School of Advanced Studies IUSS Pavia

Ph.D. PROGRAM IN COGNITIVE NEUROSCIENCE AND PHILOSOPHY OF MIND

The argument from revelation against physicalism

Doctoral Dissertation

Candidate: Bruno Cortesi (XXXVI cycle)
Supervisor: Prof. Tomasetta Alfredo
Co-Supervisor: Prof. Di Francesco Michele

[...] and when I do not know the 'quid' of anything how can I know the 'quale'?

Plato, *Meno*, 71b.

Questo lavoro è dedicato a mia madre Cesarina ed a mia sorella Giulia.

Introduction

Human beings are conscious beings: we have emotions, perceptions, thoughts, memories, desires. We are not "mere things": while *prima facie* it makes no sense to wonder whether, say, a rock, or a piece of wood, or a book, or a television is sad, or remembers what happened yesterday, or feels sick, or has an orgasm, those questions become suddenly meaningful when they are applied to human beings, among other kinds of beings. We have an "inner life", each of us is "the center of a world", the world presents itself to us. But what is consciousness?

When it comes to the nature of consciousness, physicalism[1] – roughly, an approach whereby the nature of consciousness may be fully captured in physical terms – is the prevailing attitude among contemporary philosophers, and not just among them[2]. Overall, it seems fair to deem physicalism as the *Weltanschauung* of our time (see Gillett and Lower 2001, ix; Zanotti, PhD; Tomasetta, 2015_{a}).

Yet although most – though not all – philosophers nowadays take for granted that some version of physicalism ought to be true, the work of a number of authors in the eighties and the nineties of the last century pave the way for some anti–physicalist views to be rehabilitated[3], and those views are now deemed as at least worthy of

[1] In this work, unless otherwise specified whenever I will talk about physicalism I will have physicalism *about consciousness* in mind. I offer a more precise characterization of the view in 1.2

[2] As Chalmers (2018: 2) notes, there is a significant amount of psychological evidence concerning the "dualist intuitions" of laypeople, both children and adults. However, physicalism is mostly taken for granted both among philosophers of mind and in the philosophical debate that is not explicitly concerned with the nature of the human mind as well as among neuroscientists, psychologists and other researchers with empirically–oriented postures.

[3] Here I am mostly thinking about authors like Nagel (1974), Kripke (1980), Jackson (1982), Levine (1983), Chalmers (1996) and Nida–Rümelin (1997; 2007), among others. This is not meant to imply, of course, that before the works of these authors were published no anti–physicalist views had been proposed or discussed. There is an *enormous* amount of both dualism and idealism in pre–20th Century Philosophy (see Chalmers, 2019). Still, there is no denying that the *contemporary* debate on the ultimate nature of consciousness has been deeply influenced by the works cited in

being taken seriously alongside physicalist alternatives[4].

As Robert Stalnaker (2008: 26) has noted, most of the major anti–physicalist philosophical arguments that have been proposed, at least in the last decades, if not all of them, attempt to draw metaphysical conclusions – conclusions about the ultimate nature of consciousness – from epistemological premises – statements which revolve around the idea that our relation to the phenomena of consciousness is characterized by the distinct intimacy Philip Goff (2017: 108), points to in the passage below:

> Surely, you know exactly what your pain is – what it is for someone to feel pained in precisely that way – just by attending to pain and thinking about it in terms of how it feels. There is nothing in any way hidden from you about the reality of how you're feeling; nor is it possible that you're not really feeling that way. And that's because the feeling is "right there" for you, in such a way that its reality cannot be doubted.

Most notably, the attempt to draw metaphysical conclusions from epistemological premises lies beneath Saul Kripke's (1980) so called *modal argument* against physicalism, Frank Jackson's (1982) so called *knowledge argument* and David Chalmers' (1996; 1999; 2003_{b}; 2004; 2006a; 2009; $2010_{a,b}$) so called *conceivability argument* – also known as the *zombie argument* or the *two–dimensional argument*. Likewise, most of the major physicalist rebuttals of those arguments have attempted to somehow decouple items of knowledge – allegedly knowable facts, for instance – from metaphysically possible situations in which those items obtain – metaphysically possible worlds (Stalnaker, *Ibid.*).

this footnote, among others.

[4]Various versions of dualism are presented and/or defended in Chalmers (1996), Nida–Rümelin (e.g., 1997; 2007; MS) Foster (1981) Jackson (1982) Kripke (1980), Robinson (1982) and Swinburne (2013), among others. Forms of panpsychism are presented and/or defended in Bruntrup and Jaskolla (2017) Chalmers (2013), Mathews (2003), Rosenberg, (2004, Skrbina (Ed.) (2009), Strawson (2008_{a}), and Goff (2017), among others. Panqualitysm is defended in Coleman (2014; 2017). Idealism is presented and/or defended in Adams (2007), Kastrup (2017; 2018), Albahari (2019), Chalmers (2019), Builes (forth.), Robinson (2022) Bolender (2001), Foster (1982; 2008), Goldschmidt and Pearce (forth.), Meixner (2017), Pelczar (2015), Yetter–Chappell (forth.), among others. Of course, this list is far from being exhaustive.

One of the aims of this essay is to offer a diagnosis of why it is the case that, when it comes to conscious phenomena, epistemology and metaphysics seem so deeply intertwined, and why this has been so often thought to generate peculiar problems for a naturalistic viewpoint about the mind and about reality in general.

The fact that we have, or seem to have, a privileged access towards the nature of our own phenomenally conscious mental states – mental states there is something it is like for someone to be in (Nagel, 1974) – has been given a name. The thesis whereby the nature of experiences is immediately and directly presented to subjects just in virtue of their being had or experienced by those subjects (and thought about in terms of the way it is like for someone to have them) has become known as the thesis of *revelation* – just revelation for brevity. This essay will be about a philosophical argument – the argument from revelation – which from the truth of revelation plus a number of other premises infers the falsity of physicalism. I believe the argument from revelation may be *the single one most fundamental argument against physicalism which may be offered*. Of course, examining in detail every single anti–physicalist philosophical argument which may in principle be proposed, or even those that have actually been proposed, would require way more than a single dissertation. More modestly, and for obvious reasons, I will narrow down the scope of my analyses to the three most renowned and most discussed anti–physicalist arguments that have been proposed in the last 50 years, namely Kripke's modal argument, Jackson's knowledge argument and Chalmers' conceivability argument. I am going to make a case that they may all be reduced to the argument from revelation, meaning that no physicalist response to any of those arguments is compatible with revelation, hence it is at least necessary, and possibly necessary and sufficient, that the thesis of revelation is denied in order for physicalist theories of (almost) any variety (that I know of) to avoid the threat the three arguments pose.

The term 'revelation' was introduced by Mark Johnston (1992) to refer to Galen Strawson's (1989) claim whereby the nature of colors is fully revealed in

color–experiences, even though already in Russell (1910; 1912: 47)[5] one can find what is arguably a version of the thesis.

As Chalmers (2018, fn. 17) notes, while numerous anti–physicalist philosophers – most notably Philip Goff (2011; 2015; 2017) and Martine Nida–Rümelin (2007; MS)[6] – have identified a version of revelation as the central source of their anti–physicalism, other physicalist philosophers – most notably David Lewis (1995)[7] – have identified revelation as a claim that needs to be refuted if physicalism is to be true. Lewis' (1995) *Should a materialist believe in qualia?* is probably one of the *loci* where this tension between revelation – which he labels as *the identification thesis* – and physicalism emerges most clearly. There (*ivi:* 141–2) Lewis writes:

> Unfortunately there is more to the folk–psychological concept of qualia than I have yet said. It concerns the modus operandi of qualia. Folk psychology says, I think, that we identify the qualia of our experiences. We know exactly what they are—and that in an uncommonly demanding and literal sense of 'knowing what' [...] If qualia are physical properties of experiences, and experiences in turn are physical events, then it is certain that we seldom, if ever, [know the nature of] the qualia of our experiences. Making discoveries in neurophysiology is not so easy!

The main idea Lewis wants to convey is rather straightforward. On the one hand we seem to be in a rather peculiar, not to say unique, epistemic situation with respect to our own experiences[8]. One the other hand, it is obviously not the case that one can get to know anything about the complex physiology of her brain *just by being in a state of pain and thinking about it as painful*. In light of this *impasse*, apparently

[5]As well as, most likely, in Descartes, e.g. (1641[1985] esp. Med. 2 and Med. 6), Brentano (1874[2009]) and Husserl (e.g., 1912[1983] 10–16; *ivi:* 24–36; 1936[1970]. See also Tomasetta (forth.)

[6]But see also Liu (2020; 2021; forth); Horgan and Tienson (2001) Coleman (2019) Majeed (2017) and Trogdon (2017), among others. Strawson (2008_a; 2009; 2016) is one of the pioneers of revelation – at least within the contemporary debate in analytic philosophy – but he identifies himself as a physicalist, though admittedly of a very peculiar kind.

[7]But see also Stoljar (2006; 2009): Stalnaker (2008: 99) and Damnjanovic (2012), among others.

[8]The ways it is like for someone to undergo certain experiences have been very often referred to as *qualia*, even though in 1.1 I will suggest that this might not be the best way to capture the obvious (to my eyes) fact there is a difference, from a subject's viewpoint, between having and not having a certain experience, namely a conscious mental state.

one is left at a crossroad between two options:

$\neg$*Rev*: Nothing truly essential about the experiences of subjects is nor could be revealed to those subjects just in virtue of having those experiences and thinking about them in terms of they way it is like for someone to have them.

Rev: Subjects are put in a position to know (part of) the nature of their experiences just by having them and thinking about them in terms of the way it is like for someone to have them.

Lewis' point may be laid down in the form of a philosophical argument for the incompatibility of revelation and physicalism. Such an argument might plausibly look like this:

[1] If revelation is true, then having an experience and thinking about it in terms of the way it is like to have it is sufficient to grasp at least part of its nature.

[2] In order for physicalism to be true – that is, in order for experiences to be nothing over and above a kind of physical states – the nature of experiences should fully captured or at least capturable in principle by the physical sciences. That is to say, for every (kind of) experience there should be some piece of physical knowledge such that by coming to have that piece of knowledge – by making the relevant physical discovery/ies – one would get to know what the relevant (kind of) experience essentially consists in.

[3] By thinking about an experience in terms of the way it is like for someone to have it and as a physical phenomenon, one is not put on that basis alone in a position to rationally judge that she's thinking about one and the very same thing.

[4] If both revelation and physicalism were true, by thinking about an experience in terms of the way it is like for someone to have it and as a physical phenomenon, one should be put on that basis alone in a position to rationally judge that she's thinking about one and the very same thing.

Conclusion: If physicalism is true revelation must be false. Hence, if revelation is true, physicalism is false.

Provided the argument just presented is valid, by combining it with a further premise asserting the truth of revelation, one obtains an argument for the falsity of physicalism[9].

[1] If revelation is true, then having an experience and thinking about it in terms of the way it is like to have it is sufficient to grasp at least part of its nature.

[2] Revelation is true.

[3] In order for physicalism to be true – that is, in order for experiences to be nothing over and above a kind of physical states – the nature of experiences should fully captured or at least capturable in principle by the physical sciences. That is to say, for every (kind of) experience there should be some piece of physical knowledge such that by coming to have that piece of knowledge – by making the relevant physical discovery/ies – one would get to know what the relevant (kind of) experience essentially consists in.

[4] By thinking about an experience in terms of the way it is like for someone to have it and as a physical phenomenon, one is not put on that basis alone in a position to rationally judge that she's thinking about one and the very same thing.

[5] If both revelation and physicalism were true, by thinking about an experience in terms of the way it is like for someone to have it and as a physical phenomenon, one should be put on that basis alone in a position to rationally judge that she's thinking about one and the very same thing.

[6] If physicalism is true revelation must be false. Hence, if revelation is true, physicalism is false.

[9] A version of this argument is presented in Nida–Rümelin (2007). A more detailed presentation of the argument, as well as a discussion of each and every premise of it, will be offered in 1.8.

Conclusion: Physicalism is false [from 2 and 6]

This is the argument from revelation. The argument from revelation will be the topic of this dissertation: it is to this argument, I claim, that Kripke's modal argument, Jackson's knowledge argument and Chalmers' conceivability argument against physicalism may all be reduced.

Moreover, and crucially, I will attempt to show that revelation is *the* premise of the argument that needs to be denied in order for physicalism to be safe and defended in order for physicalism to be refuted. In fact, most of the physicalist attempts to rebut the argument have attacked either its fourth premise or its fifth one. I think neither is a profitable strategy for the physicalist. Let me explain.

It has been claimed that there is no physico–phenomenal explanatory gap (Levine, 1983) at all, hence there is nothing Mary learns once set free from her black–and–white environment (see Jackson, 1982) and so called phenomenal zombies are neither conceivable nor possible (see Chalmers, 1996; 1999; 2003_{b}; 2004; 2006a; 2009; $2010_{a,b}$). This is the type of response to the argument so called *a priori* or *type–A* physicalists would offer[10].

By contrast, so called *a posteriori* or *type–B* physicalists[11] do accept that there is something Mary learns upon her release and that phenomenal zombies are conceivable. Hence, they accept the fourth premise of the argument from revelation. However, as physicalists they contend that all the facts are physical facts, and that phenomenal zombies, although conceivable, are not metaphysically possible. Some type–B physicalists, in turn, though not all of them, deny the fifth premise of the argument from revelation, claiming that it is possible to grasp the nature of one and the very same entity in two or more conceptually distinct ways. This view has

[10]See Chalmers, 2003_{b} for the distinction between type–A, type–B and Type–C physicalism. See Ryle (1949) Armstrong (1968), Lewis (1966; 1988; 1994) Harman (1990) Dennett (1991), Dretske (1995) and Rey (1995) as prototypical examples, among many others, of a type–A approach. See also 1.8.2. and 3.1.

[11]See Tye (1995), Lycan (1996), Hill (1997), Hill and McLaughlin (1999), Block and Stalnaker (1999), Papineau (1993; 2002) and Perry (2001) as prototypical examples of a type–B approach, among *many* others. See also 1.8.2 and 3.1, even though several type–B views will be discussed in more detail throughout the essay.

become known as *dual carving*[12]. I will argue that there are good reasons to doubt dual carving may be true.

