction between the phenomenal and the physical is a conceptual one, which depends on the two distinct perspectives that an agent can occupy with respect to neutral reality, one subjective and inward-directed for the phenomenal and the other objective and outward-directed for physical, and that this distinction is not necessarily reflected on the ontological level.

A reasoning on the conceptual rather than metaphysical distinction of the mental and the material underlies all version of neutral monism. It can be found in the philosophies of Mach (1886), James (1912) and Russell (1927), for instance, where all agree that there is no real distinction between the mental and the material on the ontological (fundamental) level; but we can also find it in the more contemporary approaches such as Coleman's (2017) and Heil's (2013). I therefore align myself with fellow neutral monists with respect to the conceptual distinction of the phenomenal and the physical. My view, however, differs substantially from theirs as I introduce the notion of neutral perspectives based on the enactive approach to reason on the distinction between the mental and the material.

In this chapter I take up my own exploration of the conceptual distinction of the mental and the material. I expound a popular existing strategy, the Type-B physicalist's *Phenomenal Concept Strategy*, and propose to apply it more widely by extending it to physical concepts. This is, in effect, an application of *my fresh eyes* methodology. In a nutshell, the Phenomenal Concept Strategy (Stoljar's, 2005, terminology) tells us that phenomenal properties appear distinct from the fundamental physical reality, on the grounds that the concepts we use to capture phenomenal states have a *special* nature. That is, possessing phenomenal concepts is different than possessing other types of concepts, and this makes it look as if the properties these concepts pick out are distinct from the properties that physical concepts pick out. The defender of the Phenomenal Concept Strategy then argues that we cannot draw a conclusion about the ontological distinctness of the phenomenal from the physical simply because we use distinct (*special*) phenomenal concepts for it.

Following the moves of the strategy I argue that physical concepts also have a peculiar nature, which makes them *special* in their own right. From this, I then explain how the two special sets of concepts correspond to the two distinct perspectives that an agent can occupy with respect to reality and that the distinction between these concepts seems to depend on the availability of two distinct perspectives that an agent can occupy. My reasoning is analogous to the Phenomenal Concept Strategy in finding that concepts have a special nature, specifically expanding the strategy to physical concepts, and also in that we cannot draw any conclusion about the ontological distinctness of the phenomenal and the physical realms from the possession of these equally special concepts. The purpose of this chapter, therefore, is not to argue that phenomenal and physical concepts refer to the same property (I will focus on this extensively in the next chapter, as I move toward positing a neutral reality). Rather in this chapter I aim to build upon an existing view by applying my fresh eyes methodology to expand it, in order to open up the possibility that we have two systems of concepts with their own peculiar 'special' logic, that each conceptual scheme captures reality from a specific subjective/ objective perspective that an agent can occupy toward reality, yet that this conceptual distinction does not entail an ontological distinction.

Bear in mind that at this stage I am not yet committed to any specific ontology, I have not yet argued for a neutral ontology to be specific, thus I will bind my analysis to concepts without

talk of properties. Moreover, we have not yet developed our notion of perspectives as neutral (rather than mental) states yet, and thus that the discussion here is only meant to introduce certain elements that we then build on in the rest of the thesis.

The chapter is organised as follows: in section 4.1 I introduce the Phenomenal Concepts Strategy to understand what a phenomenal concept is and how it exists in an ontology that is fundamentally physical. In section 4.2 I re-trace the strategy and expand the reasoning to the special nature of physical concepts, I discuss how the special nature of phenomenal and physical concepts corresponds to their being perspective-dependent and that these concepts seem to pick out properties in the world *as encountered* from a specific perspective. In section 4.3 I introduce the idea that perspectives are inherently limited.

4.1 The Phenomenal Concept Strategy

The Phenomenal Concept Strategy revolves around the idea that phenomenal and physical concepts are distinct, although both ultimately refer to the same physical properties. Sundström (2001) summarises this idea by making an analogy with his conception of himself. He writes:

> this difference in ways of thinking can deceive us into supposing that the thinkings must concern different and dissociable things. But there is no reason to so suppose, because, familiarly, different ways of thinking might concern the same thing; as when I think about myself in an inner way, as I, and also think about (what turns out to be) myself as the shopper who set off the alarm (Sundström, 2001, p.268).

To translate this to an example within the phenomenal/ physical vocabulary, the phenomenal concept of pain is distinct from the physical concept of C-fibers firing, although both ultimately refer to the same physical property, namely C-fibers firing. So, what produces this putatively wrongful thinking about C-fibers firing and pain as different and dissociable things, which then leads to the Explanatory Gap?

From the standpoint of the Phenomenal Concept Strategy the Explanatory Gap is a result of phenomenal concepts having a certain *special* nature. Defenders of the strategy develop different approaches to flesh out what is special about the nature of phenomenal concepts. Carruthers and Veillet (2007) helpfully summarise the three main proposals:

1) Phenomenal concepts as recognitional concepts of experience (Loar, 1990; Carruthers, 2000; Tye, 2000)
2) Phenomenal concepts as types of indexicals (Perry, 2001; O'Dea, 2002)
3) Phenomenal concepts as quotational concepts (Papineau, 2002)

All the accounts above tell us something slightly different about what phenomenal concepts are and how they are employed, i.e. what makes them *special*. Each one has been criticised for a different reason, although this need not concern us here. For our purposes I follow the methodology employed by Stoljar (2005) and Chalmers (2007), who focus on what all these instances have in common.

First, all phenomenal concepts refer to experience where experience is understood phenomenologically. Stoljar (2005, p. 468) writes that a phenomenal concept is a concept that picks out a very specific sensation, say a RED SENSATION. However, a RED SENSATION should not be confused with the concept RED, which merely qualifies the object in question. Neither it is the concept SENSATION THAT REPRESENTS THINGS AS RED, because it is possible that a red sensation may not represent things as being red, nor is it the concept THE SENSATION ONE GETS FROM LOOKING AT RED THINGS, for that may not produce a RED SENSATION at all but a PINK SENSATION for example. Generalising from this, a phenomenal concept is one that captures the first person perspective of an experiencing subject, namely that *there is something it is like* (Nagel, 1974) for a subject to undergo a certain phenomenal experience.

Second, the possession of a phenomenal concept implies that there is an intimate epistemic relation between the subject who holds the concept and the phenomena it picks. According to Balog (2009), for instance, possessing a phenomenal concept means that we are directly and incorrigibly acquainted with the mental state that the phenomenal concept in question picks such that, for some, this acquaintance ultimately reveals the essence of these

mental states. I must admit that not everyone, including myself, might want to agree with the claims about incorrigibility and the revelation of the essence of the property, for example in chapter 5 I argue that we should not assume that concepts tells us exhaustively or accurately about the metaphysical nature of the property they pick. However, I think that the privileged epistemic position that is implied in the possession of a phenomenal concept, and thus the direct access these afford us to our own experience, is an essential feature of what makes phenomenal concepts special. We can be wrong about a physical concept, for instance if we misunderstand its meaning or reference, precisely because we lack this direct intimate relation with the phenomena they pick; however, we can never be wrong about the acquisition and application of a phenomenal concept such as RED SENSATION in the same way, because of us being directly acquainted with the experience in question.

Stoljar (2007) grasps the essence of this idea in what he calls the experience thesis:

> *Experience Thesis*: S possesses the phenomenal concept C of experience E only if S has actually had experience E

The *Experience Thesis* captures the idea that a subject, such as myself or you, can come to possess the phenomenal concept C only upon having had the experience E that the concept C refers to. For example, if I have never had a RED EXPERIENCE, say I neither saw anything red nor I have hallucinated a red patch, then I cannot be said to have the concept RED SENSATION. I can be told about it, I can understand what it means to undergo an experience of red as opposed to yellow, but I simply cannot have the concept RED SENSATION itself. Let us contrast this with the concept TABLE by supposing I have never seen a table, but I have seen a chair. If I get a description of what a table looks like and of its function is and then compare the table to a chair (considering for instance that a table is similar to a chair but has taller legs and lacks a back and an arms rest), then I can form a tentative concept of a table even though I have never seen one. This is not the case with RED SENSATION. I simply cannot form the concept by means of deduction or as a result of being explained what it is like to undergo it. Moreover, there is no comparison with a YELLOW SENSATION or an ORANGE SENSATION that can provide me with what a RED SENSATION is like. In this sense, phenomenal concepts are unique as compared to

physical (or other types of) concepts, because they attach to a specific experience that needs to be had by a subject for her to have the concept.

In addition to this, and maybe even more importantly, I can give a description of all the physical components of having a red sensation, yet I simply cannot deduce the red sensation from knowing about all of its components, its similarities and differences with other phenomenal concepts, unless I have undergone that experience. It is in this sense that phenomenal concepts are said to be *conceptually isolated* (Carruthers and Veillet, 2007) or *inferentially isolated* (Sundström, 2001), and what Loar (1990) labels the *conceptual independence* of phenomenal concepts.

The Type-B physicalist then proceeds to explain that the Explanatory Gap between the mental and the material depends on the special nature of phenomenal concepts on the grounds that certain features of phenomenal concepts, such as their close ties to experience and their conceptual isolation from physical concepts, makes it look as if the two sets of concepts are so distinct that they must pick distinct kinds of properties. For example, take P to be the physical-functional property that is associated with the experience of pain, then 'what explains the "appearance of contingency" is that a phenomenal conception of pain and a conception of P in physical-functional terms can be cognitively independent- can have independent conceptual roles- even while introducing the same property' (Loar, 1990, p. 85). Loar thus agrees that there is an epistemic gap between the two sets of concepts, but denies that this translates to a distinction on the ontological level. It follows that this explanation of the appearance of contingency between the physical and the phenomenal shows that the connection between mental and material is *a posteriori*, and thus it does not threaten the supervenience of the phenomenal on physical for the Type-B physicalist. Much more can, and indeed should, be said for the purposes of appropriately doing justice to the strategy, however this would lead us astray here. This brief outline should nonetheless have been sufficient to highlight the central points of the strategy, which are useful for us to expand and build upon it.

4.2 The Strategy and Neutral Monism

We have already stated that the neutral monist supports conceptual dualism in conjunction with ontological monism, like the Type-B physicalist. As opposed to the Type-B physicalist the neutral monist does not endorse the view that the world is fundamentally physical and that phenomenal properties are reducible to physical ones. Accordingly, the strategy needs some tweaking to be adapted to the neutral monist's framework. In order to do this, I tackle the tacit assumption, evident in the physicalist's attitude, that only the first person perspective is special on the basis that it is subjective and thus connected to a single point of view. Contrast this with the third-person perspective that is putatively centreless, a so-called 'view from nowhere' (Nagel, 1986), and thus not generally deemed special. I challenge this idea by developing an account of the special nature of physical concepts, which reveals that an objective perspective, that grounds physical concepts, is just as special as the subjective.

The central idea is, in a nutshell, that both mental and material concepts are special because each of them captures a distinct yet equally essential feature of our existence in the world, namely either our existence as subjects of experience or our existence as causal structures within the causally closed world. Each of them is special because it is conceptually isolated from the other, to use Carruthers's (2007) terminology, or cognitively independent to use Loar's (1990). Each of them is special because it is acquired via a distinct process. Yet, both refer to the same (ontologically neutral, as we will see in chapter 5) reality.

In the next few paragraphs I focus specifically on a description of physical concepts and how they are acquired, an account that is not available in the Phenomenal Concept Strategy literature, thus extending the scope of the physicalist's strategy to physical concepts; this enables us to draw the distinction between one type of concept and the other more clearly, thus to expand on the idea that there is a multiplicity of concepts available to us and that each has its own logic and genesis.

Let us first, however, begin by indicating the features of the Type-B physicalist's account of phenomenal concepts we share. The perspectival neutral monist agrees with the Type-B physicalist that a phenomenal concept captures that there is something it is like to undergo an

experience such as a RED SENSATION, that they afford us a direct access to our inner life and that a phenomenal concept can be acquired by a subject only upon having the relevant experience. She also recognises that these features make phenomenal concepts 'special' and cognitively isolated, not only from physical concepts but also from other phenomenal concepts, such that it is not possible to acquire a phenomenal concept by means of inductive or deductive reasoning. These special features of phenomenal concepts enable us to pick out an intimately subjective aspect of reality which captures the essence of what it means to be a subject, and they are acquired and applied when the subject looks *inwards*. In addition to this, the perspectival neutral monist also recognises, with the Type-B physicalist, that phenomenal concepts are deeply distinct from physical concepts.

4.2.1 The Special Nature of Physical Concepts

In essence, physical concepts are all those concepts that refer to the way the world is "out there," irrespective of the experience of any single subject, hence regardless of how it appears or feels to me, to you, to a dolphin, or to my dog. Physical concepts are specifically acquired by turning our attention to the (so-called) mind-independent world and then distilling our knowledge of it from our experience of it. The genesis of physical concepts is thus founded on an *intellectual* rather than experiential process of discovery, understanding and characterisation of that reality that we share and which prescinds our experience of it.

The peculiarity of the physical conceptual framework is that it produces what we consider an *objective* explanation of the structure of the world. This is deemed objective because the concepts in question can be obtained, accessed and validated by third parties, as they capture phenomena that transcends the personal phenomenal state of the investigating subject. Tables and chairs, trees and stones, the brain and the cardiovascular system are all good examples of physical concepts, and so are quarks, the mass and charge of an electron, or the force of gravity, in that they are characterised by their capability to be equally obtained, accessed and verified by the public at large.

Note that the instances of physical concepts given above include some macro-physical examples as well as some micro-physical ones. This helps us uncover a further feature of

physical concepts that contributes to their special nature: the absence of an epistemic gap amongst them. Let us recall the theory-based definition of the term 'physical' given above: physical properties are all those properties that physics tells us about, or those that supervene on them with metaphysical necessity. We can then understand physical concepts as all those concepts that are based on, or pick out, physical properties such as the mass and charge of an electron, and those that adequately supervene on them with metaphysical necessity such as trees and tables, where both are necessary for us to describe the structure and mechanics of the world. The peculiarity of this conceptual system is that common-sense physical phenomena such as heat, namely that phenomena which causes mercury to expand in a thermometer and makes us sweat, is intelligibly explained with reference to the corresponding micro-level phenomena, namely a specific movement of molecules. It follows that the common-sense macro-physical and the scientific micro-physical coherently fit within the same physical conceptual scheme, without producing an epistemic gap analogous to the one the physicalist finds in her attempts to reduce the phenomenal to the physical.

Given this peculiarity of the physical conceptual scheme, in the rest of this section, I analyse the special nature of physical concepts with particular attention to the micro-physical, rather than the supervenient macro-physical. There are two further reasons why I choose to focus on physical concepts as developed by our physical theories. First, while it is true that tables and trees are physical objects, it is also true that our conception of what it is for something to be physical, as paradigmatically understood in our times, is based on the development of physics following the scientific revolution. In other words, our conception of the physicality of a table or of a tree is deeply informed by our scientific conception of what it means for an object to be physical (we hinted at this in chapter 1 and will return to it in more depth in chapter 5). This epistemic dependence prompts me to focus on micro-physical concepts as developed by physics rather than the macro-physical ones. Second, it is a less thorny task to illustrate the contrast between a micro-physical fact and a phenomenal one, as opposed to distinguishing a macro-physical fact and from a phenomenal one. This is because our conceptions of a macro-physical object are soaked with our experience of it, especially in the light of the cognitive phenomenology debate. An in depth discussion of this would be inappropriate for our purposes and most adequately done else where, but mentioning it should

emphasise why I lead my discussion with reference to the micro-physical rather than the macro-physical.

Before diving further, it is also important to note that, in analytical philosophy, talk of the physical refers to a world that is causally closed. (Papineau, 2006). That is: all properties that supervene with metaphysical necessity on the properties of physics are properties that supervene on those that describe a causally closed world. When using physical concepts, therefore, we must keep in mind that we are invoking notions that explain the mechanics of our perceivable world and which can be derived from or reduced to the causally closed physical world, thereby constituting a peculiarity of the physical conceptual framework in and of itself.

This takes us straight to the next special feature of physical concepts: the distinctive way they are acquired. Specifically, they are usually obtained within rigidly controlled settings or through carefully calibrated equipment, with the aim of producing a certain set of *reliable* data about the world. When lab equipment reaches its limits, or the scientist needs to work out information about the world that does not presuppose the rigid setting of experiments within science labs, physical concepts can in turn be acquired by means of studious induction or deduction. In this sense, therefore, physical concepts are acquired through an intellectual process of investigation that is in stark contrast to the process a subject undergoes to acquire phenomenal concepts (by means of having the very experience the concept captures and thus holding a direct relation to the phenomena in question).

Furthermore, physical concepts can be said to hang together as held by a single "logical thread". In other words, while being cognitively isolated from phenomenal concepts as described above, they are not isolated from other material concepts (at least not in principle). Think about the example of understanding what a table is from §4.1. We can grasp the concept of a table by describing its use and function, its similarities or its difference with other objects such as a chair for example. This translates into the constitution of a framework of investigation, or conceptual scheme, where concepts depend on each other for their meaning. This applies to the relation of micro-level concepts amongst themselves, macro-level concepts amongst themselves, as well as across levels, as seen above. An important entailment of the connection amongst physical concepts is that when new advances in the field of physics are made,

previously accepted concepts and theories may change or lose their value in the view of a newer and more appropriate one, therefore trumping the veridicality of the prior concept and what they were supposed to explain. It follows that physical concepts are inherently corrigible and modifiable, as opposed to the putative incorrigibility of phenomenal concepts, in that they pick out entities in the world which may successively be understood differently.

4.2.2 Phenomenal and Physical Concepts, Subjective and Objective Perspectives

The above analysis of phenomenal and physical concepts yields some valuable conclusions. First, that physical and phenomenal concepts are radically different, almost opposite in a certain sense, and isolated from one another. This is particularly important as it may help shed some light on why we might have an explanatory gap to fill, at least on a conceptual level, as suggested by the Type-B physicalist. Second, that each conceptual framework corresponds to a distinct *perspective* that an agent can occupy to make sense of her existence as experiencer-in-the-world: one subjective and inward-directed for phenomenal concepts, the other objective and outward-directed for physical concepts. In other words, when the agent occupies a subjective and inward-directed perspective she obtains and uses phenomenal concepts and when the agent occupies an objective and outward-directed perspective she acquires and employs physical concepts. For example, when I burn my finger taking a hot tray out of the oven and I feel pain it is the inward-directed and subjective perspective that allows me to capture what it is like for me to feel pain and I do this by means of phenomenal concepts. However, if I want to capture the causal-structural-functional mechanism of the physiological process I am undergoing I have to step out of the subjective perspective and step into the objective perspective, thus using physical concepts.

The fact that each conceptual scheme corresponds to a specific perspective is of crucial importance, because it seems to me that the possession and use of physical and phenomenal concepts expresses not only the property encountered (where we remain silent on the nature of the property), but also the perspective through which it is captured. Take indexicals such as “here” and “there” or “now” and “then”, these are helpful analogies that can help us better

understand this point, on the grounds that their application is perspective dependent and thus their uttering expresses the position of the subject with respect to the object in question. For instance, I can refer to spatial location L as "here" or "there" depending on the spatial position I assume with respect to L when referring to it. Thus when I say "here" I express that there is a spatial location L together with expressing that I am positioned nearby it, and when I say "there" I express that there is a spatial location L together with expressing that I am positioned far away from it. The inward-directed and the outward-directed perspectives function analogously with respect to the uniform reality, in that the acquisition and use of phenomenal and physical concepts does not merely express the existence of the object in question, but also expresses the "position" or perspective that an agent occupies with respect to it when using the concept. Therefore, talk of being in pain, in this sense, expresses that I am occupying a subjective inward-directed perspective toward the property in question, whereas talk of C-fibres firing expresses that I am occupying the distinct objective and outward-directed perspective with respect to the same property. This is an essential point to my view, which we will develop further throughout of this thesis (particularly in chapter 7 where we develop our neutral understanding of perspectives based on the enactive approach).

From the discussion above we can then extract a further conclusion: having these two sets of concepts, each with a distinct logic and genesis, does not entail that this distinction is reflected on the ontological level, following the heart of the Phenomenal Concept Strategy. Rather, that the distinction is a fact based on the agent's availability of two distinct and special perspectives. A crucial element of my thinking is that the kind of perspectival duality we are exploring here, which the two types of concepts are rooted in, is a duality that necessarily arises from being a subject-in-the-world, myself and other, as opposed to there being a single universal entity that has no other-ness to look onto, and thus having no chance of swapping her perspective from inwards to outwards. In other words, the conceptual duality we are concerned with is one that is intrinsic to any (conscious) subject that exists amongst other entities, thus in a world that consists of more than the entity itself (this is so even if the entity in question were to be the only entity other than the world she exists in), and who has the capability of interacting with the world. We will discuss how this duality of perspectives arises at length in chapter 7, in the context of the enactivist approach to cognition. For now, nonetheless, note that this reasoning about the duality of perspectives suggests that the dualism of concepts can

intelligibly be said to stem from the dualism of perspectives available to the agent, while it seems to tell us nothing about a putative distinction on the ontological level.

4.3 Introducing the Limited Nature of Perspectives

Before concluding, I want to spend some words on the *limited* nature of perspectives. This is an important feature of perspectives, which we return to repeatedly from now in the rest of the thesis. In this section, I focus particularly the commonly (and often silently) held assumption that only the first perspective is special and limited, while the objective is not.

In effect, that a subjective perspective is limited is mere common sense: it is embedded in the definition of "subjective" that such a perspective will tell us something about the world that is peculiar to one single view-point, grasping what the world appears or feels to the subject herself, a restricted window on reality. Stoljar's (2005) Experience Thesis makes this point clearly, as he explains that a phenomenal concept is acquired by the subject only upon having the relevant experience, namely when a subject occupies an inward-directed perspective onto the world. The subjective is thus limited to the one. On the other hand, the objective perspective is one that provides us information in a way that specifically transcends the experience of the single subject and captures a reality that is shared amongst subjects, amongst the many. It is thus supposed to inform us about what the world is (putatively) like, rather than what the world appears or feels to myself or to you. The objective is therefore deemed centreless, a *view from nowhere* as Nagel (1986) labels it. What I want to highlight here is that, though a centreless, it is certainly not a view from nowhere. The objective perspective, in effect, constitutes a specific point of view: it is the perspective of the third-person, of the 'outsider,' of a human being looking out into the world. The crucial point here is that occupying a centreless perspective for a subject such as myself or you, does not equal occupying a perspective-free position, one that could be only be occupied by some Eternal Being. Rather, an objective perspective is merely a perspective that differs from the subjective only with regards to how reality is encountered and investigated. It is nonetheless a perspective and as such limited by its very definition, by its very nature. In my opinion, this emphasises how the special nature of phenomenal concepts and the special nature of physical concepts really are analogous.

In chapter 5 we further unpack this notion. Specifically, in the first part of the chapter I focus on the philosophical limitation and incompleteness of the scientific picture of the world, which relates to the limits of the objective perspective just outlined; while in the second part of the chapter I discuss how these limitations of perspectives can be used as guides to ontology and woven into our metaphysical theories as virtues, rather than as vices. Moreover, in chapter 7, I flesh out the idea that occupying a perspective is a neutral state of the subject, based on the enactivist model.

4.4 Summing-up

In this chapter I have extended the scope of the Phenomenal Concept Strategy to physical concepts; I have discussed the special nature of both phenomenal and physical concepts based on the fact that they are obtained in wholly different, almost opposite, ways as well their ability to capture wholly different features of our existence as experiencers-in-the-world. I have discussed how these two distinct conceptual frameworks ultimately rest on two distinct perspectives that an agent can occupy with respect to reality: one inward-directed and subjective and the other outward-directed and objective. We have thus opened up the possibility that we have two distinct conceptual schemes, each with their own peculiar logic and each resting on a distinct perspective, and that the availability of these does not entail that this distinction is reflected on the ontological level, following the Type-B physicalist. Rather, it should be emphasised that concepts are linked to specific perspectives and it seems that the possession and uttering of a phenomenal or physical concept expresses that a property is encountered from a specific perspective. That is, when I use phenomenal concepts I capture that I have encountered a property in the world from the subjective perspective and when I use physical concepts I express that I have encountered a property in the world from the objective perspective (like indexicals), where this does not entail that there are distinct properties in the world. Lastly, I have introduced the idea that both the subjective and the objective perspectives are limited.

CHAPTER 5:

FROM SUBJECTIVE AND OBJECTIVE PERSPECTIVES TO NEUTRAL REALITY

The question of consciousness often revolves around the question of how experience fits within the physical world. Physicalists, for instance, can be said to attempt to "beat" the phenomenal into the physical, to use Nagel's (1980) colourful wording, the subjective into the objective in order to explain consciousness, while simultaneously defending their assumptions as to the truth (and fundamentality) of physics. In this chapter I try something different: I attempt to weave consciousness and the physical world together, without assuming at the outset the fundamentality of either. Instead, I start from what we know about reality from the subjective and objective perspectives and use this as a *map* to ontology. Therefore, I analyse what they tell us about reality, their origins, their contexts, their value as taken together, and only then allow the most adequate metaphysics to gradually emerge out of it. The conclusion we find in this chapter is that a commitment to a neutral (monist) ontology is a coherent result.

Chalmers (1995, pp.13-14) writes that to answer the of the question of consciousness

> we must seek to systematise the information we have, to extend it as far as possible by conceptual analysis, and then make the inference to the simplest possible theory that explains the data while remaining a plausible candidate to be part of the fundamental furniture of the world.

This captures the essence of my motivation and reasoning here, as well as a point of contention. The essence is to systematise what we know about the world from the two perspectives,

therefore the data we capture with the physical and the phenomenal conceptual schemes, and to extend it by means of conceptual analysis. This, for me, entails looking at the genesis of the conceptual schemes, a project we started in chapter 3, especially with reference to the physical. More specifically, I review how the great success of the sciences has led to their becoming the privileged description of reality, thus an intellectual paradigm, the philosophical shortcomings attached to this kind of approach and whether and how this can be re-framed or re-positioned within our intellectual landscape to allow for the data we get from the subjective perspective to also be used as a useful guide for discovering reality. Systematising what we know by conceptual analysis will, therefore, revolve around the notion that what we know from science belongs to merely one of the two perspectives, in conjunction with the idea that both perspectives are essentially limited points of view onto reality (a discussion we began in the previous chapter and which we expand on here). Extending it as far as possible by means of conceptual analysis ultimately land us on a neutral monist ontology.

The point of contention with Chalmers' quote above, on the other hand, revolves around the idea that the information we have should remain part of the fundamental furniture of reality. What we know about the world is captured by phenomenal and physical concepts. These have been assumed by traditional dualists and monists to constitute the fundamental (or derivative) furnishing of the world, such that mental and/ or material concepts have been posited as metaphysical entities. However, I want to raise the question of why this should be so. I feel that an adequate systemisation of what we know, and its adequate extension by conceptual analysis, must not necessarily lead to having mental and material metaphysical counterparts to the concepts. In other words, that an adequate explanation of the reality of experience and of the world that physics tells us about does not entail we must postulate (bona fide) mental and material properties.

The core idea of the chapter, in summary, is the following: both the outward-directed perspective and the inward-directed perspective tell us something different about our existence as experiencers-in-the-world. Specifically, they tell us something *true* about reality: together they indicate that there are properties in the world and that we can encounter these properties either in a subjective or in an objective manner. When we burn our hand in the oven we feel the scorching pain from the inward-directed subjective perspective and capture the fact that the

heat is damaging the cells of our skin by occupying an outward-directed objective perspective (as explored in the previous chapter). Only both taken together, however, produce a full account of the event in question; thus we admit that the scientific account should not be pushed aside for the benefit of producing a view that adequately explains consciousness (as in mental monism), and that consciousness deserves an equal treatment (the pitfall of material monism). What I have in mind here is that given that both perspectives inform us about reality equally and, given that we know that both perspectives are limited and special in their own right, there is no reason to subordinate one conceptual framework to the other. Moreover, if we conceive of the phenomenal and the physical with reference to perspectives, the idea that they should be ontological categories shakes. From this, we move on to posit a neutral ontology, discussing in depth how an acceptance of the limited nature of perspectives motivates a commitment to an ontology that is neutral, thus weaving limitations into our theory as a virtue rather than as a vice.

In section 5.1 I take up a reflection on the objective perspective, in particular with reference to the success of physics, the choice of the physical conceptual scheme as a privileged description of the world, the philosophical shortcomings associated it and the need for acknowledging the limited nature of the materialist paradigm. In section 5.2 I begin to lay out how we should think about bridging the gap between the two conceptual schemes, I question the validity of adopting a single conceptual scheme as a privileged description of reality, I lay out the elements of our map from concepts and perspectives to ontology and proceed by looking for an appropriate truth maker for our phenomenal and physical concepts. In section 5 I discuss the limitations of being human, manifest in our subjective and objective perspectives, and I argue that these limitations motivate why a commitment to a neutral ontology may be desirable and show how these can be embedded in our theory as virtues.

5.1 A Considerations on the Scope of Science and its Philosophical Outcomes

Since the scientific revolution, modern science, and physics in particular, has proved to be an extremely successful explanatory project. In *Science and the Modern World* (1925) Whitehead surveys the development of modern physics from its 17th century origin, pinning the

essence of the revolution on the shift from the priorly established metaphysical method (inspired by the Aristotelian tradition) 'in favour of the study of empirical facts of antecedents and consequences' (Whitehead, 1925, p.39). Whitehead brings attention to a number of features that characterise the newly elaborated scientific method, such as the role of measurement and inductive reasoning, as wholly expressible in the language of mathematics, as also discussed in chapter 1. The power of these empirical tools has effectively laid the ground for the great progress of science we enjoy today, not only in the lab but also in our daily life, enabling technological developments such as the atomic bomb, smart phones and electric cars.

Describing the various developments originated in the 17th century with reference to motion and mass, for example, Whitehead (1925, p.48) comments that 'having regard to this triumph, can we wonder that the scientist placed their principles upon a *materialistic basis*, and therefore ceased to worry about philosophy?' By asking this question Whitehead highlights the achievements of science, but he also introduces two key issues:

(i) over the centuries scientific thinking has developed to be independent from philosophical thinking
(ii) science has come to constitute the (material) paradigm that underpins our intellectual inquiry (everything is to be examined in objective scientific terms)

These two points are, in effect, a preamble to his core claim: that science, and more specifically the assumption that he calls *scientific materialism* is 'entirely unsuited' (Whitehead, 1925, p.17) to adequately explain and understand nature as a whole, hence including consciousness.

> There persists (…) throughout the whole period [17th century] the fixed scientific cosmology which presupposes the ultimate fact of an irreducible brute matter or material, spread throughout space in a flux of configurations. In itself such a material is senseless, valueless, purposeless. It just does what it does do, following a fixed routing imposed by external relations which do not spring from the nature of its being. It is this assumption that I call 'scientific materialism'. (Whitehead, 1925, p. 22)

Although we may no longer be committed to an irreducible brute matter in physics, the qualities that Whitehead lists with reference to the material are certainly the same that still characterise the subject matter of physics today: senseless, valueless and lacking in purpose (or anything else that could help us make consciousness intelligible). Whitehead's quote is as relevant now as ever, 100 years later. In addition to this, it is also true that the subject matter of physics regards what matter does, the external relations that hold between the various entities that physics posits. In effect, the very strength of physics is precisely its ability to model the *causal structure* of the world and thus explain how things ultimately *work*. This is the root of the causal closure principle discussed in chapter 1, and also the reason this assumption carries the strength it does. But this idea also represents the pitfall of scientific materialism when the paradigm is used as a tool for explaining nature as a whole. In other words, that is why the materialistic assumption gives rise to the Hard Problem of consciousness: how can our rich experience arise from senseless matter? And also to the Explanatory Gap: why should this physical structure correspond to this* phenomenal event? These question capture the reason why Whitehead argues that the materialistic assumption is 'entirely unsuited' (1925, p.17) for producing a complete picture of the world in the contemporary landscape. Whitehead's sentiment is shared by a number of other scholars working at the beginning of the 20th century, such as Eddington (1928) and Russell (1927), and it can be said to be shared by many in the contemporary discussion on the nature of consciousness, particularly by defenders of more innovative views such as panpsychism (Goff, 2017) and Russellian monism (Alter and Nagasawa, 2012), panprotopsychism (Chalmers, 2013) who attempt to solve the inadequacy of the materialist paradigm to explain consciousness by injecting it with an extra ingredient (as seen in § 24).

Russell (1927) specifically discusses the issue that physics, while being accomplished for its own scientific empirical purposes, is lacking in terms of its *philosophical outcomes*. Thus, while physics successfully provides us with some explanations, it has limited power to produce an adequate picture of the world which requires us to think philosophically, or otherwise more holistically, about the world. In his introduction to the *Analysis of Matter,* Russell (1927) points out that physics severs the empirical world of science from the world of perception, based on the fact that physics does not account for the world of perception, rendering the two fields of enquiry discontinuous. This mode of reasoning is fallacious according to Russell, since the data

we obtain from science really is data that we obtain *because* we are experiencing subjects. In this sense Russell (1927), like Whitehead (1925), believes that the philosophical outcome of science as a whole is not as successful as the empirical results that it produces strictly for scientific enquiry. Russell (1927, p.7) tells us that we need to find an interpretation of physics that gives due space to perceptions, if we are to coherently resort to empirical evidence. However, when the materialist paradigm is adopted without interpretation, this maintains the incommensurable distinction between the world of physics and that of perception that we find in the problem of consciousness. We have an example of this in the physicalist ontology and their difficulty to explain the phenomenal.

Recall that in the introduction to this chapter I mentioned I would use the objective and the subjective perspectives as maps to our ontology. The reason that motivates me to do so is precisely a desire to avoid this kind of fallacy and be able to think philosophically about the issue, thus embrace the scientific picture although with adequate interpretation. This is also an example of my fresh eyes methodology at work.

5.1.1 The Materialist Paradigm: The Historical Origin of the Problem

Goff (2019) suggests that the root of the issue with the scientific paradigm may in effect be of historical nature, he labels this *Galileo's Error*. I agree with him and I deem it important for us to also jump back in time and analyse the origin of the error in modern science.

In the first chapter we referenced Galileo's suggestion that the language of the universe is expressed in mathematical-geometrical terms and that, in order to explain the mechanics of the world, secondary qualities (namely those that characterise *what it is like* to undergo an experience) had to be *removed from the field of scientific enquiry,* because they belonged to the works of the soul of subjects rather than to the external world. Specifically, while primary geometrical qualities were certainly part of the external world, sensory qualities like the sweetness of the smell of lavender or the redness of a rose belonged to the inner realm of subjects. Galileo's suggestion was then calcified in our philosophical thought by the hand of Descartes, as seen in chapter 1. I have discussed that, with the aid of the Evil Demon, Descartes (1996) sheds doubt on the value of the data we obtain by means of sensory perception and

argues that the only thing we cannot doubt is our intellect (the *Cogito*). This has an interesting result: it places the knowledge we obtain intellectually, by means of reasoning, at the top of our knowledge-hierarchy and assimilates knowledge about our bodies or our perception as a form of intellectual knowledge. While we could doubt that I feel pain because I am truly being hurt, that is we could doubt that I am undergoing an experience of pain resulting from my skin being burned by a hot pan for instance, I cannot doubt that I think (or believe) that I am in pain. The epistemic value of the phenomenal (as a felt and embodied experience) is thus cast out of the equation, to the benefit of the purely intellectual and objective. As a result, our self-knowledge, or knowledge *about* our phenomenal state, is placed on the same level as scientific knowledge on the grounds that they are inherently intellectual processes, while rendering our purely phenomenal knowledge epistemically redundant with respects to producing truths about reality.

From my analysis, we distil two issues that restrict the scope of the physical sciences: (i) phenomenal data has to be removed from our analysis of the world because it belongs to the soul/ inner world and (ii) phenomenal knowledge is unreliable for the purposes of producing a reliable explanation about the world. It follows that only scientific knowledge is reliable for describing the world. We therefore come to have a scientific framework that excludes the subjective perspective, and thus phenomenal data, from its method as a necessary step for us to understand the mechanics of the world.

I want to draw attention to the fact that the removal of phenomenal experience from scientific enquiry was *epistemically necessary* for successfully doing science. The extensive successes of science, however, do not justify the adoption of the materialist assumption as *the* intellectual paradigm for producing a coherent explanation of reality as a whole (one that includes consciousness) and thus neither as an assumption for a monist ontology (as for the physicalist). Galileo explicitly states that phenomenal properties had to be discarded to uncover the mechanics of the universe, but that does not mean that the removal of secondary qualities is essential to produce a complete explanation of the world which comprises of consciousness. This is the crux of the problem. Moreover, Descartes' claims about the reliability of phenomenal properties refer to the epistemic value that phenomenal knowledge carries with reference to explaining the extended mechanical world only, rather than on their value for the production of

a complete understanding of the world that includes us as subjects of experience. The essence of Descartes' project in the *Mediations* (1996) is to shed doubt on phenomenal knowledge with reference to how much our experience as subjects can, or cannot, tell us about the mechanics of the world specifically (Patterson, 2000), but this should not be understood as the idea that phenomenal properties had to be subclassed in a complete metaphysical picture of the world. This Galilean-Cartesian reasoning, thus, concerns specifically the epistemic rather than the ontological status of phenomenal experience. The epistemic and the metaphysical reasoning are very starkly distinguished here. Both Galileo and Descartes are striving to build a reliable framework for understanding the mechanics of the world and *in order to do this* sense perception cannot be taken into account. But if we are to give an account of the universe as a whole I am doubtful either of them would agree that the materialist paradigm is adequate for the task. Descartes himself introduces the immaterial *res cogitans* alongside the extended *res extensa* precisely to produce an ontology that can accommodate for the phenomenal *on par* with the physical.

To sum up,

> If we confine ourselves to certain types of fact, abstracted from the complete circumstances in which they occur, the materialistic assumption expresses these facts to perfection. But when we pass beyond the abstraction, either by subtler employment of our senses, or by the request for meanings and for coherence of thoughts, the scheme breaks down at once. The narrow efficiency of the scheme was the very cause of its supreme methodological success. (Whitehead, 1925, p 17).

5.1.2 Acknowledging the Incompleteness of the Materialist Paradigm

We have seen that the materialist paradigm, and thus what we know about the world from the third-person perspective, 'has been so irresistibly attractive, and has so dominated ideas of what there is, that attempts have been made to beat everything into its shape and deny the reality of anything that cannot be so reduced' (Nagel, 1989, p.80). However, 'this bleached-

out physical conception of objectivity encounters difficulties if it is put forward as the method for seeking a complete understanding of reality' (Nagel, 1980, p. 80), precisely because the objective transcends the subjective; it expressly excludes it.

The scientific paradigm is built upon the objective perspective, hence as an agent occupies an outward-directed perspective towards reality. As we have seen in chapter 4, this is distinct and even opposite from the subjective perspective and we know that, according to any monist for example, an agent can switch between the two with respect to the same. In this sense, the split of the mental and the material and the exclusion of the phenomenal from physical enquiry reflects the natural states of affairs of an agent having these two manners of interacting with reality and our ability to capture reality accordingly. Confusion arises, however, when the *view form nowhere* that the objective perspective putatively affords us is mistaken with regards to what and how much it can tell us about the world. As Nagel (1980, p.90) specifies, 'the objective, not to be false must include an acknowledgement of its own incompleteness.' But, generally, when we take the materialist paradigm as our privileged description of the world, and furthermore posit its fundamentality, it is precisely the admission of its limit, thus of its incompleteness, that we lack.

On the other hand, if we take the scientific paradigm as merely only one of the possible interpretations of reality, one that is rooted in only one of the two perspectives that an agent can occupy with respect to reality, then this renders the scientific conceptual scheme a framework with strengths and limits, which can in turn be woven into a broader picture. The physical perspective can thus serve as a guide for us to discover something, although not everything, about the world. This should then provide us with an essentially more flexible space for inquiry. That is, certainly the materialist paradigm should be privileged for doing science, but its privilege dropped for the purposes of accounting for a reality that includes consciousness. This is because (i) consciousness was specifically omitted from the scope of its enquiry and (ii) the contents of consciousness just cannot be held in the third person perspective. Admission of the limits of a conceptual scheme can mean that the limits themselves become important elements to guide us towards a coherent metaphysical theory that can adequately accommodate both perspectives. This is how I want to take conceptual schemes generally, and the materialist paradigm particularly, in the remainder of this thesis.

5.2 Bridging the Gap

Seeking a coherent ontology, one that can adequately accommodate the phenomenal and the physical, Russell (1927, p. 7) suggests 'we should assimilate the physical world to the world of perceptions and to assimilate the world of perceptions to the physical world.' This is no ground-breaking idea, rather it is precisely the desire that motivates any form of monism. However, traditional monists have taken Russell's suggestion above in a one-directional sense, thus either assimilating the world of perceptions to the world of physics for the material monist, or assimilating the world of physics to that of perceptions for the mental monist.