At any rate, the arguments I will draw to the case that revelation is incompatible with physicalism are neutral with respect both to the fourth premise of the argument above and to the fifth one. Hence, I hope to establish that even if both the fourth premise of the argument and the fifth one happened to be bracketed or even denied, revelation would still be sufficient *alone* to confute physicalism. By doing so, my ambition is to offer a radical simplification of the philosophical debate on whether physicalism should be our default metaphysical stance in consciousness studies or not, for as far as in can see that debate rotates around a single fundamental idea, or it has rotated around this idea in the last five decades at least: *whether conscious phenomena are such that their essence is revealed to subjects immediately and directly in the very having of those experiences, or not.*

Here is how I will proceed throughout the essay. The first chapter is mostly introductory. In this chapter, I set the stage by elaborating on a number of working assumptions that will be on the background throughout the work. I also give a closer look at the argument from revelation and examine each and every premise of it in more detail. The second chapter deals with Chalmers' conceivability argument. In this chapter, I make a case that, provided the two–dimensional semantic apparatus Chalmers deploys to draw the argument is in place, revelation threatens physicalism no matter whether so called phenomenal zombies are taken to be ideally conceivable or not. Hence, it is necessary and most likely necessary and sufficient that revelation is denied in order for physicalists to avoid the threat Chalmers' two–dimensional semantics pose. As far as I can see, the arguments I draw in the second chapter to the case that revelation is incompatible with physicalism may be reiterated as for what concerns Jackson's knowledge argument, which is the topic of the third chapter. Still, in this chapter I offer some further remarks on a number of possible

[12]See Goff (2011, 200–201; 2017: 125–32), Díaz–León (2013; 2016: 1193), Taylor (2013), Damnjanovic (2012: 76–77), Elpidorou (2015), Levin, 2019, and (possibly) Loar, (1990/1997; 2002). Dual carving will be discussed extensively in 3.4

physicalist reactions that have been offered to the thought experiment on which knowledge argument is based, arguing that it is necessary and possibly necessary and sufficient, in order for each of them to go through, that revelation is denied. This chapter contains also an interlude, in which I deal with Saul Kripke's modal argument suggesting that it may be reduced to the argument from revelation as well.

For the most part, it won't be my aim to disprove physicalism. Rather, what I intend to do is to show what is it that should be denied in order for physicalism to be safe and defended in order for physicalism to be confuted. In other words, I will mostly be reasoning hypothetically: *if* revelation is true, physicalism is doomed no matter whether certain other assumptions are in place or not. However, towards the end of the essay I will suggest that we may have very strong *prima facie* reasons to believe revelation is true, and hence physicalism should not be our default metaphysical stance in consciousness studies. I will conclude the essay by offering *few* remarks on what, I believe, may be the best metaphysical framework we have available to accomodate the obvious datum of phenomenal consciousness (plus revelation).

Contents

1 Assumptions **15**

1.1 What are phenomenal properties? 15

1.2 What is physicalism? 29

1.3 What are fundamental properties? 36

1.4 What are essential properties? 39

1.5 Interlude: some historical background on concepts: intensions, extensions and Kripkean rigid designators. 41

1.6 What are physical concepts? 44

1.7 What are phenomenal concepts? 45

1.8 The argument from revelation: a closer look 49

1.8.1 Premise 1 51

1.8.2 Premise 2 52

1.8.3 Premise 3 56

1.8.4 Premise 4 57

2 The conceivability argument is not an argument about conceivability 60

2.1 Introduction 60

2.2 Chalmers' two–dimensional semantics: a brief exposition 63

2.2.1 Conceivability 63

2.2.2 Scenarios 65

2.2.3 Primary and secondary intensions 68

2.3 The CP principle under an epistemic and under a metaphysical approach to scenarios . . . 71
2.3.1 The CP principle under an epistemic approach to scenarios . . 72
2.3.2 The CP principle under a metaphysical approach to scenarios 74
2.4 Primary ideal negative conceivability and the cognitive independence of physical and phenomenal concepts . . . 75
2.5 The anti–zombie argument . . . 79
2.6 The four versions of the zombie argument . . . 83
2.6.1 Version I . . . 84
2.6.2 Version II . . . 85
2.6.3 Version III . . . 87
2.6.4 Version IV . . . 89
2.7 Tacking stock (I) . . . 90
2.8 Why there are good reasons to believe that the claim that phenomenal concepts have identical primary and secondary intension and phenomenal transparency are logically equivalent: some hints . . . 91
2.9 More on why the claim that phenomenal concepts have identical primary and secondary intension and phenomenal transparency are logically equivalent: the NARD principle . . . 92
2.10 Why phenomenal transparency plus the denial of the cognitive independence of physical and phenomenal concepts is incompatible with physicalism . . . 98
2.10.1 Option 1 . . . 100
2.10.2 Option 2 . . . 104
2.10.3 Options 3 and 4 . . . 105
2.11 Tacking stock (II) . . . 106
2.12 Why Chalmers' overall two dimensional framework is incompatible with the existence of radically opaque concepts . . . 108

2.13 Why phenomenal translucency plus the denial of the cognitive independence of physical and phenomenal concepts is incompatible with physicalism . . . 109
2.14 Why Revelation threatens physicalism also under the assumption that physical and phenomenal concepts are cognitively independent 113

3 The knowledge argument, the modal argument and the argument from revelation 117

3.1 Introduction . . . 117
3.1.1 There is nothing new Mary learns upon her release, and all the facts about color vision are physical facts . . . 122
3.1.2 There is something new Mary learns upon her release, but all the facts about color vision are physical facts . . . 123
3.2 More on the old–facts–new–ways strategy, plus an interlude on Kripke's modal argument against physicalism . . . 124
3.3 Phenomenal concepts as indexicals and/or as recognitional/demonstratives 128
3.4 On dual carving . . . 129
3.5 The Powerful Qualities view: Mørch (2018) on why dispositionalism may entail panpsychism. . . . 141
3.6 The ability hypothesis . . . 156
3.7 The acquaintance hypothesis . . . 158

4 Looking back, and then forward 169

4.1 Part one: looking back . . . 169
4.2 Part two: looking forward . . . 180
4.2.1 Revelation: an argument from obviousness? . . . 181
4.2.2 What now? . . . 183

Chapter 1

Assumptions

1.1 What are phenomenal properties?

It will be useful to start by providing some working characterizations of some the terms and concepts I will deploy and elaborate on throughout the essay. Let me begin with conscious (or phenomenal, or experiential interchangeably) properties.

I will adopt a characterization of states (or facts, or events, or phenomena interchangeably) as exemplifications of properties over spans of time[1]. In particular, conscious states (or phenomenal states, or experiential states, or experiences interchangeably) will be defined as those events which consist in the instantiation of at least one phenomenal property, to be characterized presently, over a certain interval of time.

I will follow Eric Schwitzgebel (2016) in trying to define phenomenal properties, in turn, as innocently as I can manage[2] I will start by offering a list of what I take to be some uncontroversial examples of conscious states as well as some examples of states

[1]For a defense of this account of events see Kim (1976)

[2]Ostensive definitions or definitions–by–examples of the kind, following Schwitzgebel, I am pursuing here are not uncommon in the philosophical literature on consciousness and related topics. Martine Nida–Rümelin (e.g., MS; 2016) deploys both positive and negative examples to capture both the notion of an experiencing subject and the one of an experience. Likewise, Angela Mendelovici (2018) defines ostensively the notion of intentionality – the "of–ness" or "about–ness" of mental states. Among the contemporary philosophers that I can think of who deploy either a list of examples or a mixture of examples and synonyms (plus possibly other strategies) to capture the notion of phenomenal consciousness at least at an initial, superficial, stage there are John Searle (1992: 83), Ned Block (1995/2007: 166–168), David Chalmers (1996: 4), Philip Goff (2019) and Galen Strawson (2016), just to mention a few.

which, at least intuitively or *prima facie*[3], do not involve the instantiation of any phenomenal property. My ambition is to isolate a notion of phenomenal properties so obvious as to make it almost impossible for it to be misunderstood. This may be a vain hope, as Strawson (2016: 86) rightly but unfortunately points out[4]. Then, I am going to offer an explicit definition of phenomenal properties adapted from the one Martine Nida–Rümelin (MS; 2018) offers (but see also Taylor, 2020) and elaborate on why I think my proposed definition is the best–suited to capture the examples provided. To conclude this section, I am going to argue that the definition of phenomenal properties I will buy into meets some *desiderata* which, arguably, any suitable definition of phenomenal properties should accommodate. These *desiderata* are the following:

1. The primary goal of a definition–by–examples is to isolate a "obvious" or "natural" target reference such that almost everyone will latch onto it once the relevant positive and negative examples have been provided. Moreover, ideally there should be *only one* "natural" target reference for people to latch onto, so as to avoid ambiguities (see Schwitzgebel, 2016: 5–6).

2. The sought definition of phenomenal properties should be as parsimonious as possible: if it defines phenomenal properties in terms of other notions or reduces phenomenal properties to them it should do so for a reason, without

[3]At this initial stage, my only aim is to isolate a sufficiently intuitive notion of phenomenal properties, so that anyone, including philosophically naive readers, could latch onto it. Towards the end of the essay I will suggest that the view David Builes (forth.) calls *modal idealism* may be the best metaphysical framework we have available to accommodate the obvious (to my eyes) datum of phenomenal consciousness (plus revelation). According to modal idealism, necessarily every fundamental property (see 1.3 for a definition of fundamentality) is a phenomenal property (to be characterized below). Modal idealism entails the admittedly very counterintuitive conclusion that, since consciousness is fundamental and ubiquitous in the material realm, even the states that figure in the list of negative examples I will offer may involve necessarily the instantiation of some phenomenal property, and thus, possibly, «subjects, or aggregates of subjects, or parts of a wider subject» (see Albahari, 2019: 1). This is what explains the 'at least intuitively or *prima facie*' bit in the sentence above. For the same reason, below I may refer to the negative examples I will provide as examples of states that *supposedly* do not entail phenomenal consciousness.

[4]Strawson (2016: 86) says that to be a realist about experience is «to to continue to take color experience or taste experience or pain experience, considered just as a mental occurrence, to be exactly what one took it to be, quite unreflectively, simply in having it, before one did any philosophy». I could not agree more. *This* is the kind of properties I will try to fix reference on in this section. Those for whom this passage of Strawson is already sufficient to get a grip on the kind of properties I'm after, may skip this section.

multiplying entities without necessity.

3. Although the relevant definition should be substantive enough so as to include at least some of the interesting features consciousness is typically associated with, it should also be as little theory–laden as possible, that is to say, it should try to avoid as many potentially troublesome metaphysical, meta–theoretical, methodological, epistemological *etc.* commitments as possible: possibly, the notion of phenomenal properties we are after should not acquire its reference via any theoretical assumption or stipulation.

4. Moreover, the sought definition should capture, at least superficially or in a very minimal way, the sense of puzzlement and of mysteriousness consciousness is often thought to generate: *if* a reductive[5] naturalistic explanation of phenomenal properties is possible, this should not immediately and uncontroversially follow from their definition (see Schwitzgebel, *ivi:* 3).

Here is my list of examples. Positive examples – namely, examples of conscious states – include:

Sensory–perceptual states of various degrees of complexity and "determinability" such as: hearing the loud sound of a trumpet, hearing a symphony, smelling coffee, tasting red wine, tasting french onion soup, touching a very soft sweater, running one's hands over a piece of sandpaper, seeing a patch of red paint, watching a painting *et cetera.*

Other bodily, somatosensory, algedonic etc. sensations of various degrees of complexity such as: feeling the warmth of the sun on one's skin, feeling the touch of water of one's skin while having a swim in a very cold lake, having a sharp pain at one's foot, having an itch at one's leg, having a nasty headache, having an orgasm, feeling nauseous, being thirsty, being hungry, having an hangover, being tired but also relaxed after a long run *et cetera.*

[5]Here the word 'reductive' makes all the difference in the world. Following Strawson (e.g., 2008_{a}; 2016) I take myself to be a full–blown naturalist, maybe even a materialist (though not in the sense most people give to 'materialism', which will be explored in depth in this essay). Still I don't believe consciousness may be reduced to any more fundamental kind of substance.

Emotions, moods and other affective states of various degrees of complexity such as: being angry, fearing the growls of a dog, being anxious for one's incumbent *viva*, being happy for having won a scholarship, mourning for the death of one's grandmother, being depressed, being enraptured looking at a Van Gogh painting, enjoying the company of one's friends on an evening, *et cetera*.

Occurrent evaluative states such as: liking the taste of the ice cream one is having, disliking the smell of rotten fish one is smelling, *et cetera*.

Occurrent thoughts and other cognitive or propositional attitudes such as: entertaining the thought that today is Friday, coming to understand the meaning of a sentence, coming to grasp a mathematical proof, wondering whether tomorrow will rain, doubting that the conservative party will win the elections, hoping that the next week will be a bit cooler that the present one, remembering that one has to step by the secretariat to drop a message, remembering one's 18th birthday, remembering what one has had for dinner yesterday, *et cetera*.

Occurrent appetitive states such as: wishing an ice cream, craving for coffee, *et cetera*.

I could have included many more examples in my list – e.g., conscious imagery, dream experiences, (Schwitzgebel, 2016: 6–7) and so on – but hopefully you get the idea. I have deliberately avoided potentially controversial examples such as being *constantly* aware of the position of one's feet in the shoes (Schwitzgebel, 2016: 10; 2007), being *inattentively* aware of the noise made by the fan of one's laptop (*Ibid.*), being aware of oneself as conscious (Howell, 2023), conscious free will (Wegner, 2018), *et cetera*. It has been contested that these and/or other phenomena really exist and/or involve phenomenal consciousness, and I want my working definition of phenomenal properties to be acceptable by the broadest possible audience[6]. If some

[6]Some might want to object that I have included occurrent thoughts and other occurrent

readers have independent reasons to believe that these phenomena exist and entail phenomenal consciousness, they may add them to their own list of positive examples, but I think my list to is already ample enough to allow almost any reader to latch onto one and only one "natural" target reference. At the same time, nothing of what I will say implies *per se* or even suggests that these or other potentially controversial phenomena do not exist or do not involve phenomenal consciousness.

Negative examples – examples of phenomena which, at least intuitively or *prima facie*, do not involve the instantiation of any phenomenal property – include: the growing of one's fingernails or of one's hair, lipid's absorption in one's intestines, a piece of wood burning, any kind of chemical reaction, weighing exactly 57 kilograms[7], a car traveling at 70 kilometers per hour, positive or negative electric charge, snowfalls, storms and any other kind of meteorological phenomenon, earthquakes, earth's rotation, cell division *et cetera*.