It seems to me that this one-directional process is rooted in the assumption that there must be a privileged description of reality, such that only one conceptual scheme can adequately inform us as to what reality is really like. This is then paired with the idea that these privileged concepts (a) stand in a one-one correspondence with facts in the world and (b) are structurally isomorphic with the facts they pick out. Here I'm readapting this from Dyke's Strong Linguistic Thesis (2012) where she criticises the overwhelming reliance on linguistic analysis on the part of contemporary metaphysicians to discover the nature of reality. This, in effect, constitutes the transposition of a scientific-type methodology, or more specifically a methodology that attempts to offer a degree of objectivity that is comparable to the explanatory power of the scientific method into philosophy (Farkas, 2019), where the philosophical viability of this methodology is questionable. Take the physicalist position as a case in point. The physicalist chooses the physical conceptual framework as the privileged interpretation of reality and takes concepts to pick out specific entities in the mind-independent world, such that these entities are the way our concepts describe. We thus have the one-one correspondence plus structural isomorphism pointed out above. The physicalist then attempts to explain how phenomenal concepts, which appear so different, ultimately refer to the same physical reality, and they generally do this by means of reduction. This is problematic on many counts, and not just because it attracts its bespoke versions of the mind-body problem. Rather, I feel the issue with this process lies in a choice that is arbitrary, however well motivated, of a privileged description of reality. Specifically, the issue is that choosing a single conceptual scheme, such as the physical, as the basis of our ontology automatically defines and heavily constrains the logical

space for inquiry, as well as the nature of the ontological landscape we are developing. The result is that a whole other set of very informative concepts, such as phenomenal concepts, becomes explanatory redundant for informing us about ontology.

Note, however, that Russell (1927) suggests that the assimilation process should be bi-directional: the world of physics is to be assimilated to perception *just as much as* perceptions need to be assimilated to the world of physics. He continues by saying that 'physics must be interpreted in a way that tends towards idealism, and perception must be interpreted in a way that tends towards materialism' (Russell, 1927, p.7).

In the rest of this chapter I want to take up Russell's suggestion and attempt to assimilate the two worlds bi-directionally. First, I begin my enquiry without committing to any privileged description of reality. Here I take a fresh eyes stance, again. This broadens the logical space available for me to think about reality, because it enables me to investigate the value of the concepts and the nature of the phenomena they pick without bias. This, in turn, puts me in the position to ask further questions such as: what do mental and material concepts ultimately pick out? What should ultimate reality be like, if we have two sets of concepts available to us to describe it? Do concepts stand in a one-one relation to the entities they describe? And are they structurally isomorphic with the reality they pick? I thus turn my attention to phenomenal and physical concepts, and the objective and the subjective perspectives taken conjunctively, to draw a map from what I know towards a coherent ontology that can adequately accommodate both.

5.2.1 Begin with What We Know: The Elements of our Map

The starting point is to begin with what we know, following Chalmers's (1995) suggestion. I use this section to summarise what we know, as explored in the the thesis so far.

First, we know that we are conscious beings and that there is a world "out there" that we share with our peers. We know that we are experiencers-in-the-world.

Second, we know that we can occupy two distinct perspectives with respect to realty, one inward-directed and subjective, and the other outward-directed and objective. Moreover, we know that we are experiencing subjects when we occupy the inward-directed perspective and we know that we exist in the world when we occupy the outward-directed perspective. We also know that each of these perspectives allows us an access to reality that is limited and incomplete, thus each informs us differently about reality.

Third, we know that there are two distinct conceptual schemes available to us to capture our existence as experiencers-in-the-world. We also know that each of these two sets of concepts corresponds to, or reflects, the world as seen from one of the two perspectives. When we stand in the subjective perspective we characterise the world as phenomenal, thus with phenomenal concepts, and when we stand in the objective perspective we characterise the world as physical, thus with physical concepts.

In addition to this, we also know that standing in an objective perspective and adopting a physical conceptual scheme allows us to depict a world that is causally closed, and that this is the output of a project devised to analyse nature by overtly omitting an important feature of it, i.e. phenomenal experience, while focussing on the causal/ mechanical features of the universe. This is a crucial point as it highlights that the physical conceptual scheme is limited just like the perspective it is rooted in. Similarly, we know that phenomenal concepts allow us to capture that *there is something it is like* to undergo an experience, thus the richness of my inner private life, however they cannot inform us about the structural features of reality, no matter how rich they might be. This is also a crucial point as it highlights that the phenomenal conceptual scheme is limited just like the perspective it is rooted in. Moreover, we know, from chapter 4, that phenomenal concepts cannot be derived from physical concepts and vice versa, as each is obtained only from the relevant perspective, and they are thus essentially distinct. It follows that phenomenal and physical concepts are *mutually exclusive*.

From the above, we can infer that if we want to give a satisfactory account of our existence as experiencers-in-the-world we must employ both sets of concepts, for each will explain an essential and distinct feature of it, while being unable to capture the whole story because of its limitations. For instance, when I burn my hand I can capture and express the story

of my pain from the subjective perspective and rely on phenomenal concepts, but this will tell nothing more than what the events feel like *to me*; whereas I have to step into an objective perspective and rely on physical concepts if I want to capture and express the causal story of the burning event that is accessible *to all*, yet this will reveal nothing about my subjective experience.

Lastly, we know that taking either the mental or the material conceptual scheme as a privileged interpretation of reality, and positing the fundamentality of the entities they pick out, leads us to the problems that the traditional monist faces. We know that the root of the issue is precisely that the concepts are the result of two distinct and incommensurable perspectives. Once we design a concept specifically to exclude the other, especially in the case of two contrasting things such as internal and external, the task of deriving the latter from the former or vice versa is an impossible one by definition. The concepts are mutually exclusive, but also *complementary*. It follows that if one posits a stuff that is defined by either of the conceptual schemes, deriving the non-fundamental property defined by the opposing conceptual scheme is unlikely, to say the least.

I want to highlight that nothing in the above discussion suggests that there might be any metaphysical distinction between what mental and material concepts pick out, neither anything that point us in the direction of mental and material concepts picking out ontological categories. Moreover, knowledge of the problems that the traditional monist encounters, in conjunction with the fact that an adequate description of our existence as experiencers-in-the-world is achieved by employing both mental and material concepts, incentivises us to think about reality without privileging any of the conceptual schemes.

5.2.2 Finding a Common Ancestor for Phenomenal and Physical Concepts

The question now is: how can we combine two sets of concepts that are mutually exclusive? The key idea is that while concepts can be guides to ontology, elements of our map, we should refrain from thinking that they appropriately uncover the nature of reality, in the sense

that these might be isomorphic and hold a 1:1 relationship with reality, as this would mean to allow language to 'define things into existence' (Van Inwagen, 1990, p.9). I therefore agree with Heil when he argues that:

> To imagine that to every significant predicate there corresponds a property, is to let language call the shots ontologically. It is true that there is a billiard ball and that the billiard ball is spherical, red, and has a particular mass. But what makes these truths true – their truthmakers – is a particular arrangement of substances of particular kinds (Heil, 2013, p. 204)

The heart of the matter is the following: if we think that talk of the phenomenal and talk of the physical are true, then these concepts need to pick out some real entity in the world. That is, they need to have a truth-maker, as Heil (2013) points out. But the fact that a truth-maker can be described with distinct, and in this case mutually exclusive, concepts does not entail that the concepts are made true by distinct properties. Neither should we assume that these concepts tell us exhaustively or accurately about the metaphysical nature of the truth-maker they refer to. In other words, distinct concepts allow us to describe reality differently, on the basis of the perspective-orientation in our case, but this does not entail that each concept picks out some distinct entity or that it tells us about the metaphysical nature of the entity in question. Moreover, in a monist ontology the truth-maker for both physical and phenomenal concepts must be the same. But we do not intend language to call the shots ontologically and thus we question whether any of the two sets of concepts ultimately describes fundamental reality, as well as whether phenomenal or physical concepts may pick out isomorphic properties/ entities.

Rather, by following the map of what we know, and especially the failures of traditional monists, we reject the idea that either phenomenal or physical concepts pick out (isomorphic) entities and thus we hold that neither mental nor material entities act as truth-makers for phenomenal and physical concepts. We merely use talk of the phenomenal and of the physical as a guide for discovering reality, based on the fact that they provide us with distinct but equally important information about it. This prompts us to look for a truth-maker that must equally and adequately make true both sets of concepts, therefore pointing us in the direction of something

more "primitive." The route towards the discovery of this sort of truth-maker is the following: both sets of concepts pick out properties in the world, but they crucially do so only in relation to the perspective they are associated with. We therefore do not have a direct relation between a concept and a property in the world, rather this relation is mediated by the perspective that the agent occupies when she talks in physical or in phenomenal terms. That means that phenomenal concepts pick out properties *as acquired* by occupying an inward-direct perspective, thus as experienced subjectively, and that physical concepts pick out properties *as acquired* by occupying an outward-directed perspective, and thus by means of objective inquiry. This gestures us towards a conception of property P that must equally be accounted for by both sets of concepts, where neither can be said to be a privileged explanation of it. The addition of perspectives to the story of the relation between concepts and entities relaxes the requirement that there must be a privileged description of reality, as well as the idea that our concepts pick out isomorphic properties. In other words, having a mediator between the concept and the property broadens the logical space for thinking that the properties in question may be inherently different than what we capture by means of phenomenal and physical concepts.

It seems to me that the most appropriate truth-maker for our phenomenal and physical concepts is a neutral property, on the grounds that it adequately fills the role of a more primitive kind of stuff, that *common ancestor* that Russell (1921, p.11) suggested we should seek. This is because a neutral entity is something that can equally be accounted for in both mental and material terms and, therefore, can in principle make it intelligible why talk of the mental and talk of the material are true. Of course, the details of how this may be the case can vary greatly. For instance, for the *big three* it is the arrangements of neutral entities that function as a truth-maker for mental and material concepts, where mental and material concepts can be held to stand in a 1:1 isomorphic relationship with the arrangement of entities they pick out. In the case of Heil (2013), on the other hand, mental and material concepts are made true by the same neutral properties, albeit differently considered by a process of mental abstraction, where concepts do not stand in a 1:1 isomorphic relationship with the entities they pick. For the perspectival neutral monist, as we will see in more detail in chapter 7, physical and phenomenal concepts are made true by the same neutral property as encountered by a specific (neutral) perspective, although neutral entities will not turn out to stand in a 1:1 and isomorphic relationship with the concepts.

From this it seems that neutral entities are good candidates to fill the role of truth-makers for phenomenal and physical concepts, in that they coherently allow us to assimilate the phenomenal to the physical and the physical to the phenomenal, thus to carry out the assimilation process bi-directionally as suggested by Russell (1927) above.

5.3 Limitations as virtues: committing to a neutral ontology

A reflection on our limited nature as human beings is a further, yet central, motivation for committing to a neutral ontology. That is specifically a reflection on the fact that our perceptual, cognitive and physical capacities are limited, such that there are specific, albeit flexible, ranges of actions we can carry out, thoughts we can think, emotions we can feel, processes we can undergo, and so on. Our perspectives, accordingly, are limited, as introduced in the previous chapter, §4.3. We live in a world to which we somehow grasp the grandeur of, but cannot wholly hold knowledge of, at least not consciously.

Being limited is, generally in common-sense, taken as a negative trait, namely as a form of restriction on how much we can do, know or feel, a prison that looms over us, making us small, incapable, ignorant. Limits are inherently something that we strive to overcome. Doing science and metaphysics are good examples of how we attempt to overcome our intellectual perceptual limitation, for instance, and to grow beyond these limits. However, this being limited is also precisely the trait that makes us properly human. We are not gods, instead we have bodies and we can feel, we can run, we can have opinions, we can make mistakes and then try to do things better, again and again, until our time on this earth expires. As such, limits inherently constitute the very essence, and thus beauty, of being human. It is precisely what allows us to feel so good when we fall in love, or to feel pleasure when we smell the sweet scent of a rose, or enjoy the first heat of spring on an early March day.

There is a powerful and overwhelming humility that comes with appreciating these limits, that comes with appreciating that we are not perspective-free, eternal beings. And this humility is one that I strive to embed in my ontology-making and that encourages me to posit neutral entities, rather than the (more restrictive) mental and material ones. What I mean to say is that, as

a human being doing ontology, I want to weave into my theory the fact that I have a specific knowledge of nature that depends on my human-ness, which yet fails to capture the whole of reality. This, in practice, is well exemplified with reference to the perspectives. The perspectives allow me to capture, and to a certain extent even shape (as we will see in more detail in chapter 7), a reality that is per se much greater than what we grasp: the worlds I find in science and in my experience are real, but incomplete with respects to the totality of ontological facts about the universe. The notion of neutrality, as paired with the perspectives, is therefore my attempt to embed these limits as a *virtues* in my ontology making, by giving them a centre-stage position in my theory. These limits inherent in our human nature constitute a further motivation why I feel a commitment to a neutral ontology is coherent and possibly desirable, especially if we are after an explanation of reality that includes consciousness together with the physical world. We are limited, therefore why should we assume that the concepts we have from the two perspectives have a one-one isomorphic relation to reality? Committing to a neutral reality is therefore best seen as a gesture towards the fact that I am limited and can only grasp part of what reality is. There is a beauty in the notion of a neutral entity, an elegance, inherent to the fact that we can grasp a part of it, but a part of it remains outside of our knowledge.

Stoljar (2006) is moved by a similar sentiment in his *Ignorance Thesis*, briefly outlined in the introduction. He argues that we simply do not have all the relevant facts about fundamental reality necessary to explain consciousness, specifically that we are ignorant about a set of (non-standard material) properties that are consciousness-relevant, the knowledge of which would allow us to understand why consciousness obtains at the macro-level. Stoljar reasons that certain challenges such as the Conceivability Argument (explored chapter 2) impose on us very high epistemic demands, namely that all the (physical) facts about the world lay right in front of us. Accordingly, adopting the *Ignorance Thesis* and positing non-standard material properties should ease those high epistemic demands. We flesh out the details of this strategy at the end of the thesis, in §7.4.1. For now, I simply want to draw attention to the fact that this move extends the scope of the concepts that are available to us, the physical to be specific, and therefore affords us a broader logical space for thinking about consciousness. Stoljar's (2006, 2019) strategy could thus be interpreted as putting into practice Nagel's (1980, p.90) suggestion that 'the objective not to be false must include a mention of its limits,' a point we agree with. It should now be obvious how Stoljar's position bears similarities to my own, in that it embeds ignorance

as a virtue of his view, and furthermore how this motivates him to posit unknown (non-standard physical) entities as constituting (part of) the fabric of the world. However, Stoljar's (2020a) argument is different from ours on two counts. First, it pairs unknown (non-standard material) properties alongside the set of properties that physics tells us about, whereas in our reality all *bona fide* properties of the world are never wholly captured by science (or phenomenal concepts). Second, his strategy is based on ignorance and it is thus a discourse on the strictly epistemic position we occupy with respect to these unknown consciousness-relevant facts, hence a discourse based on what we know and what we do not know. On the other hand, as we will see in chapter 7, the limits I am concerned with are not strictly epistemic (although closely related to them), but depend on how we define the subject as a whole in a neutral reality, thus how much of the neutral reality can be available to the subject from the two neutral (non-mental) perspectives.

With this in mind, we can look at the notion of neutrality with fresh eyes, expanding on the *Neither View* and the *Both View* defined in chapter 3. We can begin to see the neutral stuff as something unknown to us, specifically as something that is more than what we know from each perspectives or both combined. In effect, I feel that a proper characterisation of the neutral stuff is one that necessarily prescinds the perspectives of an experiencing subject and, rather, could only be given from a perspective-free position akin to those that held by gods or eternal beings and thus similar to Spinoza's *Sub Species Aeternitatis* (1677). Seager (2011, p.87) describes this as 'an infinite all encompassing substance which was itself neither mental nor material. The substance possessed an infinite number of attributes, only two of which our minds can comprehend: matter and consciousness.' From the eternal perspective, a perspective-free position, we could describe and enjoy reality in its multi-dimensionality, as a whole, but as human beings all we can do is describe and enjoy reality as available from our two perspectives. From this we can expand our notion of the neutral stuff as an infinite stuff, something much broader and much greater than what is available to us in the two perspectives (even combined); a *multitude*. In my view, this notion of neutrality as a multitude can complement well the orthodox *Neither View* and *Both View*, rendering the notion more intelligible as well as functionally more flexible. In the rest of the thesis, therefore, when I talk about the neutral stuff I will be referring to a stuff of the world in accordance with the Neither View and Both View outlined in §3.2 alongside the idea that it is a *multitude* as explored here.

In addition to the above, please note that my act of humility in this case should not be mistaken for the idea that we cannot know anything about what the world is really like, nor that the data I get from the two perspectives ultimately represents an iron curtain which we are unable to peak through, and thus that doing metaphysics is wrong. Quite the opposite, in fact. 'Our thoughts do not constitute a veil or curtain interposed between us and the things we are endeavouring to think of, somehow making them inaccessible or inscrutable to us. On the contrary, things are accessible to us precisely because we are able to think of them' (E. J. Lowe, 2002, p.14). With Lowe, I do believe in the power of our intellectual capacities to push through what each perspective singularly tells us, to think philosophically, and be able to create theories that exceed what we learn from within the chinks of our perspectives. Science is indeed a proof of that, and so is our metaphysical thinking. However, I do strongly believe that this sort of humility has to be built upon and set in the centre of the stage as we try to understand what the world is like. This captures the heart of the neutral ontology, as well as of the perspectival neutral model I build in this thesis, which I flesh out in chapter 7.

Hence, quoting Broad again, 'we cannot be acquainted with Reality as a whole, as we can with a tune or an emotion' and this is something that we weave into our definition of the unknown neutral stuff and thus, I feel that if we embrace this discussion on limits, then the postulation of a neutral stuff becomes natural, rather than arbitrary.

CHAPTER 6:

QUESTIONING THE ONTOLOGICAL HIERARCHY ASSUMPTION

It is customary in metaphysics, and especially in the consciousness literature, to think about nature being hierarchically ordered. As we have seen in chapter 2 and 3, for instance, many views partaking in the debate attempt to explain consciousness by reference to what exists at the lower level, thus in terms of hierarchical structures. This is what I call the *ontological hierarchy assumption*, which generally grounds reduction. In this chapter I explore what an ontological hierarchy is and I question the adequacy of assuming an ontological hierarchy in the context of searching for a coherent explanation of consciousness. I therefore cast my set of fresh eyes onto the issue, and suggest that assuming an ontological hierarchy for the purposes of investigating consciousness specifically may not be effective.

Please note that my critique of ontological hierarchies does not extend to all fields of enquiry, as I acknowledge that its application has been successful in other fields such as physics and chemistry for instance. My reflection here, therefore, does not amount to a blanket dismissal of ontological hierarchies in general, but it is geared solely towards questioning the use of a hierarchical framework in the area of consciousness metaphysics. In particular, I propose that we should avoid committing to any specific ontological hierarchy to develop a coherent method and model of the nature of consciousness.

6.1 Ontological Hierarchies Explained

My use of the locution *ontological hierarchy* refers to the orthodox understanding that the world is ordered into levels of ascending ontological fundamentality, presenting one fundamental level and a series of gradually less fundamental levels that *depend* on the prior levels for their existence, hence constituting a hierarchy of dependence. It follows that an explanation of the less fundamental levels must be carried out with reference to the most fundamental level. For example, that the macro-level of reality that we enjoy (and that we capture with mental and material vocabulary) should be explained in terms of the fundamental stuff that we posit. The fundamental level is conventionally thought to be the micro-level, where quarks exits. However the converse approach, namely that the fundamental level of reality is the cosmic/ all-encompassing level, has also been widely adopted both historically and in the contemporary debate (I will look at this in more details in § 6.2.1 below).

We have encountered the notion of hierarchy many times in this thesis already: in our discussion of the mind body-problem for the physicalist and the alternative views in chapter 2, but also in the context of the adoption of the materialist paradigm in chapter 5. The discussion in chapter 5 is particularly useful to understand why hierarchical frameworks have become so popular in the field of metaphysics, based on the overwhelming success of the sciences that has driven a generation of philosophers to adopt the materialist paradigm, and therefore the scientific hierarchical model, as the basis of our ontology-making. In effect, the notion of ontological hierarchy is one of the most effective tools in the sciences and thus a major point of strength of the materialist paradigm, constituting one of those methodologies that have been imported from the sciences to philosophical enquiry with the hope they would produce objective and reliable results.

It is useful to distinguish between two types of ontological hierarchies: the micro/ macro that we find in the sciences and the mental/ material that we find in philosophy of mind. An example of the former is the explanation of heat as the movement of molecules or of water as H2O; whereas an example of the latter is the reduction of the mental to the material by the physicalist, discussed in chapter 2. It is crucial to distinguish between these two types of reduction because each works differently when put to work. To further clarify, recall our discussion on Levine's Explanatory Gap (1983) in chapter 2. We saw that the description of heat

in terms of the movement of relevant molecules at the micro-level renders the macro-phenomena of heat intelligible, such that heat is wholly explained with reference to the micro-physical phenomenal. Then contrast this with the gap we find when we attempt to explain phenomenal experience, such as pain, with its associated micro-physical base understood as the relevant causal-functional story. The application of the same strategy to the two areas, where the latter is inspired by the achievements of the former, clearly seems to have very different outcomes. In micro to macro case, thus in matters that belong to the same third-person perspective, we can achieve a successful result, namely the macro-level phenomena in question is adequately explained with reference to the micro-level (although questions can question can be raised regarding the reduction of third-perspective macro phenomena such as biology. I will briefly touch upon this § 6.3). On the other hand, in the second case the result is questionable, i.e. the macro-level phenomena of consciousness is not adequately explained by reference to the more fundamental physical base; such that the subjective perspective does not adequately reduce to the objective.