The first feature which I guess would strike (almost) everyone is that necessarily, whenever any of the states that figure in my list of positive examples occur a *subject, a some–one – as opposed to a some–thing –, an individual* is involved[8]. Storms do

cognitive/propositional attitudes in my list of positive examples. The existence of a kind of cognitive phenomenology (see Bayne and Montague Eds., 2011) has been put into doubt and oftentimes even explicitly rejected (see Braddon-Mitchell and Jackson, 2007: 129, Tye, 1995: 4, Carruthers, 2005: 138–9, Nelkin, 1989: 424 as prominent examples of this attitude, among many others). So, the opponent would say, my list is not, after all, so neutral as I am claiming it to be. However, crucially, I am not implying that cognitive states involve a proprietary, distinctive, individuative or irreducible phenomenology: although as a matter of fact I believe they do, nothing of what I will say in this work hinges upon this. Hence, both my list of positive examples and the definition of phenomenal properties I will offer below are compatible with the phenomenology of cognitive states being reducible to that of sensory states, emotional ones, or other kinds of phenomenal states.

[7]There may be a phenomenology of being severely overweight as opposed to being severely underweight. Yet I doubt that there is a phenomenology of weighing *exactly 57 kilograms*. Moreover, the property of weighing exactly 57 kilograms seems to be not such that necessarily that whenever it is instantiated there is a subject involved (see below), for a car, or a rock, or a piece of wood, *et cetera* may also weigh 57 kilograms.

[8]This work won't be devoted to the investigation of the nature of conscious subjects or experiencing individuals. I think a rather broad understanding of what an experiencing subject is should suffice for my purposes. Again, a list of both positive and negative sufficiently uncontroversial examples should do the job of allowing readers to latch onto one and only one "natural" target reference. Things like you, me, your cat, a horse, a chimpanzee, are experiencing subjects. Things like a piece of wood, a rock, a book, a television, a car, *et cetera* are not (supposedly) experiencing subjects. A worry might raised at this point. In fact, it may be argued that the most obvious feature which the positive examples have while the negatives ones lack is that the former are capable at least in principle or have the potential to be in conscious states, that is to say, to instantiate phenomenal properties: while a cat might be in pain or fear the growls of a dog, a

not happen *to someone*: they just happen. By contrast, there are no headaches or depressive moods freely floating around, nor thoughts without a thinker. Moreover, say, the property of weighing exactly 57 kilograms is not such that necessarily whenever it is instantiated a subject is involved[9], for a rock, or a table, and so on may also weigh exactly 57 kilograms.

car or a television may not. This, the objector would continue, may lead to a definitional circle whereby phenomenal properties are defined in terms of experiencing subjects and experiencing subjects are defined in terms of phenomenal properties. However, Nida–Rümelin (MS) makes the case that there may be acceptable definitional circles. Roughly, her argument goes as follows. Definitional circularity would raise a concern only if the definitions concerned intended to achieve conceptual or ontological reduction. However, this isn't the case insofar as experiential properties and experiencing subjects are concerned. Instead, their circular nature underscores a theoretical standpoint. When considered together, these definitions assign both conscious individuals and experiential properties both conceptual and ontological fundamentality. There is no way, according to Nida–Rümelin, to understand what an experiencing subject *is* without implicit reference to experiential properties, and *vice versa*. For what it's worth, I am sympathetic with a view which construes subjects as "mental lives", namely as "experiential streams" or as "sums" or "collections" of experiences. The historical reference for this kind of view is David Lewis (e.g., 1971; 1976; 1983a). It has also been defended by Anthony Quinton (1962: 398; 1973: 97–105), Carol Rovane (1997: 172), and Scott Campbell (2006) (See also Tomasetta 2015_{b}). Combined with an account of experiences as instantiations of phenomenal properties over spans of time, this view would collapse on a version of the so called bundle theory whereby subjects are nothing over and above bundles of phenomenal properties, possibly suitably held together via a relation of "co–consciousness". On this view, strictly speaking conscious subjects do not exist: only phenomenal properties do. Or better, conscious subjects supervene upon the experiences – instantiations of phenomenal properties – they "have", but do not constitute a substratum occupying a distinct ontological category with respect to experiences/instantiations–of–phenomenal–properties themselves. Hence, talk of subjects – as well as talk of subjects *having* or *undergoing* experiences, and/or talk of the experiences/experiential–properties *of* conscious subjects would amount to a "useful fiction", so to say, to refer to the bundles of suitably arranged experiential properties the subjects essentially consist in. Hence, again, there would be no ontological reduction, for there would be nothing to reduce phenomenal properties to. Builes (2021) defends the claim that objects are nothing over and above their corresponding properties. Strawson (2021) has also defended the view attributing it to Descartes, Spinoza and Kant as well. Strawson (2009) also defends the view that conscious subjects are not distinct from the phenomenal properties they "have" (See Loux and Crisp, 2017: 82–117 for an introduction to the bundle theory. See also Builes, forth.). At any rate, as far as I can see the definition of phenomenal properties I will buy into is compatible with almost any theory of experiencing subjects that is on the market, and I won't try to offer any positive argument for the bundle theory as applied to experiential subjects.

[9]Rather than characterizing phenomenal properties as those properties that necessarily are instantiated *by* a subject, or that necessarily require a subject in order to be instantiated, I prefer to characterize them as those properties such that, necessarily, whenever one of them is instantiated a subject is involved. This is because the former two formulations suggest that subjects are ontologically distinct from the phenomenal properties they instantiate, whereas the latter formulation, though not incompatible with a view of subjects as categorically distinct from phenomenal properties, is also compatible with a version of the so called bundle theory whereby subjects are nothing and above the phenomenal properties they "have" (See fn. 8 of this chapter). Nonetheless, every now and then I may speak of subjects *having* or *undergoing* experiences. This formulation, again, suggests that subjects are categorically distinct from the experiences they have or undergo. Readers with sympathies for the bundle theory should read those passages as useful *Façons de parler* to refer to the experiential states – namely, instantiations of phenomenal properties – the subjects essentially consist in.

The subject–constraint is not enough, though, for merely defining conscious phenomena simply as those states such that, necessarily, whenever any of them occurs a subject is involved may make some of the negative examples count as conscious. In fact, some may intuit that the growing of fingernails necessarily requires someone, whose fingernails are growing, to occur; same for lipid absorption in one's intestines[10].

What really distinguishes the positive examples provided from the negative ones is that the former *feel in a certain way to the subject involved, they are experienced somehow, they make a difference for the subject involved just in virtue of their being instantiated, they necessarily involve the first–person perspective.*

In what follows I will deploy Thomas Nagel (1974) iconic expression "what it is like for someone to . . . " to refer to this second, peculiar feature which distinguishes the positive examples from the negative ones, just because I cannot come up with any better expression[11]. Alternatively, every now and then I might say that conscious events necessarily involve the instantiation of a *phenomenology* or (interchangeably) of a *phenomenal character*. Some caveat on my part are crucially in order at this

[10]It is conceivable that at some point in the future there might be humanoid robots so sophisticated as to possess real, actual digestive systems, process nutrients to derive energy, have real fingernails, real hair, *et cetera*. If this did not go along with the capability of these robots to instantiate phenomenal properties, this would be yet another reason why the fingernails example and the lipid absorption one do not count as positive examples of conscious states. Likewise, there may be very simple forms of life which do not instantiate phenomenal consciousness but yet have digestive or proto–digestive systems. Yet I am also open to the possibility of there being genuinely conscious robots in the future. But even if there were genuinely conscious robots, my definition of phenomenal properties below would not be affected. Simply, the class of possible experiencing subjects would be enlarged.

[11]Truth be told, although normally attributed to Nagel, the expression is used with a sense very close to that with which it is normally associated nowadays at least as early as in an article by Sprigge and Montefiore published in 1971 concerning efficient and final causes. Here is what Sprigge and Montegiore (1971: 169) write: «In the present state of philosophy, it is possible that I may be challenged to say what I mean by "consciousness". I shall offer a hint as to the meaning of the expression, without thinking it incumbent upon me to establish (for anyone absurd enough to doubt it) that there actually are such phases of existence. One is wondering about the consciousness which an object possesses whenever one wonders what it must be like being that object. Concerning an object deemed non–conscious one cannot thus wonder. To wonder what it is like being an object is to concern oneself with a question different from any scientific or practical question about the observable properties or behaviour of that object or about the mechanisms which underlie such properties or behaviour. A behaviouristically or physicalistically minded psychologist might be very good at knowing what a psychopath was like, without having any idea what it was like being a psychopath». Likewise, the idea beneath the renowned knowledge argument against physicalism predates Jackson by a long way (see fn. 1 of the third chapter).

point:

> *Caveat 1.* Let me underline that I am *not* attaching *any* peculiar, problematic, or potentially controversial technical or theoretical meaning to terms such as what–it's–like–ness, phenomenology, phenomenal character and the like. This work will mostly be devoted to an investigation of the nature of phenomenal properties and their relation to physical ones. For the time being, though, I am not assuming that phenomenal properties are intrinsic, intrinsically private, intrinsically ineffable, irreducible to physical states, monadic, or whatever feature you might want to add. By terms like 'what–it's–like–ness' and the like I *just* mean *one of the two most folk–psychologically obvious features which all the positive examples provided above possess and all the negative examples (supposedly) lack* (Schwitzgebel, 2016: 8).

> *Caveat 2.* Another point is worthy of being underlined. The boundaries between the kinds of experience that figure in my examples are somehow fuzzy. For instance, the experience of craving for coffee might involve appetitive, cognitive, bodily and in some cases even emotional components. Likewise, the experience of staring at a Van Gogh's painting will involve both sensory/perceptual elements and (presumably) emotional ones. I won't take a stance here on whether any of the kinds of experience listed above involve the instantiation of a proprietary and distinctive[12] phenomenology or whether, instead, the phenomenology of, say, emotional states reduces to that of bodily states, or cognitive ones, or *vice versa*, *et cetera*. For me, it is enough that we agree that there is something, *a* way, it is like for a subject to undergo each of the states that figure in my list of positive examples: describing *what* it is like for a subject to be in each of the states mentioned would be a matter for another work.

[12]The phenomenology of a given particular conscious state is proprietary iff it is distinct from the way it is like for a subject to be in any other type of conscious states and distinctive iff it is distinct from the way it is like for a subject to be in any other particular conscious states, including conscious states of the same type. See Pitt (2004).

Caveat 3. In what follows, I will bracket any kind of skepticism about other minds. Moreover, I will work under the assumption that distinct subjects may undergo the same experiences – e.g., distinct subjects may be visually presented with the color red, *et cetera*.

Eventually, I am in a position to offer a first, tentative working definition of phenomenal properties, and to verify whether it meets the *desiderata* 1, 2, 3 and 4 listed above. Both the first and the second definition below are adapted from Nida–Rümelin (MS; see also her 2018):

> *Definition 1. of phenomenal properties*
> A property P is a phenomenal property iff, if P is instantiated at time *t*, then in virtue of its being instantiated there is something it is like for a subject S to live through time *t*.

As it is worded, definition 1 is not entirely satisfying yet. For consider a property like 'being in the harbor district of Naples'. This property might be instantiated by some given subject at time *t*; moreover, it might ground the instantiation of some given property such as 'smelling freshly caught fish'. That is to say, someone might be smelling freshly caught fish because or in virtue of the fact that she is in the harbor district of Naples. Given that, in turn, there is something is like for a subject to smell freshly caught fish, apparently the property 'being in the harbor district of Naples' verifies definition 1, for whenever a subject is conscious and in the harbor district of Naples there is something it is like for that subject to live through that moment (see Nida–Rümelin, MS). Yet there is a clear sense whereby 'being in the harbor district of Naples' is not in itself a phenomenal property like 'being in pain' is: first, a house, or a street, or a car, *etc.* may also be in the harbor district of Naples, so the former property is not such that necessarily whenever it is instantiated a subject is involved. Moreover, a subject may be in the harbor district of Naples but in a coma, or in a dreamless sleep, or even dead, thus (presumably) unconscious. Someone in a coma, though, would not and could not (presumably) be

smelling freshly caught fish, be in pain or remember her 18th birthday[13]. The kind of properties we are after is such that their instantiation at time t consists in the fact that there is something it like for a subject to live through time t.

There is also a further issue with definition 1: clearly, the way it is like for someone to be in an excruciating pain is different from the way it is like for the same subject to taste strawberries. Indeed, this is what distinguishes the experience of being in pain from the one of tasting strawberries, at least from the first–person perspective. We want our definition to capture the very intuitive sense whereby phenomenal properties are such that their instantiation consists in there being, for each distinct phenomenal property, a *specific* way it is like for a subject to live through time t if it is instantiated at time t.

By taking these issues into due account, we come to a second, more accurate definition of phenomenal properties:

> *Definition 2. of phenomenal properties*
>
> A property P is a phenomenal property iff the following two conditions verify:
>
> 1. There is a specific way W such that, if P is instantiated a time t, then it is like W for a subject S to live through time t.
> 2. The instantiation of P at time t consists in it being like W for S to live through time t.

I think we can be satisfied with definition 2. Thus, in what follows whenever I will talk about phenomenal properties, elaborate on what further peculiar feature the might or might not have, and so on I will have the kind of properties captured by definition 2 in mind.

[13]Towards the end of the work I will express my sympathies for modal idealism (see fn. 3 of this section), which entails (admittedly, very counter–intuitively) that, since necessarily every fundamental property is a phenomenal property, any state of the universe at any moment involves the instantiation of some phenomenal properties, hence, strictly speaking, nothing is ever unconscious. Yet, again, at the present stage an intuitive construal of subjects as entities that, although necessarily capable of instantiating phenomenal properties, may be unconscious at some points in time, will be sufficient.

Arguably, definition 2. meets all the *desiderata* listed above. Although, as I said, one might come up any sort of weird feature to subsume all the positive examples without admitting any of the negative ones, it seems clear to me that there are two features which, when combined together, would strike almost anyone as the most obvious ones to do the job[14]. Definition 2 simply states that a given property is phenomenal iff it has those two features, namely subjectivity, understood as the property of being such that, necessarily, whenever a property of this sort is instantiated a subject is involved, and phenomenality, understood in the very minimal sense whereby the instantiation of the property makes a difference for the subject involved from her/his own first–person perspective: there is a difference, subjectively speaking, between being in pain and not being in pain, between feeling nauseous and consciously entertaining the thought that tomorrow is Friday, between smelling coffee and having an orgasm ... .