The difference between the ontological types of hierarchies can also be cashed out as follows: the macro-micro refers to a hierarchy within the same field of enquiry (the physical) and thus applies to concepts/ properties within the same physical objective perspective; the mental to material hierarchy aims to reduce one conceptual scheme to the other, as we have seen various times in this thesis. While we may want to defend physical-to-physical reduction, thus reduction of what we know from the third-person perspective to other data we know within that same perspective, it seems undesirable to defend reduction of one perspective to the other. The reason for wanting to avoid assuming hierarchy involving the mental and the material is particularly obvious in the context of material monism.

6.2 Grounding and Constitution

Historically, the relationship that holds between the micro-level physical structures and the macro-level phenomenal consciousness that bodies such as ours display have been explained in terms of identity (Lewis, 1966) or supervenience (Kim, 1982) by physicalists. Recently, however, the notion of *grounding* (Fine, 2001; Schaffer, 2009) has gained traction as it seems to more adequately capture the relation that holds between the relevant levels of reality.

Looking at the grounding relation can thus help us better understand the idea of ontological hierarchy that we are exploring here.

The notion of grounding captures a non-causal relation of dependence between a set of micro-level entities and a higher-level object, according to which the macro-level object is *nothing over and above* the arrangement of the micro-level entities which constitute it. For instance, the existence of a table is grounded in there being a set of micro-level entities arranged table-wise, this means that a table is *nothing over and above* the set of micro-physical entities arranged table-wise and which constitute it. Grounding therefore tells us the story of the non-causal *constitutive dependence* between the lower-level entities and a higher-level object, which is different from the more traditional identity relation because grounding allows us to capture the different identity (or properties) of the structure at the two different levels of reality, while at the same time informing us about the ontological hierarchy that holds between them. The grounding relation thus entails that: (i) there is a level of reality that is ontologically primary, and that it is the micro-level[5], (ii) macro-level bodies depend on their micro-level composition, and (iii) in order to understand macro-level entities we have to look "down" at their lower-level composition. This can be expressed as the idea that the parts are prior to the whole they compose, hence describing a bottom-up ontology where the top level is explained away by looking at the bottom level, which is ontologically prior. From this, it becomes clear how grounding conceals reduction at its core.

The above reasoning is then adopted by the physicalist, as rooted in the (materialist) assumption that physics tells (or can tell) us a complete true story of what reality is like at the fundamental level. While it is obvious that the traditional material monists would follow this line of reasoning and posit such a relation to explain the nature of consciousness, it is interesting to notice that the pool of non-physicalist (such as panpsychists, panprotopsychist, as well as traditional forms of neutral monism which we discussed in chapters 2 and 3) also take this assumption as their starting point. Specifically, these anti-physicalists embrace the idea that the world is the way physics tells us at the fundamental level, and obviate the impossibility to explain consciousness with physical stuff by injecting the physical story with the extra ingredient (be it

[5] Although this may be modified to allow for the fundamentality of the whole, as we will see below.

some rudimentary form of consciousness for the panpsychist and Russellian Monists, or proto-consciousness for the panprotopsychists) which is had by these physical ultimates. From this, they then explain that macro-consciousness is in effect *inherited* by macro-bodies, such as ourselves, from the rudimentary consciousness or protoconsciousness of the micro-physical entities that constitute the macro-body in question. In order to produce a satisfactory explanation of consciousness, the panpsychist thus leaves the materialist paradigm untouched. By embracing and safeguarding the completeness of physics, she resorts to physics as the true revealer of ontology and injects her fundamental reality with what physics leaves out, namely consciousness itself. This leads the anti-physicalist to resort to the ontological hierarchy imposed by physics, and thus to rely on a sort of reductive analysis and with similar results to the physicalist, as can be seen by looking at the combination problem that challenges the panpsychist and which is nothing other than a bespoke revision of zombie argument (Chalmers (2016), as seen in chapter 2). The panpsychist case just shows how widespread in the consciousness literature it is to conceive reality in terms of ontological hierarchy, which in turn creates the need to resort to some form of reductive explanation. It also shows how problematic this approach reveals itself to be in the metaphysics of consciousness, even when the ontology is revised especially to tackle the problems associated with reduction and the Hard Problem.

6.2.1 Reversing the Hierarchy: Top-Down

As mentioned above, some conceive of the directionality of the ontological hierarchy as top-down rather than bottom-up. Schaffer's (2010) *Priority Monism*, for example, expresses grounding in terms of the fundamentality of whole: it is the whole, rather than the parts, that are ontologically prior. More specifically, the view is that there exists only one structured object in the world, the cosmos, and everything else is grounded by this whole. Therefore, the analysis of any existing object is to be understood in terms of its existence as a part of the whole it belongs to, and which makes the part in question intelligible. For instance, in order to intelligibly explain a part of the human body such as the heart, we need to look at what the heart does as a part of the whole body: the analysis will focus on explaining that the heart pumps blood to into the veins so as to allow the transportation of oxygen and nutrients throughout the body. This account focusses on the input of the heart to the whole. Contrasts this with a bottom-up explanation that would instead focus on explaining the heart in terms of its parts as the

ventricles of the aortic valve. In short, it is by looking at what the heart does in the body-whole, thus capturing how the heart is grounded in the body for its existence, that we can make the heart truly intelligible.

It is interesting to note that positing an ontological hierarchy grounded in the whole places an emphasis on the teleology of the objects being analysed, which contrasts with the more composition-based approach of the bottom-up strategy seen above. As a result, an intelligible explanation of the grounding relation changes the focus from how an object is constituted, as per the bottom-up view, to what the object constitutes. Constitution therefore continues to be the focus of the grounding relation in the top-down ontological hierarchy although the change in directionality effectively shifts the focus on the teleology of the parts in relation to the whole they are grounded in, as opposed to what an object is made of at the lower-level. Capsizing the fundamental level from the micro to the cosmos crucially leaves the status of physical knowledge unscathed, while allowing the philosopher to capture the teleology of the parts.

To place this approach within the consciousness project context, it is interesting to note that some panpsychist like Albahari (2019) and Goff (2017, 2019) adopt this top-down hierarchy. They call themselves cosmopsychists, they combine priority monism with panpsychism and argue that the consciousness that macro-level entities such as ourselves enjoy is in effect inherited from a larger cosmic consciousness (*consciousness+* as Goff (2019) calls it), as opposed to being inherited by the combination of lower level entities such as quarks. The inversion of the ontological hierarchy on the part of the cosmopsychist nonetheless runs into the trouble of having to explain how the cosmic-consciousness can decompose into lots of macro-consciousnesses. This approach leads to a capsized form of the zombie argument, leaving the internal mechanics of the reasoning the same.

Our brief treatment of priority monism, with the cosmopsychist case in point, helps us further highlight the overwhelming trust that philosophers lay in the ontological hierarchy system in their search for consciousness, as well as showing that the problem with the ontological hierarchy is not one that is specific to any directionality of the dependence relation,

which thus touches those committed to the top-down just as much as those committed to the bottom-up within this field.

6.3 Discovering Physical Constitution Vs Consciousness

The core of the matter here is that studying nature by structuring it in a hierarchical system has proved itself to be incredibly efficient for certain matters, but not for others. If we want to explain the phenomena of heat, we will refer to the motion of molecules. An explanation of gravity will focus on the mass of the gravitational objects in questions and how they interact with the more fundamental spacetime. Mechanism in this sense is adequately explained by reference to the fundamental entities, their properties and relations, which essentially constitute the entity or the phenomena being analysed. Similarly, in the field of chemistry we turn to the atomic composition to explain what an element is; for instance, we explain water as being composed of one hydrogen atom and two oxygen ones, thus allowing chemistry to adequately be reduced to physics. The successes within the scientific physical field are thus largely based on the coherence of the bottom-up hierarchical assumption (as well as their belonging to the same conceptual framework), aimed at discovering the *composition* and the *structures* that underlie the objects or processes in question. As such we may want to defend positing an ontological hierarchy for the purposes of explaining composition and structures as we do in science. The great success of the scientific project may appear to be, at the outset, a good enough motivation to transpose the use of this framework into the field of consciousness research. However, as we can see from the issues that the physicalist's reductive approach encounters, it is questionable whether transposing the ontological hierarchy to the study of consciousness is effective. In other words, hierarchy is a tool effective in science, but it does not seem to be as effective for producing an adequate explanation of consciousness.

Moreover, the suitability of the bottom-up approach within the third person perspective has been questioned with reference to the special sciences, such as biology or the social sciences, where the bottom-up reduction fails to adequately characterise the phenomena in question. John Dupre (1994), for example, argues that the notion of the unity of science, which we have seen with reference to the causal closure principle, comes under threat in the face of the failure to reduce certain aspects of biology to lower-level physical phenomena. He

advocates instead what he calls a *horizontal approach*, as opposed to the vertical imposed by the unity of science, to explain biological phenomena. As such, it seems that it does not suffice for a phenomena to belong to the third-person perspective in order for it to be coherently placed into an ontological hierarchy. This is a greatly interesting topic, but discussing it further would take us away from our focus here. However, it is important to note that consciousness is merely one amongst other phenomena of our reality that resists being placed within an ontological hierarchy. The question we can thus raise is whether the phenomena of consciousness may bear close (metaphysical) resemblances to biological phenomena such as life, which makes them both reduction-resistant. I have an intuition that this may be the case and that our model may find application within this context. The idea, loosely, is that the fact that both conscious and living organisms have an interiority may be the key relevant commonality and thus that our neutral framework, which focuses on a horizontal analysis of interiority and how phenomenal occurs within a neutral reality, could be tested with respect to these issues too. I will come back to this in the conclusions. With this note on conscious and biological systems, I do not want to imply that hierarchies should be thrown away from our metaphysical toolkit, as I have no intention to undermine the success of this approach within the physical sciences, at least amongst a good portion of them. However, I propose to stick a "handle with care" notice on the box. As such, I do not dismiss ontological hierarchies in the sciences, but neither do I commit to them. I merely leave the question open.

Note aside, the focal point of my discussion in this chapter is that beginning the analysis of consciousness by committing to a hierarchical organisation, as in the case of the physicalist, means to gerrymander the phenomena into a framework that is inhospitable for it on the grounds that it was created for different purposes, namely that of providing an objective (reductive) description of the structures of nature. That is to forcibly fit consciousness into a framework that was created to capture the world from the outward-directed objective perspective and that cannot, by definition, capture the subjective realm (following form our discussion in chapter 2 on the mind-body problem and in chapter 5 on the physical paradigm). In this case, I feel that the adoption of a hierarchy, whether top-down or bottom-up, has no reason to be transposed to an honest inquiry into the nature of consciousness. That is an inquiry that does not prioritise defending a different set of assumptions, namely the truths of physics, a project that is diametrically opposed to that of explaining consciousness. The success of the

reductive approach in the fields of physics and chemistry and its parallel failure to intelligibly explain consciousness are sufficient indications of this. To adopt a hierarchy for consciousness research, therefore, in my opinion, would merely mean to constrain our possibility to understand the nature of consciousness before we even begin to analyse it. If, however, our interest does not lie in defending the fundamentality of the truth of physics specifically, then we might want to try something different. I thus question why an explanation of consciousness should rely on the idea that (i) the micro-level[6] is ontologically primary, (ii) macro-level bodies depend on their micro-level composition, and (iii) in order to understand macro-level entities we have to look "down" at their lower level composition (i.e. the three points entailed by the grounding relation).

In the next chapter, therefore, I begin to think about consciousness aside from ontological hierarchies and, instead, focus on the subjects that display consciousness in and of themselves. The project concerns understanding the perspectives available to these subjects, what they are and how they arise, and therefore how the phenomena that is available in the two perspectives, the physical and the phenomenal, are coherently made sense of as being part of the same unified reality. Specifically, we work on the notion of a perspective understood as a neutral, rather than mental, state of affairs. The relationship between the two perspectives will be one of *mutual co-emergence* and *mutual co-definition*. In this sense, our aim is not to explain how the subjective perspective relates to the objective, but to show that they are one and the same thing, the subject of experience in the physical world compound. This is a very different task from explaining phenomena in the same perspectives, where this requires us to "denaturalise" one of the two perspectives by subordinating it to the other.

[6] I run the argument with a sole reference to the micro-level as fundamental for clarity and ease of exposition, but bear in mind that the same applies to the priority monist framework.

CHAPTER 7:

PERSPECTIVAL NEUTRAL MONISM

In this chapter we build the perspectival neural monist model. This is inspired by certain notions central to the enactive approach to cognition, such as how to define an organism in terms of local processes and global state, sense-making, interiority, the multi-dimensionality of the environment and the emergence of a meaningful environment as lived by the organism. These help us explain what a perspective is, why it is neutral and how phenomenal experience and the physical world can arise in a neutral reality. With this in place, I also discuss the perspectival neutral monist model in relation to reduction. Before we dive in, however, it is useful to summarise what we have found so far, to contextualise how this model fits in the debate and why enactivism is a helpful model for the neutral monist.

In the first part of this thesis we have found that the phenomenal/ physical distinction originates in modern philosophy; that positing these as the basis for ontology gives rise to the mind-body problem and that positing neutral reality has not solved it. In the second part of the thesis, we have found that we have two distinct sets of special concepts, phenomenal and physical, that each of these is attached to a distinct perspective, subjective and objective, such that these picks out neutral properties as encountered from the two perspectives. We have questioned the validity of the materialist paradigm as an intellectual paradigm and privileged description of reality, focussing on its philosophical shortcomings and on the omission of secondary qualities from the scheme. We have then committed to a neutral ontology on the basis of following the map from what we know, together with an acceptance of our limits, thereby extending our conception of neutrality to that of an infinite *multitude*. Lastly, I have challenged the ontological hierarchy assumption, and I have suggested that this assumption may not be adequate for investigating consciousness.

In the light of these findings, I feel that the enactive approach is particularly suitable for our purposes because it develops an account of cognition, traditionally construed as a mental phenomena, in terms that are not only mental but also physical. Although they do not deal with the question of ontology specifically, and do not discuss a neutral ontology (they generally refer to a physical type of ontology, if mentioned at all), a re-adaptation of their strategy is instrumental for fleshing out a conception of perspectives as neutral (rather than mental). In addition to this, the enactivist's analysis focuses on a non-hierarchical model, which perfectly suits our desire to discard the ontological hierarchy assumption for the purposes of tackling the question of consciousness and the physical world.

The view I develop in this chapter fits within this broader picture in that it reflects deeply on the existing debate and proceeds by *extending* and *building on* what has been done beforehand. Our assumptions have emerged gradually through a process of questioning, rather than being picked up at the outset, and this may put us in the position to more adequately face the challenges that other views face, such as the Conceivability Argument or the Explanatory Gap. The work I do in this chapter is my attempt to contribute to the debate by showing that a reconceptualisation of the problematic assumptions generally embraced in the debate is possible, and that a border logical space for investigating the nature of consciousness and its relationship to the physics world can be available to us if we look at the world with fresh eyes. Therefore, offering a coherent alternative to thinking about the problem of consciousness.

In the first section, I discuss the move from the centrality of constitution in the analysis of consciousness to the centrality of *interiority*. In the second section, I discuss the ideas I take from the enactivist approach that are instrumental for us to build our model. In the last section, I flesh out the model itself and discuss it in reference to reduction and the Hard Problem/ Conceivability Argument.

7.1 From Micro-Level Constitution to Interiority

In this section, I begin the shift from the widespread composition/ grounding framework starting point towards a different approach that begins by investigating the conscious beings (at

the level in which it exists) itself. As we will see, *interiority* is the central notion here. What motivates my interest in the notion of interiority is that the possession of consciousness is, in its most basic form, nothing other than the possession of a specific type of (putatively rich) *inner* life. If we define consciousness as there being *something it is like* to undergo a certain experience, and if this marks the existence of a subjective realm, then the possession of an interiority can serve us as a guide in our understanding of consciousness.

As we will see below, having an interiority is necessary yet not sufficient for consciousness. However, the analysis of interiority is a great tool for us to (i) to distinguish a conscious from a non-conscious being, (ii) to be able to start an analysis of consciousness from the conscious being itself, rather than by looking up or down at other levels of reality, and (iii) to give us a basis for explaining consciousness as continuous with nature, thus dispelling its mysterious origin.

Focusing on an analysis of bodies that display consciousness means that we should be able to draw the line between conscious and non-conscious beings. This distinction, however, is not as easy to draw as one may hope for. The question is how widely in nature can we find consciousness? Godfrey-Smith (2017), for example, argues that sea creatures such as octopi are "aliens" who display a level of consciousness that is not dissimilar from our own. Recent research by Galliano (2015) or the Minimal Consciousness Lab (for example Frazier, 2021; or Segundo-Ortin and Calvo, 2021) explore and defend the idea that plants have cognition and possibly even consciousness. Some thinkers such as panpsychist Goff (2017) and Hassl-Mørch (2018), for example, assume that some sort of rudimentary consciousness is had by physical ultimates like quarks and electrons, or by the neurones in our brain. Do octopus undergo experience, as in is there *something it is like* for an octopus to swim in cold rather than hot water? Is there *something it is like* for a bee to pick up pollen from a rose that differs for her from picking up pollen from a tulip? Is there *something it is like* for a plant to be starved off or to be over-fed? Is there *something it is like* for an electron to be attracted to a proton and be repelled by another electron? Most of us are prone to admit that certain bodies do not seem to be conscious, these will generally be (putatively) *inanimate* bodies like electrons or neutrons, stones or mountains. Similarly, most of us will want to agree that some other bodies are definitely conscious such as myself, you or mammals generally. Ad then, most of us will agree that some other bodies seem to be more conscious than others, these will be *animate* things such as an octopus who may

seem more conscious than a bacterium or an oak tree, for example. But can we know whether animate bodies such as bacteria, oak trees and octopi are truly conscious, such that there is *something it is like* to be them? The various views certainly show that it is difficult to draw precise lines around what is (definitely) conscious and what is (definitely) not conscious, especially along the animal and, in the light of recent findings such as Galliano (2015), the plant world. It may in effect be easier to draw a line between animate and inanimate bodies, or bodies that are definitely conscious such as ourselves and those that we deem inanimate on the common-sense view, such as quarks and electrons.

It is interesting, at this point, to go back to the ontological hierarchy approach and reflect on the fact that if we walk down the ontological ladder and look at these bodies at the micro-level, their composition will be very much alike: all are made out of the same components such as quarks, electrons, neutrons, which have the same properties such as charge and mass and are related in a relevantly similar manner. More specifically, while the ordering or structuring of ultimates in a stone, a tree or a person certainly differs (and crucially does so even between two instances of the same body such as between two stones, two trees or two people), the components themselves are the same types and the types of relations that hold between them are also the same. In this sense, on the micro-level of reality there exists no real difference between an octopus and a stone or a mountain and a person, if not for the different structuring of the composing elements. The crucial point is that the difference that holds between two inanimate objects such as a stone and a mountain at the micro-level is not lower in degree than the difference that holds between a mountain and an octopus or myself, namely between inanimate and animate or conscious bodies. 'Structurally, therefore, a human being or a dog is not different in kind from a toaster or a computer' (Griffin, 2007, p.77). This further shows how much of a challenge an analysis in terms of grounding can be, when we attempt to explain consciousness. This can also be taken to suggest why biological phenomena, alongside consciousness, may resist reduction to the physical, for it seems that an account of composition in terms of the relations of micro-physical entities cannot reveal the difference between something that is animate or inanimate, for instance, thus seeming inadequate to account for this adaptive process that distinguishes an organism from a car. I will return to this in the conclusion.

With this in hand, the best strategy is to begin drawing the line between conscious and non-conscious bodies by following our common sense. We can assume that objects such as stones and mountains are not conscious, a proposition that even the panpsychist would generally want to agree with (Goff, 2017), whereas some organic bodies such as humans and animals undergo experiences and are thus conscious. For now, we leave the rest (quarks, electrons, and especially animate organisms such as amoebas, bacteria, flowers, trees, insects and so on) aside. The first obvious difference between the two groups, then, is that bodies in the former do not display any sort of internal process while those in the latter do. In other words, while mountains and stones are passively subject to external environmental phenomena, for example the wind may roll the stone over and the rain erode the mountainside, humans and animals undergo a number of internal physiological (and phenomenal) processes and actively interact with their environment, such that my heart pumps blood into my veins and I reach out for food when I am hungry. This allows us to trace a distinction between non-conscious *inanimate* and conscious bodies, where the difference is based on the fact that conscious bodies present an interiority, an inner realm of processes that non-conscious inanimate bodies wholly lack. Interiority therefore becomes the starting point for investigating consciousness by focusing on the conscious body itself.

Of course, the notion of interiority just outlined is not sufficient for consciousness for organisms such as bacteria or plants, who may not be conscious after all, also display this sort of interiority. However, having an experience equals being in a specific state or process which obtains in the subjective realm, namely which can only be had by the subject in question and cannot be accessed by third parties (as opposed to physical knowledge). In this sense, we can understand consciousness broadly (for now, at least) as those events that occur in the internal subjective sphere, and can thus be associated with the possession of an inward-directed perspective. This definition is purposely very broad, however as we develop the argument, we will narrow down the definition of consciousness so as to coincide with the Nagelian (1074) definition, hence aligning ourselves with the terminology found in contemporary literature.

To sum this up, it appears that (micro-level) constitution does not help us reveal the difference between a conscious and a non-conscious body, where the distinction is very hard to draw in the light of recent research. Instead, it is at the level of the entity itself that we find the

first meaningful difference, that is the possession of an interiority seems to offer an efficient starting point to delineate a distinction between conscious and non-conscious (inanimate) bodies, and thus to begin thinking about consciousness. At the same time, this allows us to lay the grounds for our discussion about interiority and the possession of the inward-directed and the outward-directed perspectives, to which we now turn.

7.2 Enactivism as a Guide to the Subjective and the Objective Realms

In this section I unpack some elements of enactivism to further develop the notion of interiority hinted at in the previous section, and thus begin to understand how this leads to the formation of an inward-directed perspective and the relative outward-directed perspective. For the purposes of this section I tease out the issues that are most useful for our analysis of consciousness from the enactive approach, such as the interdependence of global and local processes (which is parallel with the discussion on ontological hierarchies in the previous chapter) and the mutual co-definition of internal and external worlds in conjunction with the process of sense-making (on which we model the notions of perspectives and the relative "construction" of the mental and the material). Please note that the enactivism debate unfolds at the intersection of understanding life and cognition, rather than consciousness. We get back to a discussion focussed on consciousness when we model neutral monism, thus in §7.3.

7.2.1 Autopoiesis, Organisms and the Boundary

Maturana and Varela's notion of autopoiesis (1973), especially as discussed by Thompson (2007) in his exposition of the relationship between mind and life, is helpful for delineating the kind of interiority important for building our model, thus how the possession of this interiority translates to the possession of subjective and objective perspectives.

The notion of autopoiesis was developed by Maturana and Varela (1973) to describe the organisation of the processes that characterise living systems, namely organisms. As Thompson illustrates (2007, p.98)

> the concept of autopoietic organisation arose as an attempt to abstract from the molecular processes of a cell the basic form or pattern that remains invariant through any kind of structural change as long as the cell holds together a *distinct entity* [my italics].

The recognition of this pattern was then generalised to define the pattern of the processes that all living systems or organisms undergo and which, in turn, defines each organism as a distinct and autonomous being. According to Maturana and Varela (1973), autopoiesis identifies a form of continuous self-production on the part of living systems, a feature that also helps us distinguish them from allopoietic systems such as ribosomes or heteropoietic systems such as cars or computers. On the basis of autopoiesis, Thompson (2007) calls a "minimal cell" one that has an organisation that is sufficient for its existence to be a distinct and independent cell. This is due to the fact that the cell has a semi-permeable membrane that acts as a boundary between itself and the rest of the environment. This membrane prevents 'the free diffusion between the cell and the environment' (Thompson, 2007, p. 98), while at the same time allowing the exchange of nutrients and energy between itself and its surroundings. 'Based on the nutrients entering from the outside, the cell sustains itself by a network of chemical transformations. (…) The metabolic network is able to regenerate its own components including those components that make up the membrane boundary' (Thompson, 2007, p. 98-99). The key elements of autopoiesis are thus three:

(i) the system must have a semi-permeable boundary
(ii) the boundary must be produced by a network of reactions that take place within the boundary
(iii) the network of reactions must include reactions that regenerate the components of the system (Thompson, 2007, p.101)

.

These three points are meant to illustrate the minimal requirements for life within a body, namely an organism. This is a first step to understanding the concept of interiority: we have a boundary and thus a set of processes that delineate an inner world from the external environment (a rudimentary form of identity).

The above analysis thus shows that we have a level occupied by the set of local metabolic and chemical processes that occur within the cell and that are responsible for the self-sustenance of the living system, including the production of the boundary, and a global level realised by the membrane which separates the local metabolic and chemical processes from the environment and thus subsumes them within the same body unit. This is a feature of the autopoietic model that is important for us to reflect on. The autopoietic process is crucially a process of *self*-production. That means that the local processes, namely the set of internal chemical processes, are in charge of producing the boundary of the body as well as sustaining its life. Yet, it is because the cell is a distinct entity, one that exists as a functioning whole and separate from its environment, that these processes occur the way they do in the first place. In the absence of the boundary the local processes would "disperse" in the environment, there would be no entity to sustain. Similarly, in the absence of the local chemical and metabolic processes the boundary would not exist. We therefore have a relation of interdependence between the global level and the local processes of the living system, whereby one cannot be said to be prior to the other. This fact is of vital importance in our development of a model to study consciousness, for it sheds light on why we might prefer not to think about consciousness in terms of ontological hierarchies. In other words, the interdependence of the local and the global levels is key to understanding why I prefer to study conscious beings in and of themselves first, rather than starting my analysis by picking either the micro or the cosmic levels as grounds, as per the traditional forms of mental and material monism. The point will become clearer as we continue our discussion, especially in the final sections of this chapter, but it is important to flag at this point.

7.2.2 Sense-making and Dynamic Co-Emergence of Inner and Outer worlds

On the basis of autopoiesis, Thompson, Varela, and Roch (1991) develop their theory of enactive cognition. A key element I want to focus on is the notion of sense-making, which is in turn responsible for the dynamic *co-emergence* of the internal and the external worlds. This discussion is the basis for understanding how the inward-directed perspective and the outward-directed perspective are generated (this will successively help us explain how the perspectives

relate to the neutral stuff, thus how the mental and the material can be adequately explained with reference to them in §7.4). We begin by focusing on a looser sense of perspectives based on a rudimentary version of sense-making, which occurs throughout the biological world as discussed by the enactivist, to come back to the more sophisticated and relevant one for consciousness below.

Sense-making in this context is understood as the process of creation by an organism of its own *Umwelt:* this is the environment as lived by the organism and thus imbued of significance by the organism that interacts with it. Note that this means that the *Umwelt* is different from the conjunction of facts that describe the environment in question from a purely third-personal perspective. Rather, the creation of the organism's *Umwelt* arises as a result of the interaction between the organism and the environment. More specifically, it depends on the particular processes necessary to ensure the organism's life as subsumed by the membrane and on the value that these consequently confer to the world that lays outside the boundary, such as to the nutrients found in the ground or the composition of the air. It is thus the processes+membrane compound that affords meaning to the external world and the external world in turn *emerges*, or is *brought forward* to use the enactive terminology, as a result of the *interaction* between the organism and the environment. It follows that different types of organisms may share the same environment, but each type will have its own *Umwelt*. For example, we may have organism A and organism B living in the same environment where both interact with molecule X. Molecule X may serve organism A as a nutrient, while carrying no significance for organism B. Following DiPaolo's (2009) characterisation, we can say that the organism effectively 'casts a veil of significance on the world [and] it *narrows down the multi-dimensionality* [my italics] of an environment into the distinction of whether the encounter is useful for survivorship and flourishing or not' (DiPaolo's, 2009, p.13). The external environment, in other words, *co-emerges* together with the internal specifications of the organism as the organism naturally *narrows down* environment into a world of significance. We can therefore say that it is the organism that shapes its reality as a result of its internal processes subsumed under the membrane and its interaction with the external world. This dynamic *co-emergence* of internal and external world, according to Thompson (2007), thus functions as a marker for organic life.

From this perspective, therefore, the body is not seen as a pure mechanism in the Cartesian sense, but as something that *intelligently interacts* with what is out there. In effect, it must be emphasised that this is an organic *adaptive* process which differs from purely mechanical processes- such as the evolution of the planets around the sun, or the attraction/ repulsion of subatomic particles, or the erosion of a mountain by the weather- whereby a set of strict laws govern the behaviour of entities and the interaction between them. DiPaolo (2009) puts it clearly when he notes that

> this is not a fact of the matter of the same kind as that describing a chemical reaction. It is a relational fact impossible to appreciate unless we have an organism present for which the effects of the chemical encounter on the processes of self-construction and self-distinction make sense as being good or bad. (DiPaolo, 2009, p.13)

Here the notion of the interaction being processed as good and bad is key to the adaptive process of an organism and thus to its autopoietic and autonomous existence. This is because organisms have an *active* attitude towards their environment. They interact with it in order to keep their homeostatic state to a level that enables their survivorship and possibly flourishing, thus to maintain their ability to self-produce and self-preserve. The active adaptive attitude towards the environment is meant in terms of the process of sense-making, whereby an organism perceives the environment and acts upon it in a specific manner, which in turn allows the attribution of meaning towards its *Umwelt*.

It is also different from the standard understanding of the relationship between organisms and the outer world where organisms are subject to fixed external stimulations which the body and the mind, in the case of humans for instance, simply process and ultimately attribute intellectual meaning to. The difference here is that *the process of cognising the external world is coextensive with the body*, through the membrane and the internal processes it literally yields a world of significance, rather than being understood as cognitive in the traditionally mental sense of the word. It is the (adaptive) action of the organism as a whole within the environment that is key in cognising and thus shaping reality (I will come back to this below). In this sense the external stimulations do not happen to organisms, but are picked out with

reference to the specifications of the living system and its necessity for self-preservation and sustenance as a bodily action.

We thus have the beginning of a meaningful, albeit minimal, perspective that an organism can hold. Specifically, it seems safe to claim that the organism holds a meaningful perspective on the external world, while also having an interiority which is constrained by the membrane and which is necessary for sense-making. Whether this interiority is sufficient for it to be called an inward-directed perspective I will remain silent about, although I note that it is certainly necessary but not sufficient for consciousness. We have, thus, the inception of perspectives in terms of their co-emergence and mutual definition, which will be *key* for fleshing out perspectival neutral monism in § 7.4.

7.3 Modelling Perspectival Neutral Monism

With the relevant elements of enactivism in hand, we are now ready to begin fleshing out the perspectival neutral monist model. From this point onwards, therefore, we work within the neutral ontology we posited in chapter 5. As a brief reminder, we understand the neutral stuff in accordance to the *Neither* and the *Both* views taken together, as defined in chapter 3, in conjunction with the idea that the neutral stuff is a *multitude* which can only be wholly accounted for (or grasped from) a perspective-free position, following my reasoning on limits in chapter 5. The question for the perspectival neutral monist, thus, is: how do phenomenal experience and the physical world exist in this neutral reality? In other words, what does it mean to truthfully talk about the phenomenal and the physical within the context of her neutral ontology?

7.3.1 Identifying Subjects of Experience

We begin our own analysis on the level of the conscious body, at the level of the subject, as opposed to assuming either a top-down or a bottom-up hierarchy at the outset. We know, firstly, that for any subject to be conscious it must have some sort of *active* interiority, which distinguishes it from non-conscious beings. Second, that the interiority we are concerned with

must be an interiority of a certain kind, for we admit that not all beings with an interiority are necessarily conscious. For instance, a bacterium is a neutral body with an interiority, so are both a bee and a plant, however we may not want to claim these are conscious. The subject thus has to be identified and suitably distinguished from other organisms. The conscious subject can essentially be identified through the set of inner dynamics or processes together with their subsumption under a unified system, where the local and the global mutually define each other. To identify a subject of experience, therefore, we follow the same logic as the processes+boundary compound above, talking about inner *dynamics or processes* to refer to local states and about *subsumption* and *unity* to refer to the global state. It is essential to remind ourselves that we are operating within a neutral reality and, as such, we now no longer think about local processes and dynamics or about the global state and unity as physical in nature. Here we have to step into a fresh eyes position.

What I mean by "processes" and "dynamics", in the context of perspectival neutral monism, are all those states that occur within the subject, locally, thus within the bounds of the unity of the system. These processes can be captured by physical talk, such as by telling the causal/ functional story of some neural activity such as C-fibres firing, but they can also be captured by talk of some specific perceptual experience, such as feeling pain. The key is that there are a number of these processes occurring at the same time, as we know them to. Right now you are in a series of distinguishable phenomenal and physical states, for instance you see the black and white of the text on the page, you hear sounds coming from the other room, maybe you have the taste of tea or coffee in your mouth while your neurones fire, your heart pumps blood, your lungs breathe, and so on. These are all local processes and dynamics that capture, through phenomenal and physical vocabulary, the set of distinguishable states that occur locally within our system as conscious beings.

Similarly, what I mean by "subsumption" and "unity" is the state of all the processes existing cohesively, hence the global existence of all the local states unified as a single unit, which is distinguishable from other systems or from the environment, and that can be cashed out both as physical and as phenomenal. For instance, we can tell the story of subsumption and unity from a third-person perspective when we talk about self-sustenance, much along the lines of the more basic organisms, but we can also talk about subsumption from the first-person

perspective when we talk about the unity of experience, that cohesive phenomenal state that defines *what it is like* to be myself, right now. Processes+unity can thus be cashed out in mental and material terms, depending on the perspective one takes as we will see below, but are essentially neutral.

I want to note that I spent countless hours trying to find some terminology that carries no strong phenomenal or physical connotations in our common understanding, such that I could give a satisfactory description of these subjects in a neutral world, descriptions that would not immediately make us think of things we generally conceive as physical and phenomenal. However, it is the very nature of our vocabulary to be "stuck" in mental and material talk (talk that we classify as referring to either mental or material things). For this reason, the reader will have to bear with me until better words can be crafted to describe the neutral stuff, if possible at all (which I very much doubt given the nature of our vocabulary is based on our perspectives). Because of the very strong physical connotations, for example, I decided not to transpose talk of boundary into our definition of the neutral subject. Instead, I prefer the simpler term unity, leaving its conceptualisation more open ended. On the other hand, I decided to keep the notion of processes because a subject like you or myself do undergo a number of processes that we characterise as mental or as material, as seen above. I paired processes with the notion of dynamics for the same reasons.

DiPaolo (2021, p.799) notes that

> Two common prejudices must be abandoned to make proper sense of this idea. The first prejudice is to see bodies only in terms of anatomy and physiology. Our bodies are indeed organic, but we also have daily and phenomenologically informed reasons to conceive of them as bundles of activity, or as William James put it "bundles of habits": what they do, what they express, what they care about, their powers and sensitivities. In a non-dualistic perspective these active, minded aspects are not separate from the constitutive conditions that make a body the body it is.

For ease of exposition, from now on, I will refer to processes and dynamics as unified under a single whole as *processes+unity* or as a *neutral pattern*. I like the notion of neutral pattern because it can help us free ourselves from the hegemony of terminology that is mentally or materially characterised, thus allowing us to think in more neutral terms. The subject is, thus, best understood as the embodiment of this processes+unity, hence the embodiment of this neutral pattern.

7.3.2 Conscious Bodies and Hierarchies

Before moving onto perspectives, I want to contextualise my reasoning here with reference to the discussion of ontological hierarchies in chapter 6.

I have identified and analysed the conscious subject with reference to processes/ dynamics and subsuming/ unity, where these are taken to mutually define each other. That is, the processes can be said to define the unity-whole just as much as the unity defines the local processes[7]. We can therefore analyse the subject either as processes and dynamics first or as unity-whole first, applying a top-down or a bottom-up hierarchical structure for explanatory purposes, where these would be functional analyses aimed at obtaining a specific set of information or explanation about the details of the subject/ system. In other words, we are free to analyse the subject in hierarchical terms, free to pick which level is prior in our explanation, but this does not pick out any ontological fact about the priority of either the whole or its parts. The key notion in this context is the *mutual co-definition* of local and global states. Mutual co-definition captures the fact that the features of the system we pick out in our analysis *as* local and *as* global are in effect interdependent and interwoven, such that they cannot be teased apart from each other in reality. This way, although we could in principle infer a levelled structure, it seems that the inference to the existence of an ontological hierarchy is arduous, if desirable at all. Contrast this with bottom-up grounding: a subject (and her consciousness) is wholly grounded in the micro-level entities that compose her, such that the micro-level composition of the subject (and her consciousness) are fully responsible for the obtaining of the

[7] I recognise that this may raise the question: how? And that so much more needs to be said about the details of this, but this should only be the case in the event that the model is coherent and viable. It is thus a project that we will not undertake here, as we are concerned with understanding whether the model might viable at all.

subject (and her consciousness). Or with top-down grounding: the processes are wholly grounded in the subject (and her consciousness), such that their obtaining is fully depended on the (prior) existence of the subject (and her consciousness). The theoretical apparatus is just different in the perspectival neutral monism framework because, given mutual co-definition, it is unintelligible how one level of reality could be prior or more fundamental than the other.

Bear in mind that, right now, we are operating within a neutral reality and in the context of subjects of experience specifically, where it does truly make sense to talk about local and global processes standing in a *mutual co-definition* state. As such, this discussion is not meant to impede on the validity of ontological hierarchies within a single perspective such as the physical, which is a question I want to leave open. That is, it may well be that there are ontological hierarchies such as the ones we found when talking about water/ H2O and the heat/ movement of molecules, or as in the case of inanimate objects for instance. My discussion here is merely geared at showing that this is not the case for subjects of experience (and probably for biological organisms too given their analogous resistance to reduction to the physical, as we will see in the conclusion), and that within a neutral world we can coherently take our analysis from the vertical to the horizontal. Recall that my aim with this model is to attempt to move beyond some pre-assumed (and arguably arbitrary) distinctions that populate the literature on consciousness, and the above is a case in point: I do not work my view around the belief that that consciousness is to be inherited or otherwise derived from entities at higher or lower levels of reality. Hierarchies for us are thus great conceptual tools to tease out specific information about our subject/ system, but play no role in our ontology. At least not so far.

7.3.3 Putting Limits to Work: Narrowing Down the Neutral Multitude

The next step is to rework the enactivist's sense-making to flesh out perspectives, thus to understand how the phenomenal and the physical can be said to coherently exist in the neutral world.

According to the enactivist, the organism *brings forth* its world of significance, its *Umwelt*, by '*narrowing down the multi-dimensionality* of the environment [my italics]' (DiPaolo, 2009, p. 13), on the basis of its processes+membrane compound, and its relative interactions with the

world. This process of sense-making, therefore, marks the co-emergence of a narrowed-down external environment that is simultaneous with and complementary to the defined interiority of the organism.

This is precisely how the perspectives arise for our subject in the neutral world. The subject brings forth her reality by narrowing-down the multi-dimensionality (I will also refer to this as the *multitude*) of the neutral stuff. This occurs on the basis of the subject's processes+unity, thus on the basis of the pattern she embodies, and the associated possible interactions with her environment. The processes+unity pattern, therefore, affords us a specific interiority and a specific external world, simultaneously. In the case of subjects of experience such as ourselves, this marks the availability of two complementary perspectives: one inward-directed and the other outward-directed. The former captures reality subjectively and via phenomenal concepts, while the latter captures reality objectively and via physical concepts. It needs to be noted that sense-making here, thus occupying perspectives, is not a mental state but one that involves the organism (and her interactions with the environment) as a whole. And this is key to understanding both the notion of perspectives and the concepts associated with them. It is the processes+unity, the inherently neutral pattern, as a whole that does the sense-making, such that perspectives are neutral in nature. In other words, capturing the world on the part of the neutral pattern/ subject is something that can be cashed out experientially just as much as it can be cashed out physically, albeit being wholly neutral in nature. We are, in effect, moving towards a conception of the subject as truly embodied, where what we call the body is exactly the same locus of what we call the mind, and where mind and body can be understood as sharing the same nature. And I feel that this is one of the perks of perspectival neutral monism: making a conception of embodiment smoother. I will come back to this discussion in more details in the conclusion.