Even understood in such a minimal sense, though, subjectivity and phenomenality raise all sorts of deep philosophical puzzlements, hence I would say definition 2 is pretty substantive. Yet honestly I could not come up with any less theory–laden definition. Definition 2 does not imply, *per se*, that phenomenal properties are irreducible to physical ones, nor that they have any potentially problematic feature like primitivity, ineffability, and the like. At the same time, neither does definition 2 imply the opposite, namely that phenomenal properties are reducible to(/grounded on/identical with/ metaphysically supervenient upon/realized by ...) a kind of physical ones and/or do not have the above features, or others. Thus, as far as I can see the definition meets *desideratum* 4 as well, in that it allows one to *wonder* whether phenomenal properties are physical or not, whether they are intrinsically private or not, whether they ineffable, *et cetera*, thus capturing, although in a very superficial and fairly non–committal way, the sense of puzzlement and of mysteriousness with

[14]Notice that the fact some philosophical technical notions such as phenomenal consciousness, phenomenology, phenomenal character, what–it's–like–ness and the like are particularly well–suited to capture the feature anyone, I believe, would latch onto reading my positive examples does not imply that laypeople need master any of those technical notions in order to be in a position to latch onto one and only one natural target reference for those examples.

respect to a naturalistic viewpoint consciousness is very often thought to elicit.

It may be the case that eliminativist (Dennet, 1978; 1988; P. M. Churchland, 1981; 1988; P.S. Churchland, 1986) and so called illusionist (Frankish, 2012; 2016; Kammerer 2021; Shabasson 2022; Pereboom 2011, 2017; Graziano 2013, 2016; Humphrey, 2016) positions with respect to phenomenal consciousness are the only ones that are ruled out by the definition I am buying into. I am a phenomenal realist. I do think that phenomenal consciousness exists, and that it has some very peculiar features, most notably the fact that its core nature or essence (see 1.4 for a definition of essentiality and of essential properties) is revealed to subjects in introspection. I also do think that both illusionism and eliminativism – on the assumption that there is a difference between the two views, which is something I'm not sure about – are obviously false[15]. Yet, since at least for the time being I am not attaching any particular theoretical or technical meaning to terms like 'phenomenology' 'what–it's–like–ness', 'phenomenal character' and the like, I think readers with sympathies for so called weak illusionism (Pereboom 2011, 2017; Graziano 2013, 2016; Humphrey, 2016) may continue to read my essay with no harm. In fact, weak illusionists acknowledge the existence of phenomenal consciousness, but deny it has certain features it appears to have such as intrinsicality, non–physicality, non–functionality, non–representationality, non–intentionality, and so forth. Once again, however, Definition 2 does not imply that phenomenal properties have (or do not have) any of the aforementioned features.

[15]By the way, as both Schwitzgebel (2016) and Nida–Rümelin (MS; 2016) argue, illusionism may rest on a misconception of how phenomenal properties acquire their reference in the very first place. For reference to phenomenal properties is not (or should not be, at least) established via any theoretical assumption or stipulation. Rather, it is (or should be) established ostensively, as I have tried to do. Once reference to phenomenal properties has been established in the right way, it is dubious whether illusionism can really stand. First, the illusionist may say that none of the states mentioned in the list of positive examples has ever occurred. But this is trivially false: to say that no one has ever been in pain, or angry, or that no one has ever perceived the color of a painting is simply absurd. The illusionist might then say that although the positive examples do pick some existing states or properties, the feature or features they have in common is incompatible with whatever prior understanding of the nature of phenomenal properties we may have. But this strategy won't work, for the list of positive examples is meant to *introduce* phenomenal properties, so there is no prior understanding to refer to. The only strategy left available to the illusionist is to claim that there is no feature or set of features which all the positive examples have while the negative ones lack. However, I am confident that, going back to my list of positive examples, most readers will agree with me that it is dubious, to say the least, whether this strategy can be profitably pursued. At any rate, I won't pursue this argument any further at this stage. I advise readers to have a look at the works cited in this footnote for further elaboration on these points.

As for strong illusionists and/or eliminitavists (Frankish, 2012; 2016; 2017; Kammerer 2021; Shabasson 2022) – namely, those philosophers who contend that phenomenal consciousness is an illusion but nothing like it exists, as far as I can see they might still accept definition 2 but deny it picks any existing property, arguing that the class of phenomenal properties is in fact empty. In other words, strong illusionists may read my essay hypothetically: *if* phenomenal properties, defined as above, did exist, then – provided the arguments I will draw throughout the essay are correct – certain conclusions would follow. *But*, they would continue, phenomenal properties *do not* exist, hence no (problematic) conclusion follows.

At any rate, let me underline that it won't be my aim, in this work, at least for the most part (but see the conclusion), to disprove illusionism, nor any other version of physicalism that is on the market[16]. Rather, what I intend to do is mostly to show what is it that physicalists of any variety should deny in order to secure themselves from the three most renowned and most discussed anti–physicalist arguments that have been offered in the last decades, namely Chalmers' conceivability argument, Jackson's knowledge argument and Kripke's modal argument.

To conclude this section, Let me elaborate a bit on why the definition of phenomenal properties I am buying into is more parsimonious than possible alternative definitions that may be offered (on this see Nida–Rümelin, MS). The first of the two conditions stated in the definition – *i.e.*, There is a specific way W such that, if P is instantiated a time t, then it is like W for a subject S to live through time t – might be read as an existential quantification over ways it like for subjects to instantiate given phenomenal properties. This is not the correct reading, though. For as the second condition states, the specific ways it is like for a subject to instantiate given phenomenal properties are not something over and above phenomenal properties themselves; rather phenomenal properties *are*, *consists in* ways it is like for a subject to live through fragments of time, and experiences, under this account, are nothing

[16]As my dear friend Arianna Beghetto once told me, illusionism without physicalism would be a lame creature. I won't consider possible non–physicalist illusionist views, for I honestly cannot see any reason for why on earth one should embrace such a weird view. To my eyes, saving physicalism is *the* main motivation behind illusionism. See Beghetto (PhD).

over and above instantiations of phenomenal properties over spans of time.

Thus, the characterization of phenomenal properties I am adopting is more parsimonious than the one that is implied under what Nida–Rümelin (*Ibid.*; 2018) calls the *experience–property framework*. On the latter framework, the fact that experiences feel in a certain way to subjects is captured by saying that experiences have certain properties, typically referred to as *qualitative characters* or *qualia*. Now suppose one adopts an account of subjects as "collections" or "sums" of experiences, possibly held together by a suitable relation of co–consciousness, which happens to be the view on the market I am most sympathetic with. Under the experience–property framework, experiences *have* phenomenal properties, hence on this framework there are two categorically distinct kinds of entities, namely the experiences in which subjects essentially consist in, and their phenomenal properties, or qualia. By contrast, under the framework I am buying into experiences do not *have* phenomenal properties: rather, they *are* instantiations of phenomenal properties. Hence in this picture only one kind of entity is postulated, namely phenomenal properties themselves, whose instantiation essentially constitutes the experiences subjects, in turn, essentially consist in. But even if one adopted whatever possible alternative account of experiencing subjects – say, one in which they are taken to be immaterial souls, brains, living organisms or whatever else – the characterization of phenomenal properties I am buying into would still be more parsimonious than its rival approach, for while under the experience–property framework there would be three categorically distinct kinds of entities, namely subjects, the experiences of subjects, and the qualia of experiences, under the kind framework I am buying into there are only two, namely subjects and phenomenal properties, whose instantiation (by subjects) essentially constitutes the experiences of those subjects[17].

[17]Besides these considerations of parsimony, Nida–Rümelin (MS, 2018) also argues that there are several further reasons to prefer a version of the subject–property framework over the experience–property framework. First of all, as she argues, the subject–property framework is conceptually more fundamental than the experience–property framework, meaning that while every technical term of the latter needs some technical term of the former in order to be introduced, no technical term of the former needs any technical terms of the latter to be introduced. What is more, Nida–Rümelin (*Ibid.*) also contends that the experience–property framework (1) leads to pseudo–intuitions; (2) invites misconceptions of relevant phenomena; (3) tends to distract a theorist's attention from

1.2 What is physicalism?

When it comes to the nature of consciousness and of phenomenal properties, physicalism[18,19] is still the prevailing attitude among contemporary philosophers, and not just among them: according to the last philpapers survey[20], carried out in 2020, out of the participants, who were selected among English–speaking academic philosophers from all–over the world, 51.9 percent inclined toward physicalism, 32.1 percent leaned towards anti–physicalism, and 15.9 percent chose the option labeled as "other". The work of a number of authors in the eighties and the nineties of the last century pave the way for some anti–physicalist positions to be rehabilitated, and those positions are now deemed as at least worthy of being taken seriously alongside physicalist alternatives[21]. However, physicalism is still mostly taken for granted, especially in the philosophical debate that is not explicitly concerned with the nature of conscious phenomena as well as among neuroscientists, psychologists and other researchers with empirically–oriented postures. Overall, it seems fair to deem physicalism as the the *Weltanschauung* of our time (Gillett and Lower 2001, ix; Tomasetta, 2015; Zanotti, PhD; Tomasetta, 2015_{a})[22].

the core of certain problems; (4) tends to blind a theorist for important insights; (5) leads to conceptual confusion; (6) makes it impossible in some cases to (adequately) express claims which merit discussion; (7) tends to conceal deep connections between interrelated topics. It is beyond the scopes of the present essay to argue for or against any of these claims properly.

[18]This section is largely drawn and adapted from Tomasetta (2012: 18–36; 2015_{a}) and Zanotti (PhD). I also advise readers to look at Zanotti (2020; 2021; 2022) for an in–depth assessment of the major arguments for physicalism.

[19]In this work I may sometimes use the terms 'physicalism' and 'materialism' interchangeably. To my knowledge, the use of those terms as synonyms is widespread, especially in the literature on the mind–body problem (see Stoljar, 2010). However, let met stress that I do not whish to imply that according to physicalists the fundamental level of reality of is inhabited only by material entities, for while it seems fair to say that everything that is material is physical, whether the opposite is true is contestable, in that entities like, e.g., point particles, photons, and electromagnetic fields are also postulated in some our best physical theories (see Zanotti, PhD: 39; Brown and Ladyman: 2019). Let me also stress that in this work I will exclusively be concerned with metaphysical doctrines concerning *consciousness*. The view that everything in the universe is physical won't be considered, even though it is clear that on an overall physicalist worldview only physical properties carve nature at its joints.

[20]https://survey2020.philpeople.org/

[21]see fn. 3 and fn. 4 of the introduction

[22]This being the case, one would expect that there be several compelling reasons to prefer physicalism over possible alternative metaphysical views. Yet as Daniel Stoljar (2017) argues, the lines of argumentation in favour of physicalism are surprisingly low. This may precisely due to the fact that, physicalism being the *Weltanschauung* of our era, philosophers oftentimes deem it

Since most of this work will be concerned with physicalism, a more precise characterization of this doctrine ought to be offered. First of all, let me be clear on what this work won't be about. I will be exclusively be concerned with *metaphysical* versions of physicalism. Physicalism as a research program (Elpidorou and Dove, 2018) and physicalism as an attitude (Ney, 2008_{b}) won't be considered. Metaphysical forms of physicalism postulate the existence of a metaphysical relation of some sort between phenomenal properties and physical ones. However, although, as I have said, there is widespread consensus among contemporary philosophers that some metaphysical form of physicalism ought to be true, there is still disagreement, perhaps surprisingly, concerning the precise nature of the relation purportedly holding between phenomenal properties and physical ones, as well as on the nature of the *relata* themselves. Among the numerous possible alternative ways to construe the purported metaphysical relation holding between phenomenal properties and physical

sufficient to take it for granted that some version of it ought to be true, without feeling the need to offer any sophisticated or compelling argument for the view. Tomasetta (2015_{a}) examines three among the major arguments that have offered for physicalism, namely the argument from the success of science, the argument from the causal closure of the physical and the argument from methodological naturalism, and concludes that none of them survives careful scrutiny. What is more, the feeling of the arguments for physicalism being low and not always entirely convincing is shared even by some philosophers with physicalist sympathies. For instance, a physicalist such as William Lycan argues that the standard arguments for physicalism are just as weak and unconvincing as those for and against dualism are. Here is what he (2009: 551) writes: «Being a philosopher, of course I would like to think that my stance is rational, held not just instinctively and scientistically and in the mainstream, but because the arguments do indeed favour materialism over dualism. But I do not think that, though I used to. [...] I do not proportion my belief to the evidence». At any rate, let me stress that this work won't be devoted to an assessment of the arguments *for* physicalism. Rather, I will be mostly if not exclusively concerned with the argument *against* physicalism. More specifically, I am going to make the case that the three most discussed and renowned anti–physicalist argument that have been offered in the last decades – namely Chalmers' (e.g., 1996; 2009; 2010_{a}) so called conceivability argument, Jackson's (1982) so called knowledge argument and Kripke's (1980) so called modal argument – may be reduced to the argument from revelation, as hinted at in the introduction and to be described more accurately below.

ones there are *type–identity*[23], *realization*[24], *constitution*[25], *grounding*[26], and possibly *weak emergence*[27] (but there are certainly more): offering a detailed overview of the literature concerning each and every one of these theoretical options transcends by far the scopes of this work, as some theories on the market are indeed very baroque and sophisticated. Yet a common denominator of all or most metaphysical forms

[23]Type–identity physicalism rose in the sixties as a reaction to the behaviorist approaches that were predominant at the time. The historical references for this family of views are Place (1956), Feigl (1958), Smart (1959) and Armstrong (1968). As its name suggests, type–identity physicalism claims that types of experiences, say painful experiences, are identical to types of physical states – most likely, types of neurocortical activity, say, C-fibers firing. Notoriously, type–identity physicalism suffers from an objection raised by Putnam (1967) and Fodor (1975) known as the objection from *multiple realizability*. Roughly, the objection goes as follows. If type–identity physicalism were true, then every token of a certain experiential type would be identical to a token of a certain type of neurocortical activity, for type–identity implies token–identity (whereas the opposite is not true). It seems clear though, that organisms with nervous systems significantly different from our own, including organisms with no C-fibers in their nervous system, can and do instantiate the experience of pain, and the same point may be reiterated for (almost) any type–of–experience/type–of–neural–activation pair. Hence, the objection concludes, type–identity physicalism cannot be true. Needles to say, there have been counter–objections to the objection from multiple realizability. It has been claimed that there may be disjunctive kinds (but see Fodor, 1974, and Kim, 1992, 2005, for a critique). Also, Lewis (1969) has suggested that types should be construed as domain–specific. The problem with these latter approaches is that they may make type–identity collapse on token–identity. Overall, the objection from multiple realizability is usually regarded as a very serious objection, if not fatal, to type–identity physicalism. That is why nowadays most physicalist philosophers tend to prefer alternative, non–reductive varieties of physicalism.

[24]Realization physicalism claims that mental properties, states, processes etc. are either physical or physically realized. The prototypical examples of realization physicalism are Andrew Melnyk's (2003; 2006; 2018) view and Sidney Shoemaker's (2001; 2007, though a version of the subset view Shoemaker defends had originally been presented in Wilson. See, e.g., her 2005, fn. 13). I cannot afford here to get into the details of how the notion of realization is construed under these accounts.