Recall that the enactive organism's processes+boundary ensures that it brings forth a world of significance, an *Umwelt*, such that the organism's encounter with property X corresponds to nutrition whereas the encounter of a different organism (embodying a relevantly different process+boundary) would afford X a different significance. The case of a subject in the neutral world follows the same lines, where the world of significance human beings find in the outward perspective is the physical world. The perspectival neutral monist thus pushes this line

of reasoning further and pairs it with the inward-directed perspective. This works as follows: an organism, being that specific pattern (processes+unity), has an interiority, broadly, whereby this interiority just is the phenomenal. In other words, the pattern she embodies corresponds to her ability to occupy an inward-directed subjective perspective to reality, such that when she occupies an inward-directed perspective with respects to neutral property P she feels pain, for instance. Similarly, on the basis of the pattern she embodies, when she occupies an outward-directed perspective onto neutral property P she can give a casual-functional explanation of it. The specific pattern we embody, therefore, is just the narrowing-down of the multi-dimensionality of the neutral world into a subject that is conscious and inhabits the physical world. In the same key, relevantly different patterns will narrow-down the world in relevantly differently manners and will thus have relevantly different perspectives. Consider the different types of interiority and enacted worlds that obtain for relevantly different organisms, from humans to dogs, octopi, bees, bacteria and so on, and also inanimate non-conscious entities such as atoms whereby we have a total lack of perspective (putatively). The greater difference in pattern corresponds to a greater difference in interiorises and enacted worlds, and thus in the perspectives afforded.

The *narrowing-down* here is essential. The core of the narrowing-down is the idea that the subject is inherently *limited*. That is, on the basis that she embodies a specific processes+unity/ neutral pattern we can say that she is not the multitude, but merely *that* pattern. This, in turn, enables and constrains the scope of a subject's existence in the neutral world: the actions she can carry out, the types of processes she can undergo, the range of emotions she can feel, the scope of the mind-independent world she can inhabit, etc. That is, it constrains the scope of the two perspectives into all that phenomena that we dub mental and material. The pattern's limitations, for instance, present the world to us in three spatial dimensions, while the neutral stuff might ultimately be multi-dimensional spatially speaking. The 3-Dness of a subject's existence such as our own, moreover, is not something over and above the neutral stuff, neither is it separate nor arises from it. It is merely all that is available to the subject and characterised as three dimensional by means of the perspectives, mental and material concepts, which just cannot capture the multi-dimensionality of the neutral stuff as a whole because their nature is limited. It is thus on the basis of these limitations that the

narrowing-down of the neutral *multitude* results into the specific phenomena that is our existence as experiencers in the physical world.

A loose analogy that can be useful to further unpack what we mean by narrowing-down the multitude of the neutral is with the determinable/ determinate relation. The determinable is a certain type of property that can be further determined by instances of that determinable property. For instance, the property of being shaped or being coloured are determinable, while the properties of being square or circular are determinate instances of being shaped and the properties of being red or green are determinate instances of being coloured. Broadly, the neutral stuff is the determinable like shape or colour, and the determinate is the processes+unity, the pattern, like square or red.

To sum this up, in our perspectival neutral monist model consciousness is simply the narrowed-down, limited neutral reality *available* to us from the inward-directed, first person perspective. Similarly, the physical world is simply the narrowed-down neutrality *available* to us in the outward directed, third person perspective. These perspectives, the processes+unity, is the limiting of reality and this is coextensive with what we describe as consciousness or as the physiological states of the body. In other words, the unified processes that a subject *is* affords reality in the two perspectives, what and how much is available to us of the multitude of the neutrality, thus allowing us to bring forth our world of significance: our existence as experiencers in the physical world. A perspective is therefore not understood as a way to look at reality, but rather reflects an ontological fact about of the pattern and its interaction with the neutral reality. This is why the notion of perspective is truly neutral, as opposed to mental when understood in the common sense way as an epistemic position.

7.4 Perspectival Neutral Monism and Reduction

The question now is whether this gives rise to distinct, phenomenal and physical properties and then whether these properties are somehow reducible to the neutral base. The two questions are inter-related in that if our model gives rise to distinct mental and material properties then we can ask whether there is reduction, whereas if the model does not result in the obtaining of distinct properties then the question about reduction cannot be raised.

I see the answer to the first question as follows. Within the world that the perspectival neutral monist presents us, properties are all of one kind and this kind is neutral. The narrowing-down (correlated with the neutral pattern in question) does not carve up reality, in the sense that it does not pick out breaks, divisions or distinctions within the neutral, nor does it produce a sort of second-order reality (as in the case of the mental for the physicalist or of macro-consciousness for the panpsychist). Recall what said in Chapter 4 about phenomenal and physical concepts picking out properties *as encountered* in the two distinct perspectives. Then, talk of the mental and the material is merely a matter of perspectives and of the inherent limits these have in capturing the more expansive neutral reality. Therefore, the fact that we can characterise neutral properties via phenomenal and physical concepts simply provides us with information about our limited nature: that we are the embodiment of a specific neutral pattern, that we can stand in two distinct perspectives with respects to reality, and that these two perspectives are inward-directed and subjective and outward-directed and objective. Yet, this tells us nothing (or very little) about the deep nature of reality or the existence of distinct properties. In other words, the differently oriented perspectives capture reality as phenomenal and as physical, and via phenomenal or physical concepts, were these contrasts with capturing reality from the perspective-free position in that from the perspective-free position the world is still a uniform neutral *multitude*, populated only by neutral properties. This may be hard for us to truly to grasp because we are not eternal beings, and only the hypothetical eternal being could really comprehend the real nature of the uniform neutral *multitude*.

Nonetheless, the nature of a property is generally taken to be that it confers character onto a substance, but in the metaphysical picture we are painting the stuff/ substance retains its neutral character, ontologically speaking, as captured from the perspective free position and where talk of mental and material informs us about what perspective we occupy, rather than about the character of reality. The perspectives are thus informative about our own limits in acknowledging, encountering, interacting, processing or thinking about reality; and while we can capture and enjoy reality as mental or as material, this does not entail that there are distinct mental and material properties. In addition to this, recall that our analysis of conscious beings has taken place without the conception that the world is hierarchically ordered, focussing

instead on facts at one single level of reality, this makes it harder to understand where these properties would exist in our world.

The features outlined above constitute the key differences between the perspectival neutral monist and the other monist views around, according to which the existence of physical and phenomenal properties are expressly postulated as being instantiated at various levels of reality. The traditional neutral monist, such as Russell (1921, 1927) for instance, tells us that mental and material properties obtain as different arrangements of neutral entities at different levels of reality, such that a certain arrangement of neutral entities at the fundamental level yields mental properties at the macro-level and a different arrangement at the fundamental level yields physical properties at the macro-level. They thus adopt a layered picture of reality and populate it with properties at more or less fundamental levels. From this, the traditional neutral monist can reduce mental and material properties to the adequately arranged neutral base. The same sort of theory structure can be found in the physicalist and panpsychist views we have exposed previously. These theories present, first, a structured picture of reality and, second, the instantiation of different properties at the two levels whereby less fundamental properties are wholly defined by, and to a certain extent are thus isomorphic to, their constitution at the more fundamental level.

According to Heil this sort reasoning is due to an 'uncritical acceptance of a certain conception of language to the world' (Heil, 2003, p.207). He summarises it as follows:

(i) concepts pick out specific properties[8]
(ii) some type of concepts P (like phenomenal concepts) can be 'analysed, paraphrased or completely decomposed' (Heil, 20033, p. 216) with reference to other concepts F (like physical concepts), such that
(iii) if Ps cannot be fully analysed as Fs, then Ps are either distinct or do not exist

This assumed special relation of language to reality suggests that we 'can "read off" features of reality from our ways of speaking about it,' according to Heil (2003, p.207), thus imbuing validity

[8] Heil calls this Picture Theory and leaves it open to the reader whether, and eventually how, this might be related to Wittgenstein's own in the *Tractatus* (1921).

into a hierarchically ordered reality given by the fact that some concepts can be analysed in terms of others, as well as to the assumption that having a concept entails designating properties (I am simplifying the exposition here for ease of argument). But Heil is clear in stating that 'concepts do not carve up the world' (2003, p.216). This idea is not new in this thesis, in chapter 5 I readapted Dyke's (2012) *Language Thesis* which also criticises this approach and I have also discussed Heil's (2013) idea that two concepts can be made true by the same truth-maker, whereby possessing two concepts does not entail the existence of different properties.

In my opinion, it is true that we can read off features of reality from our way of speaking about it. For instance, in the context of a neutral ontology when we talk about the mental and the material we read off that we have two perspectives we can occupy with respect to reality and these perspectives allow us to characterise reality differently in terms of two distinct conceptual schemes. However, this does not entail that each concept reads off of reality that there are mental and material properties.

McGinn (2004, p.41), on the other hand, supports the type of approach criticised by Heil (2003) and Dyke (2012)

> two concepts C1 and C2 differ in their content if and only if there are properties P1 and P2 such that C1 connotes P1 and C2 connotes P2 and P1 is not identical to P2. One motivation for this is that a concept is a way of thinking of something, and a way of thinking of something is a way something is taken to be, in other words a property it is taken to have. Thus if two ways of thinking differ, then they differ in how the world is taken to be, in the properties that are taken to be instantiated.

I understand that in the context of a physicalist ontology this sort of reasoning is necessary to defend the coherence of the metaphysics, for if the relation between a concept and a property was not of the type desired by McGinn (2004) we would have a proliferation of mental entities that find no counterpart in the physical world, thus undermining the explanatory power of positing a physical fundamental reality. However, in the context of the perspectival

neutral monism model we know that different concepts pick out the different perspectives that an agent can occupy with respect to reality, rather than difference in how objects are, beyond the perspectives, ontologically speaking. As such different concepts do not correspond to a distinction in properties, but merely to different ways the property becomes available to us in the two perspectives.

In the perspectival neutral monist model, therefore, suitably related phenomenal and physical concepts pick out the same neutral properties, and do not refer to distinct properties that may for example exist at different levels. Rather, they refer to neutral properties as available in one or another perspective. With this in hand we can therefore answer the first question raised at the beginning of the section, namely whether the perspectives yield mental and material properties. The answer can be negative if we leave aside the methodology about the relation of language to the world defended by McGinn (2004) and, instead, decide to follow the route carved by Heil (2003). In this case, then the second question about reduction posed above simply does not arise, on the grounds that there are no properties to reduce.

7.4.1 Taking up the Critics' Position: Possible Issues of Conceivability

To further dig into the question about properties and reduction, I tried designing a perspectival neutral monist zombie. If I could design a neutral zombie, maybe I could revise my proposition that phenomenal and physical concepts effectively pick out non-neutral properties. However, I could not. In the case of the physicalist, for example, we can design a zombie precisely because she expressly posits a layered world, then adopts a reductive strategy that moves from the fundamentality of physics to why it is the case that we have experience and how this is not causally redundant. The physicalist thus expressly posits a hierarchical structure and two sets of properties that obtain at different levels, where one depends on the other for its existence. A similar consideration can be made with reference to the panpsychist for example. But how could this be applied to the perspectival neutral monist? I simply cannot construct a perspectival neutral monist zombie on the grounds that there are no higher or lower levels, but only the neutral unity+processes which *mutually co-define* each other. That is, though within the neutral landscape we can analyse the subject top-down or bottom-up, this does not reveal an ontological hierarchy. It reveals organisation and structuring, but not a hierarchy understood in

terms of priority and dependence. Moreover, we posit only neutral properties and that consciousness and the physical world is nothing other than the neutral multitude as encountered from different perspectives. Our premisses are thus very different from the ones that the Conceivability Argument challenges. This way, I feel the Conceivability Argument cannot be posed to the perspectival neutral monist. What would a perspectival neutral zombie be like? I cannot concoct one, but I am open to suggestions.

A critic may, however, prefer to adopt the Linguistic Strategy and somehow overthrow the metaphysical picture above, thus find levels of reality populated by neutral properties at lower level and phenomenal and physical properties at the higher level. Although I think that this would mean to misconstrue the view, I feel this challenge would not undermine the potential of the view. First of all, this is because I think that the source of the mind-body problem that attaches to reduction for the traditional monist is not generally based on the obtaining of distinct mental or material properties, but rather on the fact that it is unintelligible how a certain set of properties, such as the phenomenal, can arise from a metaphysical base that wholly lacks features that are conductible to them, such as the physical, as expressed clearly in the Explanatory Gap and Conceivability Argument. On the other hand, I feel that the notion of neutrality that the perspectival neutral monist introduces may help her overcome the challenge on the grounds that the neutral multitude, which we committed to by embracing the limited nature of our perspectives, can intelligibly give rise to any phenomena in the world when paired with the relevant neutral pattern.

A useful position that follows a similar line of reasoning is Stoljar's (2015) non- standard materialist response to the Conceivability Argument. His non-standard materialism is the view that all instantiated fundamental properties are either standard physical properties or non-standard physical properties. Standard material properties are those we have defined with reference to the theory based conception of the physical. Non-Standard material properties are either properties that are not available to us in our current physical theories, but that would be available in ideal or complete physical theories (the Nagelian version); or properties that are non-structural and non-disposition and not generally expresses by physical theories, but which ground those structural and dispositional properties that physics tells us about (the Russellian version). Non-Standard properties are thus best understood as properties which we can

hypothesise as existing, we can deduce their existence from all the facts we know about scientific theory, but which we just do not know what they are. He argues that this version of materialism is preferable to standard materialism on the grounds that it cannot be challenged by the Conceivability Argument.

> For suppose we formulate a third conceivability argument, similar to the two already on the table; according to this argument, it is conceivable, and so possible, that there is a possible world identical to the actual world in respect of fundamental standard and non-standard physical properties and, and yet different from it in terms of the nature and distribution of certain standard psychological properties, in particular those associated with phenomenal consciousness. This argument is unpersuasive and the reason is the one noted above, viz., that while we can describe these non- standard physical properties, we don't know what they are. Notoriously, the notion of 'conceivability' that is in operation in these arguments is epistemically demanding: you cannot conceive a situation in which various properties are instantiated unless you know, at least, in outline what those properties are. Since we don't know what they are can't conceive of them in the way required by a conceivability argument, (Stoljar, 2019, p. 225)

The response focuses on the first premise of the Conceivability Argument, thus undermining the very conceivability of zombies. In Stoljar's response, ignorance about non-standard material properties allows him to defend himself from the challenge posed by the Conceivability Argument on the grounds that if we don't know what these non-standard properties are then we cannot conceive of the scenario that underpins the argument. Ignorance thus plays a virtuous, rather than vicious, role as it enables Stoljar to defend himself from the challenge.

This reasoning is analogous to the perspectival neutral monist's, where Stoljar's ignorance about fundamental non-standard material properties translates to the reduced availability of the neutral multitude from the two perspectives. Moe specifically, the perspectival neutral monist can respond to the critic that finds a way to raise this challenge that the neutral

reality is only limitedly available to us from the two perspectives, thus we cannot reason the way the Conceivability Argument desires. In this sense, our limits become virtues of our theory in that they give us, like in Stoljar, further tools to defend ourselves from arguments that challenge a great portion of the views in the debate. Moreover, I feel that the notion of neutrality as a multitude may, in effect, be better suited for this sort of response to the Conceivability Argument than Stoljar's notion of non-standard properties, because it is not subject to the criticism as to whether the nature of non-standard material properties is truly material or not.

In this sense, the perspectival neutral monist also fares better than her neutral monist peers, because if we posit neutral entities such as those put forward by the *big three*, then we would know sufficiently about these properties to conceive of scenarios the way the argument requires us.

With reference to the Conceivability Argument, therefore, it seems that our traditional neutral monist has a number of ways to respond. Firstly, if her premisses are accepted then it seems unlikely that an argument that targets the dependence of one well defined set of properties, such as the phenomenal, on another well defined set of properties, such as the physical, may be raised at all. This is on the grounds that the view wholly lacks the features that the argument targets. However, if a critic is to find fallacies in the reasoning that leads the perspectival neutral monist to her view, thus argue that levels of reality obtain and that these are populated by physical and phenomenal properties, then the perspectival neural monist can answer by following the track laid by Stoljar: that our limited nature constrains how much we know (and enjoy) about reality, and thus that we cannot exercise conceivability the way the argument requires us to.

7.5 Summary

In this chapter we have fleshed out the notion of neutral perspectives and described how the perspectival neutral monist makes sense of our true talk of experience and the physical world within her neutral ontology. We have done this by developing the idea that perspectives are nothing other than the processes+unity compound and their interaction with the environment, and thus that interiority and *Umwelt* co-emerge as the existence of the being itself.

We have also discussed that consciousness, thus that there is *something it is like* for me to be myself, and the physical world, thus that *Umwelt* I inhabit and which differs from that of other organisms, is nothing other than the narrowing-down of the multitude of the neutral stuff on the part of the limited subject, understood in terms of it being a neutral pattern. We have thus established the idea that perspectives are neutral because they involve the subject as a whole and its interactions with its environment, and that can be cashed out in physical or phenomenal talk, but that these phenomenal and physical concepts merely pick out neutral properties *as encountered* by the subject, rather than picking out physical or phenomenal entities. From this we have argued that physical and phenomenal properties do not obtain in our landscape. We have also tested the model with reference to how issues related to the Conceivability Argument, outlining that she seems to have the tools to overcome them. It follows from the above that the neutral monist can coherently explain the mental and the material within a neutral reality without resorting to reductive strategies or the addition of extra ingredients to her ontology, keeping the landscape relatively light, where the ontological picture is composed of only neutral entities and where the perspectives understood as neutral do the explanatory work.

CONCLUSIONS:

RESOLUTIONS AND POSSIBLE FUTURE APPLICATIONS

In this thesis I have developed a model I call perspectival neutral monism, with the aim of producing an alternative framework for thinking about the nature of experience and how it can be coherently woven with the scientific picture of reality (in a monist ontology). My strategy has been to approach the debate with a set of *fresh eyes*, questioning the basic notions and common assumptions adopted by popular views; particularly I have attempted to re-conceptualise and reframe them to carve a different path to approach the question of consciousness. The position I fleshed out in the last chapter, deeply inspired by the enactivist approach to cognition, seems to have the potential to adequately explain the reality of experience and the physical world within a neutral ontology. More specifically, by presenting perspectival neural monism, I hope to have offered a model for thinking about consciousness differently, to have shown that a set of fresh eyes can be cast upon a matter that is old and seems more frustrating than it should be, thus that taking different avenues for investigating consciousness are indeed possible if we are happy to question the status quo.

There are three main features in my model that I feel are particularly interesting and may positively contribute to the debate. The first, is the reflection on the limited nature of perspectives that I have employed to justify my commitment to a neutral ontology, and which complements the existing notion of neutrality. The second, is developing the notion of perspectives as neutral, rather than mental, as based on the enactive approach to cognition. The third, is the proposition to leave behind ontological hierarchies for the purposes of investigating consciousness. These elements, especially when taken together, form the basis of the explanatory potential this model seems to have, presenting a theoretical apparatus that is rather different from those that are currently available in the literature. We have seen an example

of this in our discussion on how the perspectival neutral monists deal with issues related to the Conceivability Argument.

These features also afford the perspectival neutral monist tools to overcome the mental monism challenge and the epistemic gap. I turn to these in § (i) and (ii) below. The model however, may also have some applications beyond the metaphysics of consciousness debate. The first relates to producing an interesting framework for thinking about issues related to the embodied mind and the second relates to issues connected to the resistance of biology to reduction. I turn to these in § (iii) and (iv) below. Of course these are only cursory overviews of the potential answers to the problems and the model would need to be further developed and tested, for that is the nature of a model. By the end of the conclusion, however, we should have a better grip on the potential of the view, how this may fare better than other views in the face of common challenges and whether it has the potential to depict a more uniform picture of reality, whereby consciousness is not mysterious but continuous with other biological phenomena in the world.

i. Perspectival Neutral Monism and the Mental Monism Challenge

We have introduced the mental monist challenge, which targets all existing version of neutral monism, in chapter 3. Generally, this is the challenge that neutral monism may be collapsed into a form of mental monism, on the grounds that it seems that the explanatory work is done by paradigmatically mental entities and thus that the notion of neutrality is largely explanatory redundant. If this is the case, then it would simply be more parsimonious to adopt some form of mental monism. The critic could specifically pose the challenge to the perspectival neutral monist on the grounds that perspectives are (or seem to be) mental states, characterised by mental properties, and thus that the nature of the physical is dependent on the mental.

The defence of the perspectival neutral monist relies on her updated notion of neutrality, whereby the neutral stuff is characterised as a multitude, hence along the lines of what (Seager, 2011, p.87) describes as a stuff 'possessing an infinite number of attributes,' thus something much greater than what is available to us in the two perspectives. This notion of neutrality is then used to complement the *Neither View* and the *Both View* that traditional neutral monists adopt.

From this, the perspectival neutral monist posits neutral entities to the nature of which is unknown to us, to a certain extent. This differs from positing neutral entities that have close ties with mental entities such as percepts, sensations/ elements and pure experience. In this sense, the neutral monist does not fall into the challenge the way the *big three* did, in the sense that there are no mental properties in our landscape but only neutral entities.