[25]The most prominent defense of constitution physicalism is the one put forth by Derk Pereboom (2011). Needles to say, constitution physicalism claims that physical properties constitute phenomenal ones. The relevant relation here at play is material constitution, understood as a primitive – in that it cannot be reduced to more primitive relations – irreflexive asymmetrical relation of metaphysical dependence. The prototypical example of material constitution is the statue/lamp–of–clay one. Wasserman (2021) provides a detailed overview of the notion of material constitution

[26]The resurgence of the notion of grounding in the contemporary debate has mostly been due to the work of Fine (2001), Schaffer (2009_{a}), Rosen (2010), and Audi (2012). Grounding physicalism about the mental is discussed in Dasgupta (2014) Kroedel and Schulz (2016) and O'Conaill (2018), among others. It is typically said that the relation of grounding is captured by the expression 'in virtue of': what is grounded obtains in virtue of what grounds it. Grounding is thus construed as an irreflexive, hierarchical and asymmetrical relation of metaphysical dependence: if A grounds B – *i.e.*, B obtains in virtue of A – then A is more fundamental than B. Although grounding physicalism promises to offer a clear, explicit and hierarchical account of how phenomenal properties may be nothing over and above a kind of physical properties though somehow distinct from the latter, it should be underlined that the very notion of grounding is admittedly rather controversial: there is no unanimous consensus of what the *relata* of the grounding relation are, or should be. Moreover, some philosophers are skeptical about the very notion of grounding (e.g., Hofweber 2009; Daly 2012) or argue that there is no point in introducing a distinctive relation besides the ones we already have available such as constitution, token–identity, supervenience and the like (Wilson, 2014).

[27]Chalmers (2006_{b}: 1) characterizes the distinction between strong and weak emergence as follows:

of physicalism may be individuated. This common denominator consists in the combination of a negative *desideratum* and a positive one. Let us have a look at each *desideratum* in due order:

> *Positive* desideratum: According to this *desideratum*, in order for physicalism to be true phenomenal properties must be metaphysically supervenient upon physical ones, meaning that if a metaphysically possible world W is an exact minimal physical duplicate of the actual world, then W must be an exact duplicate of the actual world with respect to its conscious properties, namely a C–duplicate of the actual world[28]. In other words, if physicalism is to be true then for every phenomenal property M instantiated at our world there must be a physical property P such that for every metaphysically possible world W* if P is instantiated at W* then M is instantiated at W* (see Stoljar, 2010: Ch. 6). This constraint on the metaphysical strength of the supervenience relation is crucial when it comes to distinguishing physicalism from alternative – e.g., dualist (Chalmers, 1996, just to make one example among a plethora of others) – positions which postulate that the conscious is metaphysically distinct from the physical though nomologically supervenient on it in virtue

«We can say that a high–level phenomenon is strongly emergent with respect to a low–level domain when the high–level phenomenon arises from the low–level domain, but truths concerning that phenomenon are not deducible even in principle from truths in the low–level domain. [...] We can say that a high–level phenomenon is weakly emergent with respect to a low–level domain when the high–level phenomenon arises from the low–level domain, but truths concerning that phenomenon are unexpected given the principles governing the low–level domain». As far as I know, Chalmers characterization of the distinction between strong and weak emergence is pretty standard (see Kim, 1998:8; Van Cleve 1990; Nordhoof 2010). While strong emergence is clearly incompatible with a physicalist worldview, weak emergence may be compatible (with some further clarifications to be made) with at least some of the other metaphysical relations mentioned above – e.g., grounding, constitution *et cetera*.

[28]This is Jackson's (e.g., 1998) version of physicalism. The addition of the 'minimal' bit is is meant to avoid the so called problem of epiphenomenal ectoplasms, namely, pure phenomenal entities of some kind which do not interact causally with anything else there is in a given possible world. A minimal physical duplicate of the actual world is a world which duplicates all the entities of of the actual world physics deals with *without adding anything else*. According to Lewis, (1983_b) the positive *desideratum* is already sufficient for what he deems as a minimal form of physicalism. Morris (2014) also argues that the metaphysical supervenience of phenomenal properties upon physical ones is sufficient for physicalism to be true. In general, there is widespread agreement that physicalism should entail at least the metaphysical supervenience of the mental upon the physical (among others, see Jackson, 1998; Witmer, 2001; Chalmers, 1996, 2010; Levine and Trogdon, 2009; O'Conaill, 2018). To my knowledge, only Montero (2013) and Montero and Brown (2018) deny that supervenience is necessary for physicalism (but see Alter, 2021 for a reply). Overall, though, it is fair to say that the view they sponsor is very minoritarian in the debate.

of some lawful connections – e.g., Chalmers' (e.g., 1996) psycho–physical laws – holding between physical and phenomenal facts at our world. According to these latter positions, although it is true that, in our world, whenever there is, say, human pain there is such–and–such pattern of neuro–cortical activity, it is still possible that a given world W may be an exact minimal physical duplicate of our world yet not a C–duplicate of it – *i.e.*, so called phenomenal zombies (see the second chapter) are metaphysically possible.

Negative desideratum*:*Although there is widespread consensus among philosophers with physicalist inclinations that the truth of physicalism requires at least the metaphysical supervenience of phenomenal properties upon physical ones, many (see, Horgan, 1993 and Wilson, 2005 among many others) acknowledge that this is not enough. For consider a view such as strong emergentism. Chalmers (2006_b: 1) defines strong emergentism as follows: «We can say that a high–level phenomenon is strongly emergent with respect to a low–level domain when the high–level phenomenon arises from the low–level domain, but truths concerning that phenomenon are not deducible even in principle from truths in the low–level domain». Depending on whether one interpret the 'arises' bit metaphysically or only nomologically, Chalmers' definition of strong emergentism may be compatible with the metaphysical supervenience of the phenomenal upon the physical. Yet whether strong emergentism counts as a form of physicalism is controversial to say the least. Indeed, many would deem strong emergentism as an overtly non–physicalist thesis (this point has been made vivid by Horgan, 1993[29] and Crane, 2010[30], among many others). That is why most metaphysical versions of physicalism that are on the market

[29]«... [for the emergentist] all properties could be supervenient on physical properties [...] even if certain supervenience facts are metaphysical *sui generis*, unexplainable in more fundamental terms. Yet a materialistic metaphysical position should assert that all supervenience facts are explainable in some materialistically acceptable way». (Horgan, 1993: 557)

[30]«[...] the connection between the physical and the mental is a brute, inexplicable (metaphysical) necessity [...] this is the characteristic thesis of emergentism. [...] (ER) All truths can be explained in principle in terms of broadly physical truths. The emergentist of course must deny (ER) [...] for to deny (ER) is to hold that [...] the connection between the physical and the mental is a brute, inexplicable necessity [...] this is the characteristic thesis of emergentism [...]. Physicalist cannot reject (ER) on the pain of abandoning what is essential to physicalism.» (Crane, 2010: 25).

are either implicitly or explicitly committed to a further, negative *desideratum* (See Wilson, 2006, Brown, 2017 and Montero and Brown, 2018, among many others). According to this second *desideratum*, in order for physicalism to be true there cannot be fundamental experience–involving entities. Besides strong emergentism, this negative *desideratum* allows to exclude any dualist position, as well as idealism, so–called Russelian Monism (Russell, 1927; Eddington, 1928; Alter and Pereboom, 2023; Goff and Coleman, 2020)[31] *et cetera*.

Having already provided a working definition of phenomenal properties, I now wish to say something on how physical properties and/or entities ought to be construed.

As far as I know, nowadays most physicalist philosophers adopt, either explicitly or implicitly, some version or another of a theory–based approach whereby given entities(/properties/states/processes/...) count as physical if they are postulated by some of our best physical theories. Notoriously, elegant and widespread as this approach may be, it is vulnerable to what is known as *Hempel's dilemma*. Suppose – first horn of the dilemma – we appeal to the best physical theories we currently have available to fix reference on physical properties. The problem with this strategy is that it may make physicalism turn almost certainly false, for even if we embrace a moderate optimism whereby scientific research gets asymptotically closer and closer to the truth as time moves on, we have no reason to suppose that current physical theories won't be revised in any aspect. Indeed, everything strongly suggests the opposite.

The other horn of the dilemma, though, has its own issues. On this horn, physical properties and/or entities are those and only those that figure as the predicates of a future, ideally complete physical theory. The fact now is that this strategy may make physicalism a rather vague and mysterious position, for admittedly we have no clear understanding of how the properties, entities, *etc.* posited by an ideal future physical theory would look like. What is more, this strategy may make physicalism trivially true, for as Ladyman (2011; see also Zanotti PhD) argues, no metaphysical

[31] See also Feigl (1958), Maxwell (1979), Lockwood, M. (1989), Strawson (1994; 2003; 2008_a) and Chalmers (1996) among others

framework that is compatible with the empirical data should be *a priori* excluded, and a number of utterly non–physicalist views – e.g., dualist and panpsychist ones – are, as a matter of fact, compatible with empirical evidence. Indeed, there may be principled reasons to be skeptical that empirical evidence will *ever* adjudicate the mind–body problem (see, e.g., Papineau, 2002: Ch. 7; Zanotti, 2022; Chen–Wei Ng: 2021).

In general, most philosophers would be uncomfortable, and rightly so, with committing themselves to a view which will most likely turn out as false at some point. Hence, they tend to avoid the first horn of the dilemma (but see Melnyk, 2003 for a relevant exception) and prefer the second horn, biting the bullet and admitting that we are not (and may never be) in a position to offer a precise definition of what counts as a physical entity(/property/state/process . . .). However, some options are available, if not to entirely avoid, at least to mitigate the charge of vagueness. The first option consists in trying to fix some minimal constraints on how the entities posited by an ideal future physical theory should look like. For instance, one might say that such a theory should at the very least deal with the most fundamental entities and properties of our world (see 1.3 for a definition of fundamentality), or the it should at least allow to make empirically testable predictions, *et cetera* (See Ney 200; Elpidorou and Dove, 2018: Ch. 3); Dowell, 2006, 38–39; Zanotti, PhD). At any rate, the prevailing strategy among physicalists on this regard consists in the so called *via negativa*, namely, in imposing constraints on what *cannot* count as physical. In turn, most philosophers who resort to the *via negativa* impose a no–fundamental–mentality constraint (see Wilson, 2006; Worley, 2006) on their theories, as per the negative *desideratum* of conventional physicalism seen above. Indeed, some (e.g., Spurrett and Papineau 1999; Montero 2001; Montero and Papineau 2005) avoid making any explicit appeal to theoretical physics and simply equate the physical with the non–fundamentally–mental. As already said, buying into a no–fundamental–mentality–constraint also allows to exclude a number of overtly non–physicalist positions such as dualism and panpsychism, compatible

with empirical evidence as they may be, whereby the charge of triviality is arguably avoided.

As I hope to have managed to show even with this admittedly very brief overview of some of the approaches that are available, there is no univocal understanding of how physical properties, as well as physicalism as a doctrine in general, ought to be construed. However, in the following I will work under certain assumptions. In virtue of the remarks just drawn, I think it is fair to say that these assumptions may accepted by a large number of physicalist philosophers. The first of my working assumptions is drawn from Wilson (2006:72): a given entity is physical iff «it is treated, approximately accurately, by current or future (in the limit of inquiry, ideal) versions of fundamental physics». Moreover, I will work under the assumption that in order for physicalism to be true (at least some of) the properties(/states/processes/entities/...) physics deals with should be fundamental. Finally, I will assume that in order for physicalism to be true fundamental physical entities cannot be essentially experience–involving.

1.3 What are fundamental properties?

Fundamental properties will be defined as those properties such that, necessarily, they are not instantiated in virtue of anything else. That is to say, a given property P will be said to be fundamental iff the fact that X is P does not hold in virtue of any other fact[32,33]. More in general, the notion of fundamentality has oftentimes been understood in terms of dependence (see Tahko, 2023: sec. 1). Given a relation of dependence D, an entity/fact/property/... is said to be D–fundamental iff it does not D–depend on anything else in order to occur/be instantiated/exist/... (Tahko,

[32]This section is mostly taken from Builes (forth: 2–3). As far as I know, the taxonomy that, following Builes, I'm adopting here is widely accepted.

[33]See Builes (forth.). See also Rosen (2010) and Audi (2012) for a thorough examination of the 'in virtue of' relation. Other philosophers such as Sider (2011) and Wilson (2016) argue that the notion of fundamentality is primitive in that it cannot be reduced to other notions such as the 'in virtue of' relation. Readers with sympathies with this latter approach may still continue to read my essay, though, as I think that a broad intuitive understanding of what it takes for a given property to be fundamental should suffice for my purposes.

Ibid.). It is not entirely clear what kind of "things" could be said to be fundamental. Among the possible candidates there are objects, facts, and properties. I cannot afford here to get into the detail of the various theories of fundamentality that are available. Most often, I will speak of fundamental properties. Every now and then, I might also speak of fundamental entities. At any rate, as Tahko (*Ibid.*) notes, an acceptable translation of one view into another is often possible. For instance, assumed that facts are instantiations of properties, to say that the fact that 'X is P' or 'all X's are P' is fundamental might be translated as saying that P is a fundamental property. Likewise, defined fundamental entities as those entities that do not depend on anything else to exist, we might say that the non–relational properties of those entities are fundamental.

Fundamental *one–place* properties apply to a single item to be instantiated, whereas fundamental *relations* – or fundamental n–places properties – require multiple items to be instantiated. Typically, fundamental one–place properties are taken to be *intrinsic*, meaning that if a property P of this kind is instantiated, than the fact that X is P does not hold partly in virtue of any other entity besides X itself and/or some of its parts (see Rosen, 2010, for this definition of "intrinsicality". See also Builes, forth.). In this essay I will be concerned with so called *non–hacceitistic* properties, namely, properties that are not bound to a particular individual object. Properties such as 'being Bruno Cortesi' are called *hacceitistic*, in that they are bound to an individual entity – me. By contrast, properties like 'being tall', 'being red' or 'being in an hangover' are non–hacceitistic, in that they are not bound to a particular individual entity (see Builes, forth: 2). I take it that both phenomenal properties and physical ones are non–hacceitistic in this sense. A further distinction that will be of interest in the essay is the one between *dispositions* (or dispositional properties, interchangeably) on the one hand, and so called *quiddities* on the other. We may understand dispositional properties as those properties that inherently or essentially ground certain dispositional powers of their bearers. That is to say, the dispositionality of a property is the influence the property has on the behaviors

exhibited by an object that possesses it, based on specific stimuli. For instance, the fragility of an object is a dispositional property, in that its possession by that object grounds its power to be broken when hit (see, e.g., Taylor, 2013, among many others). Quiddites, by contrast, are those properties that we need not identify by the dispositional powers they confer on their bearers. In turn, the space of quiddities may be divided into so–called *bare quiddities* and so called *qualitative quiddities* or, for brevity, *qualities* (interchangeably) (see Hildebrand, 2016; Builes, forth.). Bare quiddities are defined as those quiddities whose nature is exhausted by whatever other quiddities they are identical or not identical with, whereas the nature of qualitative quiddities is not exhausted by which other quiddities they are identical or not identical with. The idea, here, is that qualitative quiddities have an underlying substantial nature which bare quiddities lack (*Ibid.*).