With reference to whether perspectives themselves are responsible for the existence of the physical realm the matter is more delicate, in that a perspective is generally understood to be a mental state. In this sense, the challenge resembles more the version that Heil (2010) is subject to. The central point of the perspectival neutral monist defence is that, as discussed in chapter 7, 'occupying a perspective' is not a mental, but a truly neutral, fact. This is because occupying a perspective involves the organism as a whole, it is dependent on the neutral pattern that a subject embodies, namely on all those processes and states that a subject embodies and that we *can* capture via phenomenal or physical concepts, depending on the perspective the agent occupies; where these concepts pick out neutral properties *as encountered* from the relevant perspective. If we admit that the ontology is neutral, and embrace the notion neutrality and of perspectives modelled on the enactive approach, then we can see how the perspectival neutral monist can easily get out of the challenge. This is because the physical realm, our *Umwelt*, arises as a result of the narrowing-down of the neutral multitude on the basis of the neutral pattern (processes+unity) it embodies and not on a mental activity. The same process explains the phenomenal. Therefore, there is no mental fact here that explains how the physical world obtains. This contrasts starkly with Heil's (2010) position that expressly places a mental process at the basis of the existence of physical properties.

From this, it seems that the perspectival neutral monist seems stronger than her neutral monist predecessors in that she seems to have the tools to counteract the mental monist challenge, it resits being collapsed into a form of mental monism. Therefore, on the grounds of her updated notion of neutrality in conjunction with that of perspectives as modelled on the enactivist approach, she preserves the explanatory potential that the neutral metaphysics is put in place to do.

ii. Perspectival Neutral Monism and the Explanatory Gap

The perspectival neutral monist can also deal with why there is such an incommensurable epistemic gap between the mental and the material. The core of the claim here is that we have two sets of concepts, the phenomenal ad the physical, and that each set is essentially based on a specific and distinct perspective. Moreover, each perspective is limited and directed either inward *or* outward. From the inward-directed perspectives we can *only* see reality subjectively, experientially, and we cannot have access to the world that the sciences tell us about. The same way, from the outward-directed perspectives we see reality objectively, as shared with our peers, yet we cannot access the subjective world of experience. The former excludes the latter and vice versa. Accordingly, each set of concepts is special precisely because it belongs to a distinct perspective, because it is obtained and used in special and peculiar ways, because it captures something about our existence as experiencers in the world that the other set of concepts, based on the distinct perspective, cannot capture. The incommensurable distinction we find between the mental and the material is, thus, merely a reflection of the fact that each set of concepts attaches to a distinct, limited and mutually exclusive perspective. The perspectival neutral monist here can be seen as supporting a conceptual dualism that reflects the duality in perspectives available to an agent such as ourselves, while also defending monist metaphysics whereby relevantly associated phenomenal and physical concepts pick out *bona fide* neutral properties as encountered from the two perspectives.

A question may then be posed about an explanatory gap between phenomenal and physical realms in relation to the neutral stuff. That would be the question of how the neutral stuff makes it intelligible that there should be experience and that the experience should feel the way it does, as associated with the neutral pattern in question. This can also be extended to the physical realm, thus why there should be a physical world, an *Umwelt*, associated with the neutral pattern in question. The inner mechanics of the perspectival neutral monist model developed in chapter 7 are a direct answer to this. Once again, briefly: we posit the neutral stuff as a multitude, something we can only know limitedly about from the two perspectives; we then explain that the neutral pattern that an agent embodies narrows-down this multitude of the neutral stuff into experience and the physical world (where relevantly different patterns are associated with relevantly different phenomena, see myself and you having different experiences, and then myself and a dolphin, for instance). I feel that this gives us a coherent

story as to why that specific neutral pattern should be associated with a specific experience and a specific *Umwelt*; this is on the grounds that a specific neutral pattern, based on its limits, narrows down the multitude of the neutral stuff into the specific experience and the specific physical world. We thus can trace a complete story of how a specific experience (or a physical Umwelt) arises based on the internal specifications of the process+boundaries and her relation to the environment as they occur within a neutral multitude.

Contrast this with the physicalist case where, according to Levine (1983), the full account of the physical-causal-functional structure, no matter how detailed, leaves us in the dark as to why that particular structure should feel, qualitatively speaking, the way it does. In the case of the perspectival neutral monist, on the other hand, the ontological picture is very different. The result is that when she gives us her story of why an experience should feel the way it does, based on the narrowing down of the neutral multitude on the part of the neutral pattern, it seems to make it intelligible why an experience feels the way it does (or a physical world is the way it is) associated with the neutral pattern in question. In this sense, I feel that, should a question about gaps arise for the perspectival neutral monist, she has better tools than her fellow physicalist to avert the challenge.

iii. Perspectival Neutral Monism and the Embodied Mind

The feature of the model I like that most is that it can actively bring the body back within the non-mechanical realm of the subject, back into the self, providing us with a good framework for conceiving of ourselves as truly *embodied minds*.

With Descartes we have seen that the body took up the connotations of a machine, with his *homme machin*e conception, and that the mind was conceived as wholly separate from it. The body is then relegated outside that I who merely *seems* to feel. We obtain a disembodied self. Although physicalists supposedly reject this strong separation and aim for a more embodied approach, I feel that their attempts may not have been enough to rescue the notion of the body from his machine-like status, and thus to properly understand how we are embodied minds. That is, even within physicalism we have a notion of the subject of experience that is fundamentally (reducible to) a set of physical facts, and thus to a body that for the most

part responds mechanically to reality. In other words, while the physicalist can adopt an embodied approach to mind and consciousness, the mechanical approach inherent in the physicalist ontology hinders the project of explaining how the mind really is embodied, rather than just subordinate, such that the mind still seems irreconcilably distinct from the body.

On the other hand, I feel our model is better suited to function as an ontological framework for thinking about ourselves as embodied minds. Specifically, I feel that the introduction of a neutral ontology as developed in this thesis can facilitate the move towards a conception of the embodied mind where the mind is inherently spread throughout the body and enactive, thus where mind and body exist *coextensively* in the active relation of the whole of the organism with its environment. The conception of the body in the perspectival neutral monist approach, as modelled upon the enactive approach to cognition, thus actively shifts from the body as *lived by a mind* to being a *living mind* itself (*Leib* and *Korper* in Merleau-Ponty, 1945/ 2012). The body as understood in this context in terms of a neutral patterns, thus becomes consciousness itself just as much as it is also understood as the physical body. This contrasts with the conception of a body that resembles purely mechanical processes, analogous to the evolution of the planets around the sun or the attraction/ repulsion of subatomic particles or the firing on neurones, whereby a set of strict laws govern the behaviour of entities and the interaction between them. On our reading, the neutral pattern can be said to i*ntelligently interact* with its environment as a whole and in the adaptive process available to it, on the grounds of interiority and sense-making, such that the neutral pattern just is a conscious body. As such I feel that theories that focus on embodiment and the embodied mind may benefit form assuming a metaphysical picture along the lines of what I have fleshed out in this thesis. Of course, whether this is so, it remains a case that needs to be tested in the future.

iv. Perspectival Neutral Monism, Biology and Life

I have mentioned a few times throughout the thesis that biological phenomena, like consciousness, resist being reduced to the physical. I have also raised the question as to whether the two phenomena may have something in common which renders them both reduction-resistant, and that if this is the case then perspectival neutral monism model may

contribute to this debate, at least by providing a framework for casting a set of fresh eyes on the matter.

The idea I have in mind is that biological phenomena and consciousness both revolve around entities that present an interiority and the associated co-emergence of their *Umwelts*, and this is what may make them reduction-resistant.

That mind and life have common features is not my own idea, rather this is the central matter in the life-mind debate. The life-mind thesis is the thesis that cognition may be a more sophisticated manifestation of life, such that the two exist on the same continuum (Godfrey-Smith, 1994; Wheeler, 1997; Thompson 2007; DiPaolo, 2021). Kirchhoff (2016) states that how to study the continuity of mind and life is still an unresolved issue, however, with contributors still working on refining the details. Moreover, note that the notion of "mind" in this debate is geared towards cognition specifically, rather than consciousness. However, as in chapter 7, the model easily translates to the case of consciousness.

The thesis is best defined with reference to Godfrey-Smith's (1994) strong continuity, which has been commonly adopted in the debate (for example Wheeler, 1997; Thompson, 2007; DeJesus, 2016), namely

> *Strong Continuity*: Life and mind have a common abstract patterns or set of basic organisational properties. The functional properties characteristic of mind are an enriched version of the functional properties that are fundamental to life in general. Mind is literally life-*like*. (Godfrey-Smith, 1994, p. 320)

The thesis unfolds over three points: (i) there are some basic (abstract) patterns or organisational principles which are necessary for life, or otherwise define what it is for an organic body to be alive, these are broadly those presented in the enactivist framework discussed in §7.2; (ii) that mind is a richer, more sophisticated, version of the basic organisational principles we find when we investigate life, (iii) that mind and life share the same characteristics, although with different degrees of richness. From this, we extrapolate the life-mind continuum which

organises living systems according to the richness of their patterns and organisational processes, from the most basic to the most sophisticated, and where mind (cognition) is understood in the rudimentary terms as form of basic (enactive) sense-making on the left end and the more sophisticated living systems, such as humans, who display the degree of mind (cognition) we are so familiar with on the right hand. In the above definition we take organisational patterns to constitute all those features that make up an organism, which in chapter 7 we have labelled processes+unity (or neutral patterns). Organisational patterns thus refer to organisms possessing a certain interiority which dynamically co-emerges with the Umwelt, as seen in the literature on autopoiesis and enactivism. It follows that biological organisms, thus all those organisms that display life are relevantly similar to conscious (human) beings.

Given these similarities, I feel that our model may provide an alternative framework for investigating biological phenomena, possibly more suitable than the traditional approach based on the physical sciences. More specifically, we have seen that a vertical analysis does not reveal anything salient about the possession of an interiority/ *Umwelt*, as it merely provides us with information about structure and constitution (as seen in chapter 7). Recall Griffin's (2007, p77) suggestion that structurally speaking a toaster is no different from a human being. Our model instead proposes that we carry out a horizontal analysis of those beings that have an interiority, and while we have discussed this issue with reference to conscious beings we do not transpose it to biological organisms more generally. From this I feel my model can provide a metaphysical framework that is more appropriate for investigating the nature of biological phenomena. This is based on the conjunction of the neutral notion understood as a multitude and the narrowing-down of the multitude on the part of a neutral processes+unity compound; specifically, we could reason that consciousness obtains as a more complex form of narrowing-down the multitude than the narrowing-down of the multitude on the part of simpler organisms which instead leads to life. Of course, I have no idea whether this is could be a successful endeavour, but I feel it deserves to be explored on the grounds of the similarities that consciousness and life seem to share. In short I feel that an approach to biological phenomena along the lines of the perspectival neutral monism model may open up the possibility to look at the matter differently, with a set of fresh eyes that see beyond the ontological hierarchy usually assumed and beyond

an ontological framework based on the adoption of the (often restrictive) scientific picture of reality.

Lastly, the life-mind thesis above also allows us to tackle another one of the aims I set out at the beginning of this thesis, namely my desire to paint a uniform picture of the world, where the phenomena of consciousness is not seen to arise mysteriously, but held to be continuous with nature at large. I feel that, in the light of the discussion above, thus placing our model alongside the life-mind thesis, this continuity can intelligibly be achieved on the grounds that consciousness is nothing other than a more complex version of life, whereby both phenomena result from the narrowing-down of the neutral multitude. Of course much more would need to be said with respects to this issue, this could be another PhD project, however I hope that the few words spent here have been enough to illustrate the direction that the perspectival neutral monist model, and her set of fresh eyes can find application in the future.

LIST OF REFERENCES

Albahari, M., 2019, Beyond Cosmopsychism and the Great I Am: How the World Might Be Grounded in Universal 'Advaitic' Consciousness, in Seager. W. (ed.), *The Routledge Handbook of Panpsychism, pp*. 119–130, New York, NY: Routledge

Alter, T., Nagasawa, Y., 2012, What is Russellian Monism? *Journal of Consciousness Studies* 19 (9-10), pp. 67–95

Aristotle, 2006, *On the Soul*, translated by Smith J. A., digireads.com (paperback)

Armstrong, D., M., 1968, *A Materialist Theory of the Mind*, London: Routledge and Kegan Paul

Árnadòttir, S.T., Crane, T., 2013, There is No Exclusion Problem, in: Gibb, S.C., Lowe, E.J., Ingthorsson, R.D. (Eds.), *Mental Causation and Ontology,* pp. 248–266, Oxford: Oxford University Press https://doi.org/10.1093/acprof:oso/9780199603770.003.0012

Atmanspacher, H. (2012). Dual-aspect monism à la Pauli and Jung, *Journal of Consciousness Studies, 19*(9-10), pp. 96–120

Baggs, E., Chemero, A., 2021, Radical embodiment in two directions. *Synthese* 198, pp. 2175–2190 https://doi.org/10.1007/s11229-018-02020-9

Balog, K., 2009, Phenomenal Concepts, in McLaughlin, B., Beckermann, A. & Walter S. (eds.), *The Oxford Handbook of Philosophy of Mind*, pp. 292-312*,* Oxford, UK: Oxford University Press

Banks, E. C., 2004, The Philosophical Roots of Ernst Mach's Economy of Thought, *Synthese* 139, pp. 23–53.

Banks, E. C., 2010, Neutral Monism Reconsidered, *Philosophical Psychology* 23 (2), pp. 173-187

Banks, E. C., 2014, *The Realistic Empiricism of Mach, James, and Russell: Neutral Monism Reconceived*, Cambridge: Cambridge University Press

Bailey, A., 1998, Supervenience and Physicalism, *Synthese* 117 (1), pp. 53-73

Ben-Yami, H., 2015, *Descartes' Philosophical Revolution: A Reassessment*, New York, NY: Palgrave Macmillan

Benovsky, J., 2018, *Mind and Matter: Panpsychism, Dual-Aspect Monism, and the Combination Problem*, Berlin, DE: Springer Verlag

Blackburn, S., 1992, 'Filling in Space', *Analysis* 52, pp. 60-65

Bohm, D., 1980, *Wholeness and the Implicate Order*, London, UK: Routledge

Broad, C. D., 1924, Critical and Speculative Philosophy published in (ed.) J. H. Muirhead *Contemporary British Philosophy: Personal Statements* (First series), pp. 77-100, London, UK: Allen and Unwin

Broad, C. D., 1925, *The Mind and its Place in Nature*, London, UK: Routledge

Burnyeat, M. F., 1982, Idealism and Greek Philosophy: What Descartes Saw and Berkeley Missed, *The Philosophical Review* 91, pp. 3–40 https://doi.org/10.2307/2184667

Chalmers, D.J., 1995, Facing up to the problem of consciousness, *Journal of Consciousness Studies* 2 (3), pp. 200-19

Chalmers, D. J., 1996, *The Conscious Mind: In Search of a Fundamental Theory*, Oxford, UK: Oxford University Press

Chalmers, D. J., 2002, Does Conceivability Entail Possibility? in T. Gendler and J. Hawthorne (ed.), *Conceivability and Possibility*, pp.145–200, Oxford, UK: Oxford University Press

Chalmers, D. J., 2003, Consciousness and its place in nature, in Stitch, S. P. & Warfield T. A. (eds.) Blackwell Guide to the Philosophy of Mind, pp. 102-142, Hoboken, NJ: Blackwell Publishing

Chalmers, D. J., 2007, Phenomenal Concepts and the Explanatory Gap, in Alter, T. & Walter, S. (eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism,* pp. 167-94, Oxford, UK: Oxford University Press

Chalmers, D. J., 2013, Panpsychism and Panprotopsychism, Amherst Lecture in Philosophy 8

Chalmers, D. J., 2016, The Combination Problem for Panpsychism, in Brüntrup, G., & Jaskolla, L. (eds.), *Panpsychism,* Oxford, UK: Oxford University Press

Carruthers, P., Veillet, B., 2007, The Phenomenal Concept Strategy, *Journal of Consciousness Studies* 14 (9-10), pp. 212–236

Cartaya, J., 2018, The Representational Fallacy and the Ontology of Time, *Dianoia: The Undergraduate Philosophy Journal of Boston College*

Coleman, S., 2013, The Real Combination Problem: Panpsychism, Micro-Subjects, and Emergence, Erkenntnis 79 (1), pp. 19-44

Coleman, S., 2016, Panpsychism and Neutral Monism: How to Make Up One's Mind, in Brüntrup, L.G.J. (Ed.), Panpsychism. Oxford University Press

Coleman, S., 2017, Neutral Monism, *Philosophy Now* 12, pp. 9-11

Coleman, S., 2018, The Evolution of Nagel's Panpsychism, *Klesis*: 41, pp.180-202

Crane, T., 1992, Mental Causation and Mental Reality, *Proceedings of the Aristotelian Society*, New Series 9, pp. 185–202

Crane, T., 1995, The Mental Causation Debate, *Aristotelian Society Supplementary,* Vol. 69 (Supplementary), pp. 211-36

Crane, T., 1999, The Mind Body Problem, in Wilson R., & Keil F. (eds.), *The MIT Encyclopedia of the Cognitive Sciences*, Cambridge, MA: MIT Press

Crane, T., 2000, Dualism, monism, physicalism, *Mind & Society* 1, pp. 73–85 https://doi.org/10.1007/BF02512314

Crane, T., 2005, Papineau on Phenomenal Concepts, *Philosophy and Phenomenological Research* 71 (1), pp.155-162

Crane, T., 2018, A short history of philosophical theories of consciousness in the 20th century, in: Kind, A. (Ed.), *Philosophy of Mind in the Twentieth and Twenty-First Centuries*, Series: *The history of the philosophy of mind,* Vol. 6, pp. 78–103, New York, NY: Routledge https://doi.org/10.4324/9780429508127-4

Davidson, D., 1970, Mental Events, in: Foster, L., Swanson, J.W. (eds.), Readings in Philosophy of Psychology, Vol. 1, pp. 207–22, Oxford, UK: Clarendon Press

De Jesus, P., 2016, Autopoietic Enactivism, Phenomenology and the Deep Continuity Between Life and Mind, *Phenomenology and the Cognitive Sciences* 15, pp. 265–289 https://doi.org/10.1007/s11097-015-9414-2

Descartes, R., 1645/1996, Meditations on First Philosophy, in John Cottingham (Ed.), *Meditations on First Philosophy with Selections from the Objections and Replies*, Cambridge, UK: Cambridge University Press

Descartes, R. and Bennett, J., 2009, *Correspondence between Descartes and Princess Elisabeth*. 1st ed. [ebook] Available at: <http://www.earlymoderntexts.com/assets/pdfs/descartes1643_1.pdf> [Accessed 6 January 2018]

Dennett, D., C. 1988, Quining Qualia, n: Marcel, A. & Bisiach, E. (eds.) *Consciousness in Modern Science*, Oxford, UK: Oxford University Press

Díaz-Leon, E., 2008, Defending the phenomenal concept strategy, *Australasian Journal of Philosophy* 86 (4), pp. 597 – 610

Díaz-Leon, E., 2010, Can Phenomenal Concepts Explain The Epistemic Gap?, *Mind* 119 (476), pp.933-951

Di Paolo, E., 2009, Extended Life, *Topoi* 28, pp. 9–21 https://doi.org/10.1007/s11245-008-9042-3

Di Paolo, E. A., 2021, Enactive becoming, *Phenomenology and the Cognitive Sciences* 20, 783–809 https://doi.org/10.1007/s11097-019-09654-1

Dyke, H., 2012, Metaphysics and the Representational Fallacy, Oxford, UK: Routledge https://doi.org/10.4324/9780203933718

Dupre, J., 1994, Against Scientific Imperialism, *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, pp. 374–381

Dewey, J., 1928, The Vanishing Subject in the Psychology of James Eddington, A. *The Nature of the Physical World*, Cambridge, UK: Cambridge University Press

Farkas, K., 2008, *The Subject's Point of View*, Oxford, UK: Oxford University Press

Farkas, K., 2019, The Boundaries of the Mind, in Amy Kind (ed.), *Philosophy of Mind in the Twentieth and Twenty-first Centuries,* pp. 256-279

Feigl, H., 1958, The 'Mental' and the 'Physical', *Minnesota Studies in the Philosophy of Science* 2, pp. 370-497

Feigl, H., 1975, Russell and Schlick: A Remarkable Agreement on a Monistic Solution of the Mind-Body Problem, *Erkenntnis*, Vol. 9, No.1 (May 1975) pp. 11–34 https://doi.org/10.1007/BF00223131

Fine, K., 2001, The Question of Realism, *Philosophers' Imprint*, 1, pp.1–30

Frazier, P. A., 2021, On the Possibility of Plant Consciousness: A View from Ecointeractivism, *Mind and Matter*, 19 (2) pp. 229-259