Dispositions and qualitative quiddities are oftentimes considered to be distinct kinds of properties. However, later on – especially in 3.5 – I will elaborate on the so called *powerful qualities view*, of which Martin and Heil (Martin, 2007; Martin and Heil 1998; 1999; Heil, 2003, ch. 11; 2010; See also Strawson, 2008_{b} and Taylor, 2013) are among the main advocates. According to such a view, qualities and dispositions are identical. That is to say, each dispositional property is qualitative and *vice versa*. Hence, on this approach qualities and dispositions are not distinct properties of given entities/properties/..., but rather distinct ways of considering what it is or would be for an entity/property/... to be instantiated, or distinct ways of characterizing one and the very same entity/property/...: it is not as if given entities/properties/... were both qualitative and powerful: their being qualitative *just is* their being powerful, which implies, of course, that their being powerful just is their being qualitative. I will suggest that, when combined with the thesis of revelation (to be presented below), the powerful qualities view may lead to a form of idealism, whereby necessarily every fundamental entity is essentially phenomenal, hence experience–involving. Idealism happens to be the view I lean towards. This is no matter to be treated here, though.

1.4 What are essential properties?

As we shall see, the thesis of revelation claims that phenomenal concepts (to be characterized presently) provide a subjects with cognitive access to at least part of the essence (or nature, interchangeably) of their referents, namely phenomenal properties. Hence, a definition of essentiality and of essential properties should now be provided. For the purposes of the present essay, I think a rather broad intuitive understanding of what the essence of something is should suffice. Nonetheless, let me elaborate a bit.

The essence of a given item will be defined as the set of all and only its essential properties. Since Fine published his (1994) (but see also his $1995_{a,b}$) and Kripke his (1980) the notion of essence has regained a prominent role within the contemporary metaphysical debate.

There are basically two main ways to construe the notion of essentiality. According to a modal account of essentiality – of which Barcan Marcus (1967) and Kripke (1980) are among the main sponsors[34] – and entity E has a certain property P essentially iff it has it necessarily, namely in all possible worlds, or at least in all the possible worlds in which it exists[35]. The relevant notions here at play is *metaphysical* necessity, not *logical* necessity, *physical* necessity *et cetera*. *Prima facie*, the modal account of essentiality aligns well with our intuitive understanding of what the essence of something is, which often boils down to the idea of some given property being *indispensable*, *required*, *necessary* for an entity to exist (see Robertson Ishii and Atkins, 2023: sec. 2). Notwithstanding, the modal account may have some

[34]Zalta (2006), Correia (2007) and Brogaard and Salerno ($2007_{a,b}$; 2013) have offered non–standard versions of the modal account in light of Fine's (1994) (purported) counterexamples to the classic version of it.

[35]The restriction of the class of relevant possible worlds to those and only those in which E exists is meant to avoid the following objection to the modal account. Necessarily, an entity has to exist in a certain possible world in order to instantiate a certain property in that world. Now take a contingently existing entity in a certain possible world, say a dog in our world, call her Emma (I'm taking this example from Robertson Ishii and Atkins, 2023: sec. 1): Emma exists in our world, but it could not have been born. From this seems to follow that the property of Emma of being a dog is not an essential property of her, for all the possible worlds in which she does not exist are world in which, trivially, she does not instantiate that property. Intuitively, though, 'being a dog' is, indeed, an essential property of Emma. Hence the restriction of the class of admissible possible worlds.

unfortunate implications: for instance, it may lead to the conclusion that the property of being such that there are infinitely many prime numbers is an essential property of any arbitrary entity, or that the property of being the sole member of the single set {2} is an essential property of the number 2 (*Ibid.*).

That is why several philosophers – most notably Kit Fine (1994; $1995_{a,b}$) – have opted for a primitivist/non–modal/definitional account of essentiality. According to this latter approach, the essence of an entity E is what makes E the entity it is or, equivalently, comprises E's most core respects. E, thus, will be said to have a certain property P essentially iff P belongs to E's most core respects, namely to the class of properties E could not cease to have without ceasing to be entity it is.

Along with many others[36], I believe the notion of essence is primitive and not analyzable in terms of other, simpler notions. Hence, I prefer a non–modal approach. However, readers that happened to be sympathetic with a modal approach for independent reasons could continue to read the essay with no harm. In fact, it is widely agreed that if a given property P is essential to an entity E in the Finean sense P is also essential to E in the modal sense. For it seems clear that an arbitrary entity must possess the property that make it the entity it is in all the possible worlds in which *it* exists. As I have said, whether the opposite is true may be contested. Hence, readers may continue to read the work no matter the approach they prefer.

Let me address a worry which may be raised at this point. Some might feel that the way in which I have defined experiences and phenomenal properties trivially implies that the way in which a given experience feels to a subject is an essential property of the experience itself. In fact, according to our definition an experience – namely, the instantiation of a given phenomenal property P at time t – consists in the fact that there is a specific way W such that it is like W for a subject S to live through time t (which, let me repeat, does not imply *per se* that W has any potentially problematic feature, including non–physicality). Notoriously, the essence of X is spelled out by a real definition of X. According to the thesis of revelation (on

[36]Besides Fine, see also Hale (2013) and Lowe (2012), just to mention a couple.

which I will elaborate momentarily), phenomenal concepts – those concepts that are deployed to think about the experiences in terms of the way they feel to a subject (see 1.7) – allow subjects to grasp the (entire) essence of the experiences they refer to. In light of all this, an opponent might contend that in assuming the definition of experiences and of phenomenal properties provided in 1.1 I am begging the question against deniers of revelation.

Let me underline, though, that as far as I can see one might very well assume, especially at the present stage of the work, the definition of phenomenal properties I am adopting to be only a *nominal* definition but not as a *real* definition. Roughly, we might say that while a real definition spells out the essence of X, a nominal definition merely clarifies the meaning of the term 'X': the chemist aims at a real definition of water, the lexicographer at a nominal definition of it (see Gupta and Mackereth, 2021: 3–4).

1.5 Interlude: some historical background on concepts: intensions, extensions and Kripkean rigid designators.

The discussion of the thesis of revelation and its role within the mind–body problem debate will inevitably require devoting a considerable amount of space and effort to an investigation into the nature of both physical and phenomenal concepts (to be characterized below) and the relations holding between the two. Before getting to that, though, a very brief overview of how certain key–figures in the history of philosophy have thought about concepts in general might be of some use, I think, in shedding light on some of the issues I will touch upon throughout the essay.

Let me briefly walk the reader through the history of what Chalmers (2004: 153–8) calls *the golden triangle*. The golden triangle links three key philosophical notions, namely meaning, reason and modality. Among the figures who played a major role in drawing the golden triangle, three stand out: Kant (1781/1787[2009]),

Frege (1892) and Carnap (1947) (see Chalmers, *Ibid.*).

Kant drew the first side of the triangle, by linking reason and modality. Notoriously, he held that a proposition is *a priori* – *i.e.*, knowable independently of experience – iff it is necessarily true, and *vice versa*.

Frege drew the second side of the triangle, by linking reason and meaning. He noted that concepts have a dual nature[37], in that the entity to which a given concept refers – what is typically called the *reference*, the *denotation*, or the *extension* (interchangeably) of the concept – does not (or at least not always) determine its meaning, its cognitive significance, the role the concept plays within a subject's reasoning, inference and knowledge. The cognitive significance of a concept is called by Frege its *sense*, but it may also be called the *intension* of the concept. That two concepts may have the same extension though a different intension or sense is demonstrated by the renowned 'Hesperus–Phosphorus' example: while the concepts <HESPERUS> and <PHOSPHORUS> are co–extensive or co–referential – in that they both refer to the planet Venus – they have a different sense, for the sentence 'Hesperus is Phosphorus' is cognitively significant, meaning that it is not trivially true for a rational subject.

Finally, Carnap drew the third side of the triangle, by linking meaning and modality. He held that two concepts have the same intension – the same sense, or the same meaning – iff they are necessarily coextensive , that is to say, if the judgment 'A is B' – or 'A iff B' if A and B are propositions (Chalmers, *Ibid.*) – is necessarily true. By combining the Carnapian link between meaning and modality with the Kantian link between modality and reason, one obtains that two concepts A and B have the same intension iff the judgment 'A is B' – or 'A iff B' – is *a priori*.

[37]Truth be told, the recognition that concepts have such a dual nature predates Frege by a long way. In the so called Port–Royal Logic (Arnalud and Nicole, (1662[1996]) we find expressions which might be translated as the *comprehension* and *denotation* of a term. Likewise, Mill (1843,[1974]) distinguished between the *connotation* and the *denotation* a term. Offering a detailed overview of how these notions have been understood throughout the history of philosophy is far beyond the purposes of this essay. Melvin (2022) provides a useful introduction to intensional logic, with some historical references. At any rate, it is fair to say that the modern understanding of intensional issues has its roots in Frege's (1892) seminal paper and, more in general, that Frege played a pivotal role within the very long story which, following Chalmers (2004), I am trying here to make *very* short.

Famously, Kripke's (1980) and Putnam's (1975) seminal work in the seventies of the last century led to the breakage of the golden triangle. Though Kripke kept the Carnapian link between meaning and modality intact, he severed the Kantian link between reason and modality, thus severing the Fregean link between reason and meaning as well (Chalmers, *Ibid*). Kripke held that many of our concepts function as rigid designators, picking the same referent in any metaphysically possible world, which implies that any true identity statement involving two rigid designators is necessarily true. Not every identity statement is *a priori* and cognitively insignificant, though, for according to Kripke there may be – there are, indeed – necessary *a posteriori* and cognitively significant judgments – 'water is H_2 O' being one example.

Nowadays it is widely agreed that Kripke's arguments against the Kantian side of the triangle are right. Still, many people, myself included, believe that there must be something Frege was right about. In the second chapter I will elaborate extensively on Chalmers' (e.g., 1996; 2004, among other works) version of so called two–dimensional semantics, which is the core of his two–dimensional argument against physicalism and, more in general, promises to somehow restore the golden triangle while at the same time acknowledging Kripke's lesson.

Whether two–dimensional semantics, either in Chalmers' or in someone else's version, has actually succeeded, to this day, in restoring the golden triangle, or whether the golden triangle can indeed be restored somehow, are very important questions. In this work, however, I won't be concerned with them, nor will I take any explicit stance on them.

For the time being, let it be sufficient to say that in what follows I will work under two assumptions: (1) both physical and phenomenal concepts, both of which I am about to describe, function as Kripkean rigid designators; (2) different concepts may be co–referential and yet characterize a given entity differently, *i.e.*, they may have different modes of presentation or present their referent under different aspectual shapes. As far as I can see, both (1) and (2) are hardly controversial assumptions anymore (see, e.g., Holman, 2013, among others).

1.6 What are physical concepts?

I will adopt a very straightforward definition of physical concepts: these are the concepts the are deployed in our best physical theories to refer to physical properties, entities, states and processes – the latter having been characterized as in 1.2. Prototypical examples of these concepts include concepts such as <MASS>, <ELECTRIC CHARGE>, <BOSON>, <ELECTRON>, <PROTON> <GRAVITATIONAL FIELD>, and so on. One question of great relevance within the debate on the metaphysics of conscious phenomena is how physical concepts characterize their referents. According to a rather widespread, *dispositionalist*, conception of theoretical physics – of which Russell (1927)[38] is among the most prominent sponsors – physical concepts characterize physical properties structurally, functionally, dispositionally, relationally or dynamically (see 1.3 for a definition of dispositional properties). As we shall in 3.5, it is not clear what the relations between these notions are. That is to say, it is not clear whether if a given property is essentially dispositional it must also be purely structural or relational, and so on. For now, let it be sufficient to say that on a dispositionalist approach physics describe the world in terms of its spatiotemporal structure and changes within such structure but is silent on what underlies the structure itself (Alter and Pereboom, 2023). In broad brushstrokes, we may say that on such a conception physics tells us what physical entities(/properties/states/processes/...) *do* but says little or nothing of what they *are* (*Ibid.*). Part of this work will be devoted to an investigation of the role such a dispositional conception of physics plays within the debate on the metaphysics of consciousness.

Another distinct but closely related issue, is whether physical properties, as characterized under such a dispositionalist conception, can be said to be intrinsic, or not.

Later – especially in 3.5 – I will elaborate on the so called powerful qualities

[38]But see also Kant (1781/1787[2009]) and Eddington (1928) among *many* others.

view, of which Martin and Heil (Martin, 2007; 2008: Martin and Heil 1998; 1999; Heil, 2003, ch. 11; 2010; See also Strawson 2008_{b} and Talor, 2013) are among the main sponsors. As I've already mentioned, according to such a view, qualities and dispositions are identical. That is to say, each dispositional property is qualitative and *vice versa*.

At any rate, I don't want to rush into anything. For the time being, the minimal definition of physical concepts I have just provided will be sufficient. As far as I can see, such a definition may be accepted by readers who happened to be inclined to reject either a dispositionalist conception of theoretical physics or the powerful qualities view, or both, as well.

1.7 What are phenomenal concepts?

The debate on the nature of phenomenal concepts is an extremely tangled one. This is partly due to the fact that understanding what concepts are to begin with is one of the thorniest philosophical issues there have ever been, old at least as philosophy itself[39], and partly due to some peculiar feature or combination of features *phenomenal* concepts in particular have been claimed to have (or lack). Be as it may, a plethora of accounts of phenomenal concepts has been offered in the recent (and not so recent) literature, some of which are indeed very baroque and sophisticated: some philosophers (e.g., so called strong illusionists, see 1.1) believe that phenomenal concepts do not exist at all or are devoid of any content; others believe phenomenal concepts have all sorts of peculiar features[40]. Overall, it is fair to say that the debate

[39]Just to timidly begin to scratch the surface of the problem, we might distinguish at least three uses of the notion of 'concept' (see Sündstrom, 2011a: fn. 6). Under a first, roughly Fodorian approach (Fodor, 1975, 1988) concepts are conceived as internal mental symbols one uses to think with. Under a second, Fregean approach (on which I have elaborated *a bit* in 1.5) concepts are conceived as abstract entities. Under a third, less rigid approach, to employ a concept is simply to think about something, and to assert that someone possesses a concept implies that they have the capacity to think something. This usage doesn't commit to specific content or processes involved in thinking (see Byrne 2005: section 1.1, and Sundström 2011:section 2.4). At any rate, it is not my ambition here to offer an in–depth analysis of any of these approaches, nor on any other theory of the nature of concepts that may be offered.