Fürst, M., 2014, A Dualist Account of Phenomenal Concepts, in Lavazza, A. & Robinson, H. (eds.), *Contemporary Dualism: A Defense*, pp. 112-135, London, UK: Routeledge

Gagliano, M., 2015, In a Green Frame of Mind: Perspectives on the Behavioural Ecology and Cognitive Nature of Plants, *AoB Plants*, Vol. 7, pp.1-8

Gherab-Martin, K.J., 2013, From Structuralism to Neutral Monism in Arthur S. Eddington's Philosophy of Physics, *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 44, pp. 500–512 https://doi.org/10.1016/j.shpsb.2013.09.006

Godfrey-Smith, P., 1994, Spencer and Dewey on Life and Mind, in Brooks, R. and Maes P. (eds.), *Artificial Life 4,* pp. 80-89, Cambridge, MA: MIT Press

Godfrey-Smith, P., 2017*, Minds: The Octopus and the Evolution of Intelligent Life,* London, UK: William Collins

Goff, P., 2006, Experiences Don't Sum, *Journal of Consciousness Studies*, 13(10–11), pp. 53–6

Goff, P., 2009a, In David Skrbina (ed.), *Mind That Abides: Panpsychism in the New Millennium,* pp. 129--135

Goff, P., 2009b, Can the Panpsychist Get Around the Combination Problem? in David Skrbina (ed.) *Mind that Abides: Panpsychism in the new Millennium,* pp. 129-135, Amsterdam, NL: John Benjamins

Goff, P., 2016, The Phenomenal Bonding Solution to the Combination Problem, in L. Jaskolla (ed.), *Panpsychism: Contemporary Perspectives,* pp. 283–302, Oxford, UK: Oxford University Press,

Goff, P., 2017, *Consciousness and Fundamental Reality*, New York, NY: Oxford University Press

Goff, P., 2019, *Galileo's Error*, New York, NY: Penguin Books

Goff, P., 2019, Cosmopsychism, Micropsychism and the Grounding Relation, in William Seager (ed.), T*he Routledge Handbook of Panpsychism,* pp.144– 156, New York, NY: Routledge

Griffin, D., R., 1997, Panexperiential physicalism and the mind-body problem, Journal of Consciousness Studies 4 (3), pp. 248-68

Griffin, D. R, 2007, *Whitehead's Radically Different Postmodern Philosophy An Argument for Its Contemporary Relevance*, Albany, NY: Sate University of New York Press

Heil, J., 2003, Levels of reality, *Ratio* 16 (3) pp. 205–221

Heil, J., 2010, Powerful qualities, in Anna Marmodoro (ed.), *The Metaphysics of Powers: Their Grounding and Their Manifestations,* New York, NY: Routledge

Heil, J., 2013, P*hilosophy of Mind: A Contemporary Introduction* (3rd edition), New York, NY: Routledge

Holbrook, D, 1992, Descartes on mind-body interaction, *Southwest Philosophical Studies* 14, pp. 74-83

Horgan, T., Potrč, M., 2012, Existence monism trumps priority monism, in Goff, P. (ed.), *Spinoza on Monism,* pp. 51–76, New York: Palgrave-Macmillan

James, S., 2000, The Emergence of the Cartesian Mind, in Crane, T. and Patterson, S. (eds.), *History of the Mind-Body Problem*, pp.111-130, London, UK: Imprint Routledge

James, W., 1890, *The Principles of Psychology*, Vol. 1 and 2, New York, NY: Dover

James, W., 1907, *Pragmatism: A New Name for Some Old Ways of Thinking*, New York: Dover Thrift Editions

James, W., 1912, *Essays in Radical Empiricism*, Reprinted: Lincoln, US: University of Nebraska Press (1996)

Lowe, E. J., 2002, *A Survey of Metaphysics*, Oxford, UK: Oxford University Press

Lewis, D. K., 1966, An Argument for the Identity Theory, *Journal of Philosophy,* 63(1) pp. 17–25 https://doi.org/10.2307/2024524

Lewis, D. K., 1972, Psychophysical and Theoretical Identifications, *Australasian Journal of Philosophy* 50(3) pp. 249–58

Mach, E., 1886, *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen,* 5th edition translated as T*he Analysis of Sensations and the Relation of Physical to the Psychical,* 1959, NY: Dover

Mindt, G., 2019, Doctoral dissertation, available at https://philpapers.org/archive/MINRUU.pdf

Frank, J., 1982, Epiphenomenal Qualia, *Philosophical Quarterly* 32 (April), pp. 127-136

Kern, A., 2020, Life and Mind: Varieties of Neo-Aristotelianism: Naive, Sophisticated, Hegelian. Hegel Bulletin 41, pp. 40–60 https://doi.org/10.1017/hgl.2019.11

Kim, J., 1979, Causality, Identity, and Supervenience in the Mind-Body Problem, Philosophy Documentation Center, *Midwest Studies in Philosophy* 4, pp. 31–49 https://doi.org/10.1111/j.1475-4975.1979.tb00372.x

Kim, J., 1982, Psychophysical supervenience, *Philosophical Studies* 41 (January), pp. 51-70

Kim, J., 1990, Supervenience as A Philosophical Concept. *Metaphilosophy* 21, pp. 1–27 https://doi.org/10.1111/j.1467-9973.1990.tb00830.x

Kim, J., 1993, Multiple Realization and the Metaphysics of Reduction, in: *Supervenience and Mind, selected philosophical Essays,* pp. 309-335, Cambridge, UK: Cambridge University Press

Kim, J., 1997, Does the Problem of Mental Causation Generalize?, *Proceedings of the Aristotelian Society* New Series, Vol. 9, pp. 281-297

Kim, J., 2005, Supervenience, Emergence, Realization, Reduction, *The Oxford Handbook of Metaphysics,* New York, NY: Routledge https://doi.org/10.1093/oxfordhb/9780199284221.003.0019

Jackson, F., 1982, Epiphenomenal Qualia, *Philosophical Quarterly*, 32, pp. 127–136

Kirchhoff, M.D., 2018, Autopoiesis, Free Energy, and the Life–mind Continuity Thesis, Synthese 195 (6), pp.2519–2540 https://doi.org/10.1007/s11229-016-1100-6

Kriegel, U., 2003, The New Mysterianism and the Thesis of Cognitive Closure, *Acta Analytica* 18, pp. 177–191 https://doi.org/10.1007/s12136-003-1020-1

Kripke, S. A. 1980, Naming and Necessity, Cambridge, MA: Harvard University Press

Ladyman, J., Collier, J. G., Ross, D., Spurrett, D., 2014, *Everything must go: metaphysics naturalized, Oxford, UK:* Oxford University Press

Langton, R., 2004, Elusive Knowledge of Things in Themselves, *Australasian Journal of Philosophy* 82, pp.129–136 https://doi.org/10.1080/713659791

Locke, J., 1690/1978, *An Essay Concerning Human Understanding,* Nidditch, P. H. (ed.) Oxford: Clarendon Press. Excerpted in Heil (2003b), pp. 59–63

Levine, J., 1983, Materialism and Qualia: The Explanatory Gap, *Pacific Philosophical Quarterly* (64) pp. 354-61 https://doi.org/10.1111/j.1468-0114.1983.tb00207.x

Lewis, D. K., 1995, Should a materialist believe in qualia?, *Australasian Journal of Philosophy* 73 (1) pp.140-44

Loar, B., 1990, Phenomenal states, *Philosophical Perspectives* 4, pp. 81-108

Locke, J., 1959, *An Essay Concerning Human Understanding Book II: Ideas*, New York, NY: Dover Publications, Inc.

Lorenz, H., 2009, Ancient Theories of Soul, *The Stanford Encyclopedia of Philosophy* (Summer 2009 Edition), Edward N. Zalta (ed.) URL = <https://plato.stanford.edu/archives/sum2009/entries/ancient-soul/>

McGinn, C., 1989, Can We Solve the Mind–Body Problem? *Mind* XCVIII, pp. 349–366 https://doi.org/10.1093/mind/XCVIII.391.349

McGinn, C., 1991, *The Problem of Consciousness: Essays Toward a Resolution*, Cambridge, MA: Blackwell

McGinn, C., 1993, *Problems in Philosophy: The Limits of Inquiry*, Oxford, UK: Blackwell

McGinn, C., 2003, What Constitutes the Mind-Body Problem? *Philosophical Issues* 13, pp. 148–162 https://doi.org/10.1111/1533-6077.00010

McGinn, C., 2004, *Consciousness and its Objects*, Oxford, UK: Oxford University Press University

Merleau-Ponty, M., 1945/2012, *Phenomenology of perception*, D. Landes (trans.). London, UK: Routledge

Montero, G., 2015, Russellian Physicalism, in Alter, T., Nagasawa, Y. (eds.), *Consciousness in the Physical World: Perspectives on Russellian Monism*, pp. 209-223, Oxford UK: Oxford University Press

Hassl-Mørch, H., 2018, Does Dispositionalism Entail Panpsychism? *Topoi* https://doi.org/10.1007/s11245-018-9604-y

McKenzie, K., 2017, Ontic Structural Realism, *Philosophy Compass* 12 https://doi.org/10.1111/phc3.12399

Nagel, T., 1974, What is It Like to Be a Bat? *Philosophical Review* 83, pp. 435–50 https://doi.org/10.2307/2183914

Nagel, T., 1980, The Limits of Objectivity, in *The Tanner Lectures on Human Values*, Vol. I, in McMurrmn S. (ed.) pp. 77-139, Cambridge, UK: Cambridge University Press

Nagel, T., 1986, *The View from Nowhere*, Oxford, UK: Oxford University Press

Nagel, T., 2012, *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False,* New York: Oxford University Press https://doi.org/10.1093/acprof:oso/9780199919758.001.0001

O'Dea, J., 2002, The Indexical Nature of Sensory Concepts, *Philosophical Papers*, 31 (2) pp. 169–81

Papineau, D., 1993, *Philosophical Naturalism*, Blackwell Publishing

Papineau, D., 2000, The Rise of Physicalism, in Gillett, C., Loewer, B.M. (eds.), *Physicalism and its Discontents*, Cambridge, UK: Cambridge University Press

Papineau, D., 2002, *Thinking About Consciousness*, Oxford UK: Oxford University Press

Papineau, D., 2006, Phenomenal and Perceptual Concepts, in Alter, T. & Walter, S. (eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism,* pp. 111–144, Oxford, UK: Oxford University Press

Patterson, S., 2000, How Cartesian was Descartes? in Crane, T. and Patterson S., *History of the Mind-Body Problem*, Edition 1st Edition, pp. 70-110, London, UK: Imprint Routledge

Pereira Jr, A., Nunn, C., Nixon, G. & Pregnolato, M., 2018, Consciousness and Cosmos: Building an Ontological Framework, *Journal of Consciousness Studies* 25 (3-4), pp. 181-205

Perry, J., 2001, *Knowledge, Possibility, and Consciousness*, Cambridge, MA: MIT Press

Pred, R. J., 2005, *Onflow: Dynamics of Consciousness and Experience,* Cambridge, MA: Bradford Books

Primas, H., 2009, Complementarity of Mind and Matter, in: Atmanspacher, H., Primas, H. (eds.), *Recasting Reality,* pp. 171–209, Berlin, DE: Springer https://doi.org/10.1007/978-3-540-85198-1_9

Richardson, R.C., 1982, The `Scandal' of Cartesian Interactionism, *Mind* 91, pp. 20–37.

Rozemond, M., 2003, Descartes, Mind-Body Union, and Holenmerism, P*hilosophical Topics* 31, pp. 343–367

Rosenberg, G., 2004, *A Place for Consciousness Probing the Deep Structure of the Natural World,* Oxford, UK: Oxford University Press

Ryle, G., 1949, *The Concept of Mind*, London UK: Hutchinson & Co

Rueger, A., McGivern, P., 2010, Hierarchies and levels of reality. *Synthese* 176, pp. 379–397 https://doi.org/10.1007/s11229-009-9572-2

Russell, B., 1921, *The Analysis of Mind*, Reprinted in1978, London: George Allen & Unwin

Russell, B., 1927, *Analysis of Matter,* London, UK: Kegan Paul

Russell, B., 1984, *The collected Papers of Bertrand Russell, Vol. 7, Theory of Knowledge_ The 1913 Manuscript,* Blackwell, K. and Ramsden-Eames, E. (eds.), London, UK: Routledge

Seager, W., 2010, Panpsychism, Aggregation and Combinatorial Infusion, *Mind and Matter* 8 (2), pp. 167-184

Seager, W., 2011, A New Idea of Reality: Pauli on the Unity of Mind and Matter. *Mind and Matter* 9 (1), pp. 37–52

Seager, W., 2013, Classical Levels, Russellian Monism and the Implicate Order, *Foundations of Physics* 43 (4), pp. 548-567

Seager, W., 2019, A panpsychist Manifesto, in William Seager (ed.), T*he Routledge Handbook of Panpsychism,* pp.144– 156, New York, NY: Routledge

Segundo-Ortin, M, Calvo, P., 2021, Consciousness and cognition in plants, *Wiley Interdisciplinary Reviews: Cognitive Science*

Schaffer, J., 2009, On What Grounds What. In David Manley, in Chalmers, D. J. & Wasserman R. (eds.), *Metametaphysics: New Essays on the Foundations of Ontology.* pp. 347–83 New York: Oxford University Press

Schaffer, J., 2010, Monism: The Priority of the Whole, *Philosophical Review* 119 (1), pp. 31-76

Seachris, J. W., 2013, The Sub Specie Aeternitatis Perspective and Normative Evaluations of Life's Meaningfulness: A Closer Look, *Ethical Theory and Moral Practice* 16, pp. 605–620

Shani, Itay. (2015). Cosmopsychism: A Holistic Approach to the Metaphysics of Experience, *Philosophical Papers*, 44(3), pp. 389–437

Silberstein, M., 2009, Essay Review: Why Neutral Monism is Superior to Panpsychism, *Mind and Matter* 7 (2) pp. 239-248

Silberstein, M., 2015, Extending Neutral Monism to the Hard Problem, *Journal of Consciousness Studies* 22, 181–194

Silberstein, M., 2017, Panentheism, Neutral Monism, and Advaita Vedanta. *Zygon* 52, pp. 1123–1145 https://doi.org/10.1111/zygo.12367

Silberstein, M., 2019, Neutral Monism Reborn: Breaking the Gridlock Between Emergence and Inherence, in William Seager (ed.), T*he Routledge Handbook of Panpsychism,* New York, NY: Routledge

Smart, J. J. C., 1959, Sensations and Brain Processes, *The Philosophical Review,* Vol. 68, No. 2 (Apr., 1959), pp. 141-156

Skrbina, D., 2009, Mind that Abides: Panpsychism in the New Millennium, John Benjamins Pub.

Spinoza, B., 1677/1985, Ethics, In: Curley, E.M. (ed.), T*he Collected Works of Spinoza,* Vol. I, Princeton, US: Princeton University Press

Stoljar, D., 2001a, Two Conceptions of the Physical, *Philosophical and Phenomenological Research* 62, pp. 253–281 https://doi.org/10.1111/j.1933-1592.2001.tb00056.x

Stoljar, D., 2001b (first edition) *Physicalism*, The Stanford Encyclopedia of Philosophy (Summer 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2021/entries/physicalism/>

Stoljar, D., 2005, Physicalism and Phenomenal Concepts. *Mind and Language* 20 (2), pp. 296–302 https://doi.org/10.1111/j.0268-1064.2005.00296.x

Stoljar, D., 2006, *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness,* New York, NY: Oxford University Press

Stoljar, D., 2014, Four Kinds of Russellian Monism, in: Kriegel, U. (Ed.), *Current Controversies in Philosophy of Mind,* p. 17, London, UK: Routledge

Stoljar, D., 2015, Russellian Monism or Nagelian Monism?, in: Alter, T., Nagasawa, Y. (eds.), *Consciousness in the Physical World: Perspectives on Russellian Monism*, pp 324-345

Stoljar, D., 2019, Panpsychism and Non-Standard Materialism: Some Comparative Remarks, in: Seager, W. (Ed.), *The Routledge Handbook of Panpsychism*, in William Seager (ed.), T*he Routledge Handbook of Panpsychism,* pp. 218-229, New York, NY: Routledge

Stoljar, D., 2020, The Epistemic Approach to the Problem of Consciousness, in: Kriegel, U. (ed.), *The Oxford Handbook of the Philosophy of Consciousness,* pp. 482-498, New York, NY: Routledge

Stubenberg, L., 2018, Neutral Monism, *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2018/entries/neutral-monism/>

Swinburne, R., 2018, Cartesian Substance Dualism, in Loose, J. J., Menuge A. J. L., Moreland, J. P., (eds), *The Blackwell Companion to Substance Dualism*, p. 133-151, Hoboken, NJ: Wiley-Blackwell

Sundström, P., 2011, Phenomenal Concepts, *Philosophy Compass* 6 (4):267-281

Tahko, T.E., 2021, *Unity of Science*, Cambridge, UK: Cambridge University Press https://doi.org/10.1017/9781108581417

Tanesini, A., 2018, Intellectual Humility as Attitude, *Philosophy and Phenomenological Research* 96, pp. 399–420 https://doi.org/10.1111/phpr.12326

Tegmark, M., 2014, *Our Mathematical Universe. My Quest for the Ultimate Nature of Reality*, New York, NY: Alfred A. Knopf

Tye, M., 2000, *Color, Consciousness and Content*, Cambridge, MA: MIT Press

Thompson, E., 2007, *Mind in Life: Biology, Phenomenology, and the Sciences of Mind,* Harvard University Press

Torrance, S., 2005, In Search of the Enactive: Introduction to Special Issue on Enactive Experience, *Phenomenology and Cognitive Science* 4, pp. 357–368 https://doi.org/10.1007/s11097-005-9004-9

Tully, R. E., 2003, Russell's Neutral Monism, in Griffin, N. (ed.), *The Cambridge Companion to Bertrand Russell*, Cambridge, UK: Cambridge University Press, 332–70

Van Inwagen, P., 1990, *Material Beings*, Ithaca: Cornell University Press

Van Riel, R., 2008, From Water to H2O - What Reduction is About, *Reduction in the Special Sciences* In: UNSPECIFIED- available online at http://philsci-archive.pitt.edu/4101/1/Reduction-paper-online.pdf [accessed January 2019]

Van Gulick, R., 2001, Reduction, Emergence and Other Recent Options on the Mind/Body Problem: A philosophic Overview, *Journal of Consciousness Studies* 8 (9-10), pp.1-34

Varela, F. J., Thompson, E., Rosch, E., 1991, *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge, MA: MIT Press

Velmans, M., 2008, Reflexive Monism, *Journal of Consciousness Studies*, 15(2), pp. 5–50

Velmans, M., 2012, Reflexive Monism Psychophysical Relations Among Mind, Matter, and Consciousness, *Journal of Consciousness Studies* 19, pp. 143–165

Weber, A., Varela, F. J., 2002, Life After Kant: Natural Purposes and the Autopoietic Foundations of Biological Individuality, *Phenomenology and the Cognitive Sciences,* Vol. 1, pp. 97–125

Weekes, A., 2012, The Mind-Body Problem and Whitehead's Nonreductive Monism, *Journal of Consciousness Studies* 19, pp. 40–66

Wheeler, M., 1997, Cognition's Coming Home: the Reunion of Life and Mind, in: *Proceedings of the Fourth European Conference on Artificial Life,* pp. 10–19, Cambridge, MA: MIT Press

Wheeler, M., 2011, Mind in Life or Life in Mind? Making Sense of Deep Continuity, *Journal of Consciousness Studies*, 18(5-6), pp. 148–168

Whitcomb, D., Battaly, H., Baehr, J., Howard-Snyder, D., 2017, Intellectual Humility: Owning Our Limitations, *Philosophy and Phenomenological Research* 94, pp. 509–539 https://doi.org/10.1111/phpr.12228

Westphal, J., 2016, *The Mind-Body Problem*, Cambridge, MA: MIT Press

Wishon, D., 2015, Russell on Russellian Monism, in Alter, T. And Nagasawa, Y. (ed.), *Consciousness in the Physical World: Perspectives on Russellian Monism,* pp. 91-118, Oxford UK: Oxford University Press

Wishon, D., 2017, Panpsychism, Panprotopsychism, and Neutral Monism, in McLaughlin, B., P. (ed.), Philosophy: Mind pp. 51-70, MacMillan Interdisciplinary Books

Wishon, D., 2021, Radical Empiricism, Neutral Monism, and the Elements of Mind, *The Monist*, Vol. 104, Issue 1, January 2021, pp. 125–151 https://doi.org/10.1093/monist/onaa026

Whitehead, A. N., 1925, *Science and the Modern World*, New York: Free Press

Whitehead, A. N., 1929, *Process and Reality: An Essay in Cosmology*, New York: Free Press