[40]Among the most influential accounts of phenomenal concepts that have been proposed there are *recognitional* accounts (Loar 1990 / 1997, 2003; Tye 2000, chap. 2; 2003; Carruthers 2000, 2004; Perry 2001; Levin $2007_{a,b}$), *quotational* accounts Papineau 1993, chap. 4; 2002, chap. 4;

on the so called hard problem of consciousness (Chalmers, 1995) and the one on what phenomenal phenomenal concepts are (if they exist), overlap to a large extent[41].

A few of the main approaches that have been proposed to frame phenomenal concepts will be discussed throughout the proceeding of the dissertation. At this stage, I don't want to rush into anything. Yet some very general remarks may already be drawn.

First of all, I take it that it to be utterly uncontroversial to say that people can have all sorts of thoughts about their own and other subjects' experiences – instantiations of phenomenal properties. Recall that here I am adopting the very minimal characterization of experiences and of phenomenal properties offered in 1.1. I may hope that taking a painkiller will make the nasty headache I am now having go away, wondering whether <THIS> is what people mean when the say that they're in love with someone, wondering whether my friend liked or disliked the dinner we had together yesterday, thinking that my dog may be hungry, *et cetera*, *ad infinitum*.

Moreover, I take it that phenomenal properties may be referred to and thought about in a number of distinct ways, or characterized differently, or presented under different modes of presentation. To see this, suppose I am thinking about an experience of my friend Giulia, say I am entertaining a certain thought about her experience of seeing something red. Arguably, there are several ways in which I might think about this, hence several concepts that would yeld a true (or false) belief (see Chalmers, 2003_a sec. 2). I might, for instance, be thinking that (1) such–and–such surface–reflectance profile is affecting Giulia's retinas, or that (2) Giulia's brain cortex

2007; Balog 1999; 2011; 2012). and *indexical* accounts (Tye 1995, chap. 6; Lycan 1996, sect. 3.3). According to Levine (2001: 84) phenomenal concepts present their referents in a peculiarly substantive and determinate way. Chalmers (2003_a) and Gertler (2011) discuss a peculiar kind of phenomenal concepts, called *direct phenomenal concepts* (Chalmers, *Ibid.*) or *pure demonstratives* (Gertler, *Ibid.*) (see also Pallagrosi and Cortesi, forth, for a discussion of the role direct phenomenal concepts play in so called introspective knowledge by acquaintance). The present list is far from being exhaustive of all the theories of phenomenal concepts that are on the market. As I have said, the debate around the nature of phenomenal concepts is really a battlefield. Sündstrom (2011) provides a useful introduction to the notion of phenomenal concepts and related issues. Later, I will elaborate a bit on a few of the accounts mentioned in this footnote, and others.

[41]Here of course, I am referring to the so called *phenomenal concepts strategy* (PCS) to defend physicalism against typical anti–physicalist arguments (Stoljar, 2005; Carruthers and Veillet, 2007; Sündstrom, 2011a). Some versions of the PCS will be discussed later on in the essay.

exhibit such–and–such pattern of activity. What is more, I might be thinking that (3) Giulia is having the experience typically caused in most people in her community by paradigmatic red objects such as blood, tomatoes *et cetera* (Chalmers, *Ibid.*), or that (4) she is having the experience typically caused *in me* by paradigmatic red objects, or (5) that she is having an experience with such–and–such phenomenal character – the notion of phenomenal character, again, being here understood in the very minimal way described in 1.1.

I take it that (1)–(5) are all distinct ways to think about one and the very same type of experiential state. Yet some of them – (1) and (2) – do not require thinking about the phenomenal property in question – seeing something red – as such that necessarily there is something it is like for a subject to instantiate it, whereas others – (5), (4) and possibly (3) – do characterize the state in question as a state such that necessarily there is something it is like for subject to instantiate it.

To see that each of the concepts deployed in (1)–(5) characterizes their referent each under a distinct, peculiar way we might do a "Fregean test", so to call it: given two concepts A and B, if the judgment 'A is B' is cognitively significant – *i.e.*, not trivially true – for a subject, then A and B have a different way of characterizing their referent, a different mode of presentation of the latter, as it were. And now it seems clear that this is the case for the judgment 'the experience typically caused in most people within my community by paradigmatic red objects is the experience typically caused in me by paradigmatic red objects': one, for instance, might not know to be affected by daltonism. Hence, even if the judgment is probably true for most people, it is not *trivially* true. Same for the judgment 'the experience typically caused in most people within my community by paradigmatic red objects is the experience with such–and–such phenomenal character'. This is shown very vividly, I think, by Nida–Rümelin's (1995) famous updated version of Jackson's (1982) knowledge argument. This updated version involves a lady called Marianna. Like Jackson's Mary, Marianna has spent all her life in a black–and–white prison. At some point, she is presented with a number of colored surfaces. Yet she is not

still told the names of the colors she is seeing. Moreover, although she knows that certain objects, say sunflowers, ripe tomatoes, *et cetera* – are called yellow, red *etc.* by people in her community, she is still prevented from seeing these objects (Nida–Rümelin, *ivi:* 221). When she's finally released and gets to see how ordinary objects are colored, Marianna will presumably come to believe something along the lines of 'Oh! So THIS is what people call red!', where <THIS> is a concept which she already possessed in virtue of having being presented with nameless colored surfaces, and refers to the way it is like for her to see red, though she didn't know it referred to the experience typically caused in her and presumably in most people within her community by paradigmatic red objects (Nida–Rümelin, *Ibid.*). Before being eventually released, Marianna might even have coined a new term for the experience referred to by <THIS>, say <SENSO> (Papineau, 2011) . Marianna's new belief is far from being cognitively insignificant or trivially true for her. Likewise, at least according to many philosophers, Jackson's original version of the knowledge argument shows at least that the judgment of Mary 'whenever someone's brain cortex is in such–and–such pattern of activity, that subject is having an experience with such–and–such phenomenal character' is cognitively significant for her.

But even if it turned out, as per so–called type–A physicalism (see the introduction, 1.8.2 and 3.1 among other places) that the judgment of Mary 'whenever someone's brain cortex is in such–and–such pattern of activity, that subject is having an experience with such–and–such phenomenal character' is not cognitively significant for her, meaning that it is trivially true for her – which is something I'm skeptical about – this would not undermine the claim that there is a way of thinking about our own experiences – instantiations of phenomenal properties – in terms of the way it is like for someone to have them: As long as the latter is accepted, wether such a way of thinking about phenomenal properties is cognitively independent from physical concepts or not, is something that does not count at this stage of the work.

So, we've got that certain concepts refer to phenomenal properties and may be used to attribute them to subjects in all sorts of circumstances. Moreover, these

concepts characterize phenomenal properties as such that there is something it is like for a subject to instantiate them. Plus, this way of characterizing phenomenal properties is (arguably) distinct from the way in which other kinds of concepts would characterize them, which would also yeld true or false beliefs about them. This brings us to a possible definition of phenomenal concepts:

> *Definition of phenomenal concepts*
>
> A concept C is a phenomenal concept iff:
>
> 1. It refers to a phenomenal property P and may be used to attribute the state of being in the state in which the instantiation of P consists in (*i.e.*, having an experience E/being in E/ undergoing E/ ...) to other subjects.
> 2. It characterizes P as a property such that it is like W for a subject to live through time t iff P is instantiated at time t, where W is the specific way it is like for a subject to undergo the experience in which the instantiation of P consists in.

This is the definition of phenomenal concepts I will assume from now on.

1.8 The argument from revelation: a closer look

Eventually, I am in a position to offer a more detailed exposition of the argument from revelation against physicalism, which will be the main topic of this dissertation. My own preferred version of the argument is a slightly modified version[42] of the

[42]In the version of the argument proposed here, the thesis of revelation is spelled out in terms of Goff's (2011; 2015; 2017) *phenomenal transparency*. While Goff talks about transparent/translucent concepts, where a concepts is transparent iff it is *a priori* for a subject, in virtue and just in virtue of possessing that concept, what it is for the entity the concept refers to to be part of reality, Nida–Rümelin (2007) does not put the matter in terms of phenomenal transparency/phenomenal translucency. Rather, she (*Ibid.*) argues that phenomenal concepts allow subjects to *grasp* the nature of the experiential properties they refer to, where to grasp a property via a concept is to understand what that property essentially consists in, and to do so without any prior background knowledge besides the one that is provided by the concept itself. This seems to me to be mostly a terminological issue. At any rate, Nida–Rümelin (MS.) argues that, under the metaphysical presupposition that experiences are instantiations of experiential properties – namely, instantiations of properties by subjects at given times – Goff's phenomenal transparency is sufficiently close to

argument Nida–Rümelin develops in her (2007). This slightly modified version looks like this:

[1] *Phenomenal transparency/full–revelation*: phenomenal concepts are transparent in Philip Goff's (e.g., 2011; 2015; 2017) sense.

[2] *Cognitive independence of physical and phenomenal concepts*: given an arbitrary phenomenal concept C_{phen} there is no physical concept C_{phis} such that a subject mastering both C_{phen} and C_{phis} and accepting an arbitrary amount of background physical knowledge as true would be on that basis alone in a position to rationally judge that C_{phen} and C_{phis} are necessarily coextensive.

[3] *Cognitive inter–dependence of transparent co–referential concepts*: if two concepts are transparent with respect to the same entity – *i.e.*, necessarily co–referential and both transparent – then they cannot be cognitively independent.

[4] *Cognitive accessibility of phenomenal properties via physical concepts*: in order for physicalism to be true, for every experience E there must be at least one

her own phenomenal essentialism. As Nida–Rümelin (*Ibid.*) writes, the only further substantial difference between her own account and Goff's is whether we should restrict the thesis of phenomenal transparency/phenomenal essentialism to a subclass of phenomenal concepts, namely to what she calls maximally specific pure phenomenal concepts, or not. The reason why I chose to stick with Goff's formulation of the thesis rather than with Nida–Rümelin's is that the former allows me to make more easily and more directly a comparison with another thesis, namely *phenomenal translucency*, or the idea that phenomenal concepts reveal only part of the nature of their referents, *i.e.*, phenomenal properties, but not the entirety of it. Another option would have been to deploy Trogdon's account (2017). According to Trogdon (*ivi:* 3), a given concept C provides a partial essential characterization of their referent just in case «there is some Q such that P has Q essentially and C characterizes P in terms of Q» (*Ibid.*). By contrast C is said to provide a full essential characterization of its referent iff (*Ibid.*) «for any property Q, if P has Q essentially then C characterizes P in terms of Q. ». The notion of a characterization is just aimed at capturing the idea that, as per a post–Fregean/two–factors approach, concepts always do characterize their referent as being in a certain way or present it under a peculiar aspect. As far as I can see, Trogdon's distinction between concepts which provide a full essential characterization of their referents and concepts which provide a partial essential characterization of their referents is perfectly equivalent with Goff's distinction transparent concepts and translucent ones. Unlike Trogdon's taxonomy, Goff's one allows me to talk about mildly opaque and radically opaque concepts as well. In the second chapter I will make the case that, provided certain assumptions regarding Chalmers' two–dimensional semantics are in place, the claim that phenomenal concepts have identical primary and secondary intension (see 2.2.3 for a definition of primary and secondary intensions) is logically equivalent with Goff's phenomenal transparency. At any rate, the thesis of revelation has been proposed and/or discussed in a plethora of slightly differently nuanced formulations. See, among many others, McLaughlin (2003: 478); Stoljar (2006; 2009) Stalnaker (2008: 99) Liu, (2020; 2021; forth.); Horgan and Tienson, (2001); Coleman (2019); Majeed (2017); Damnjanovic (2012); Chalmers (2016: 2018).

transparent physical concept of E.

[5] Either physicalism is false or phenomenal transparency is false [in virtue of 2, 3 and 4].

Conclusion: physicalism is false [in virtue of 1 and 5].

Let me now walk the reader through each premise of the argument in due order.

1.8.1 Premise 1

Goff (2011; 2015; 2017) distinguishes between four kinds of concepts a subject might come to have. *Transparent* concepts reveal the entire nature of their referents, namely all their essential properties. *Translucent* concepts reveal part of the nature of their referents, *i.e.*, some of their essential properties, but not all of them. *Mildly opaque* concepts do not reveal any essential property of their referents, but reveal some property which uniquely identify their referents in the actual world. *Radically opaque* concepts do not reveal any property of their referents. Opaque concepts, that is, merely denote their referents, but say little or nothing about what it is for their referents to be part of reality.

Given a concept C, C is said to reveal a property P of its referent(s) iff it is *a priori*, for a subject possessing C and just in virtue of possessing it, that C's referent(s) has/have P.

In line with a very standard characterization, here I will take it that a given piece of knowledge of a subject is *a priori* iff experience plays only an enabling role, but not a justificatory role regarding that piece of knowledge: it is or may be necessary for forming the concepts deployed in the proposition one knows, but it plays no role in justifying it. We can also say that a given proposition D is *a priori* true or false iff a subject is in a position to see that it is true or false just in virtue of possessing the relevant concepts or understanding the meaning of the notions deployed in it, regardless of whether experience has been necessary to coming to master the relevant concepts or not.

1.8.2 Premise 2

Nida–Rümelin (2007: 328) defines cognitive independence as follows:

> The concepts C1 and C2 are cognitively independent relative to background knowledge K iff a rational person who accepts K as true in the actual world and understands C1 and C2 is not on that basis in an epistemic position to rationally judge that C1 and C2 are necessarily coextensive

The second premise of the argument assumes something more specific, namely that taken an arbitrary amount of physical knowledge as one's background knowledge K, there is no phenomenal concept C_{phen} such that there is at least one physical concept C_{phis} that is not cognitive independent from it relative to K.

Notice that this principle is fairly weak. Sponsors of so called Type–B physicalism (see Chalmers, 2003_b), among others, believe that physical concepts and physical ones are *inferentially isolated*, in that one cannot infer a phenomenal characterization of experiences from a physical one, however extensive[43] (see Sundström, 2011). This claim constitutes the core of what has been called the *phenomenal concepts strategy* to defend physicalism (see Stoljar, 2005; Carruthers and Veillet, 2007; Sundström, 2011).

Now, while inferential isolation implies cognitive independence, it is not straightforward that the latter implies the former. That inferential isolation implies cognitive independence is clear: if one cannot *a priori* infer that two concepts are *actually* – in the actual world – coextensive, then clearly she is not in an epistemic position to rationally judge that they are *necessarily* – in all possible worlds – coextensive. As for the reverse relation, there might be cases in which one could be in a position to

[43]Truth be told, it is not always clear whether on given type–B physicalist approaches inferential isolation is interpreted in a stronger sense – namely as equivalent to Nida–Rümelin's cognitive independence, meaning that given an arbitrary phenomenal–concept/physical–concept pair a subject could not *a priori* infer that the two subjects are necessarily (*i.e.*, in all possible worlds) coextensive – or in a weaker sense whereby an (ideal rational) subject could not rationally judge that the two concepts are coextensive in the actual world. But even if inferential isolation was interpreted in a weaker sense, it would still imply cognitive independence – whereas the opposite is not straightforward. See also fn. 44 of this section.

infer that two concepts are coextensive in the actual world but not be in a position to rationally judge, given what she knows about the actual world, that those two concepts are necessarily coextensive, which suggests that, at least in some cases, inferential isolation might false even though cognitive independence is true[44]. In light of this, people who accept that phenomenal concepts and physical ones are

[44]This point has been made vivid, I think, in Chalmers (1996: 32–5). There, Chalmers draws a distinction between *logical* (or conceptual) supervenience and *natural* (or nomic/nomological, or empirical) supervenience. Some given B–properties are said to be logically supervenient on some A–properties iff there is no logically possible world in which the A–properties are instantiated but the B–properties are not. Thus, we may also say that the B–properties logically supervene on the A–properties iff it is inconceivable that the A–properties may be instantiated without the B–properties being instantiated (I will elaborate on what is, for Chalmers, the sort of conceivability here at play in the second chapter) or, equivalently, iff the A–facts – instantiations of A–properties – logically entails the B–facts. By contrast, some given B–properties are said to be only naturally supervenient on some A–properties iff no two given naturally possible situations may share all their A–properties without sharing their B–properties. A naturally possible situation is a situation which may occurr in nature, that is, a possible situation which does not violate the natural laws of our world. So, on natural supervenience, although it is conceivable that A–properties may be instantiated without B–properties being instantiated, it is still true that *in nature* – that is, in our world – the situations in which the A–properties are instantiated have the same distribution of those in which the B–properties are instantiated (see Chalmers, *Ibid.*). Now, as far as I can see, the logical supervenience of B–properties on A–properties is logically equivalent with the cognitive inter–dependence – *i.e.*, the denial of the cognitive independence – of A–concepts and B–concepts, meaning that the two are either both true or both false (I will elaborate extensively on this point throughout the essay). First, suppose for the sake of argument that A–concepts and B–concepts were cognitively inter–dependent – *i.e.*, not cognitively independent – relative to a certain amount of background knowledge K – *i.e.*, a subject mastering both of A–concepts and B–concepts and accepting K as true would on that basis alone be in a position to rationally judge that every A–concept is necessarily coextensive with a least one B–concept. Such a subject could not conceive a world in which all the A–properties – *i.e.*, all the referents of the A–concepts – are instantiated without the B–properties being instantiated. So, cognitive inter–dependence implies logical supervenience. Likewise, and rather straightforwardly, if it is inconceivable for a subject that the A–properties may be instantiated without the B–properties being instantiated, then every A–concept must be cognitively inter–dependent with at least one B–concept, meaning that for every A–concept there must be at least a B–concept such that a subject mastering both of them and accepting the relevant amount of background knowledge K as true would on the basis alone in a position to rationally judge that the two are necessarily coextensive, for otherwhise the would be at least one B–property – call it B – such that one could conceive all the A–properties being instantiated without B being instantited, but then logical supervenience would be violated. So, cognitive inter–dependence implies logical supervenience. Now, clearly logical supervenience – and hence the cognitive inter–dependence of A–concepts and B–concepts – entails natural supervenience: if it is inconceivable that A–properties may be instantiated without B–properties being instantiated, then a subject should rationally judge that if the A–properties are instantiated in our world, the B–properties must also be instantiated in our world. However, natural supervenience does not imply logical supervenience by definition. Moreover, it seems clear that if natural supervenience is false – that is, if there may be a natural situation in which all the A–propertyes are instantiate without at least one of the B–properties being instantiated – then a (ideal rational) subject would not infer that if the A–properties are instantiated in our world, the B–properties must also be instantiated in our world: if natural supervenience is false, inferential isolation as defined above must arguably also be false, which by *modus tollens* implies that if inferential isolation is true, natural supervenience must also be true. To sum up, we have that it is not true that if a subject is in a position to infer that some given A–concepts are coextensive with some given B–concepts in our world then she must be in a position to rationally judge that they are necessarily coextesntive: inferential isolation

inferentially isolated, as phenomenal concept strategists do, should accept that they are cognitively independent as well (see Nida–Rümelin, 2007: 329–30).

Let me say something on the 'however extensive' bit in the characterization of type–B physicalism: type–B physicalists believe that phenomenal and physical concepts are *and will keep on being* inferentially isolated. This is what distinguishes type–B views from what Chalmers (2003_{b}) calls type–C ones: the former hold that, although physico–phenomenal identity statements are necessarily true, due to some peculiar features of phenomenal concepts they will always retain an appeareance – indeed, an illusion, according to a physicalist worldview – of contingency and mysteriousness (see, among many others, Papineau, 2002 and Loar, 1990/1997; 2002). That is to say, they believe that phenomenal zombies are both *prima facie* and ideally conceivable, though not metaphysically possible, and that there is a sense whereby there is something Mary learns upon her release from the black–and–white–prison, though all facts about color vision are physical facts.

By contrast, Type–C theorists hold that although there is *now* a physico–phenomenal epistemic gap, the gap will be closed at some point in the future, or at the very least it is closable in principle, *i.e.*, they think that zombies are *prima facie* but not ideally conceivable, and that although it seems to us now that Mary lacks some phenomenal information, in the limit it will become apparent that there is no information she lacks.

So called type–A physicalists, in turn, hold that there is no physico–phenomenal epistemic gap at all, hence there is nothing Mary learns upon being released from her black–and–white–prison, and phenomenal zombies are neither conceivable nor metaphysically possible.

Now, both type–A theorists and type–C ones are compelled to deny the second premise of the argument from revelation. In fact, by assumption Mary knows everything neuro–physiology could tell us about the visual experiences of colours.

does not imply cognitive independence. But if is not true that inferential isolation implies cognitive independence, then it is not true that if cognitive independence is false, inferential isolation must be false as well. So, inferential isolation might be true even though cognitive independence is false.

Combined with the claim that there is nothing new Mary learns upon being released, this implies either (1) that phenomenal concepts do not exist at all or are devoid or any content or (2) that although phenomenal concepts do exist, a phenomenal characterization of experiences may be entirely inferred from a physical one. Both (1) and (2) clearly imply the falsity of the second premise of the argument. Indeed, (2) merely states the denial of the second premise. As for (1), if phenomenal concepts do not exist, obviously they cannot be cognitively independent from physical ones.

Regarding type–B physicalists, as I have said they claim that physical and phenomenal concepts are inferentially isolated, which implies that they are cognitively independent. Hence, they are committed to the truth of the second premise of the argument.

I think any physicalist should accept the fourth premise of the argument (I shall say something more on this momentarily). In light of this there are only two options type–B physicalists might resort to in order to secure themselves from the argument. They might (3) deny the first premise, namely phenomenal transparency or (4) they might deny the third premise. That is to say, they might contend that two concepts might be necessarily co–referential and both transparent or at least translucent with respect to the entities they refer to and yet cognitively independent. This second strategy has been labeled *dual carving* (see Goff 2011, 200–201; 2017: 125–32; Díaz–León 2013; 2016: 1193; Taylor, 2013; Damjanovic 2012, 76–77; Elpidorou, 2015; Levin, 2019, and possibly Loar, 1990/1997; 2002 among others).

In 3.4 I will cast doubt on whether defending physicalism via dual carving is a profitable strategy. I will suggest that none of the purported counterexamples which have been offered thus far in the relevant literature against the principle stated in the third premise – namely, to the principle whereby two co–referential concepts cannot be both transparent or at least translucent and cognitively independent – is uncontroversial. Although this does not suffice *per se* to prove that the principle is true, it still give a strong *prima facie* reason to believe that it, or something very close to it, holds.

What is more, one of the aims of this essay is to show that even if both premise 2 and premise 3 of the argument happened to be bracketed or even confuted, both phenomenal transparency and phenomenal translucency – the claim that phenomenal concepts are translucent – would be sufficient *alone* to threaten conventional physicalism. Hence, provided the arguments I will draw are correct, physicalists are committed to deny both phenomenal transparency and phenomenal translucency, claiming that phenomenal concepts are either mildly opaque or radically opaque. I will suggest that the latter is the case regardless of whether one believes that physical and phenomenal concepts are not cognitively independent – as per type–A physicalism – or that they are and will keep cognitively independent – as per type–B physicalism. Being type–C physicalism a sort of "middle view" with respect to type–A and type–B ones, if the latter two approaches are denied, *a fortiori* it is denied as well[45]. Thus, both phenomenal transparency and phenomenal translucency would confute, if true, almost any variety of conventional physicalism that is on the market.

1.8.3 Premise 3

The third premise of the argument from revelation states that given two arbitrary concepts, they cannot be both transparent with respect to the same entity and cognitively independent. It is of crucial importance to understand something at this point: the principle of cognitive inter–dependence of transparent co–referential concepts does *not* state that for any arbitrary entity there is one and only one transparent concept of it. In other words, it does not state that there is just one way to grasp the (entire) nature of a given entity, so that if two subjects grasp the (entire) nature of that entity they must have the same concept. Rather, this principle asserts that if two concepts are co–referential and both transparent, then they cannot be cognitively independent.

[45]What is more, I find Chalmers' (2003_{b} 23–4) critiques to type–C materialism rather compelling. Basically Chalmers (*Ibid.*) argues that type–C materialism is an inherently instable position, in that I may easily collapse into either type–A physicalism or type–B physicalism as defined above or, possibly, into some non–physicalist view.

On this regard, I disagree with Goff. Goff (e.g., 2011: 198) labels the claim whereby the essence of a given entity or set of entities may be revealed in two or more conceptually distinct ways as *the thesis of dubious intelligibility* (TDI). I believe that, at least in certain cases, we *can*, indeed ,make sense of TDI. Yet TDI does not contradicts the principle of cognitive inter–dependence of transparent co–referential concepts. Indeed, as far as I can see the only way to make TDI intelligible is to combine it with that principle and, as we shall see in 3.4, no uncontroversial counterexamples to this have offered thus far, at least to my knowledge.

What is more, as said I am going to suggest that both phenomenal transparency and phenomenal translucency are sufficient to threaten physicalism regardless of the truth of falsity of both premise 2 and premise 3. Hence, even if premise 3 happened to be confuted – but I think there are strong *prima facie* reasons to believe it holds – physicalists would still be committed to the denial of both phenomenal transparency and phenomenal translucency.

1.8.4 Premise 4

Although some physicalists – type–B ones – may accept that physical concepts are cognitively independent from phenomenal ones, as far as I can see they are still committed to claim that every arbitrary phenomenal concept must be necessarily coextensive with at least a physical one. That is to say: although some physicalists may concede that subjects, even on ideal rational conditions, may be never be in a position to rationally judge that physical and phenomenal concepts are necessarily coextensive, they are still committed to say that those concepts are, indeed, necessarily coextensive. For suppose for the sake of argument that not all phenomenal concepts were necessarily – *i.e.*, in every metaphysically possible world – coextensive with at least a physical one, meaning that there was at least one phenomenal concept that was not necessarily coextensive with any physical concept. If this were the case, then there would be at least one possible world in which all the physical properties would be instantiated but at least one phenomenal property would not be instantiated – for

otherwise all phenomenal concepts would be necessarily coextensive with at least one physical concept, *contra* what we are assuming for the sake of argument. But a world in which all the physical properties are instantiated is an exact physical duplicate of our world. So, there would be at least one exact physical duplicate of our world which is not a duplicate of our world with respect to at least one experience, hence the positive *desideratum* of physicalism, whereby the conscious must be metaphysically supervenient upon the physical, would be violated.

So, we have that in order for physicalism to be true, every phenomenal concept must be necessarily coextensive with at least a physical one, regardless on whether subjects are in a position to rationally judge (in ideal conditions) that the latter is the case or not.

Suppose now for the sake of argument that physical concepts, or at least those physical concepts that, being necessarily coextensive with a phenomenal concept, refer to some phenomenal property, were either mildly opaque or radically opaque, *contra* what the fourth premise of the argument from revelation states. This would imply that such concepts would not characterize their referents, phenomenal properties, via any essential property of them, but rather either characterize their referents via some merely contingent property of them, or do not characterize their referents via any property at all, merely pointing to their referents or denoting them without characterizing them in any way. As far as I can see, such view would be blatantly at odds with physicalism, for it implies that while it is still true that phenomenally conscious mental states are physical, it is only contingent that they have whatever feature or combination of features described by physics and/or play whatever causal/functional/structural role(s) physics attaches to them. So, there might be metaphysically possible worlds in which phenomenally conscious mental states keep on being the states they are without, however, exhibiting any of the features and/or causal/structural/functional roles physics attaches to them. This would something very weird for physicalists to say. So, as far as I can see physicalists would better say that physical concepts are either transparent or translucent.

What is more, In 2.13 I will suggest that physical translucency is most likely at odds with most mainstream forms of physicalism (that I know of), for it leads to one of the following views: a form of property dualism – whereby there are two ontologically distinct essential aspects to the nature of physical properties – a form of mysterianism – whereby no combination of concepts, no matter how extensive and no matter the concepts it deploys, could ever get us to know the entire nature of physical properties themselves – or a form of pluralism – whereby there are several ontologically distinct aspects to the nature of physical properties.

Hence, I take it that physicalists are most likely committed to say that in order for physicalism to be true, physical concepts must be transparent.

At any rate, later on – 2.10, 2.13 and 2.14 – I will argue that even if physical concepts were taken to be either mildly opaque or radically opaque, both phenomenal transparency and phenomenal translucency would confute physicalism anyway.

Notice that, as Nida–Rümelin (2007: 326) argues, we need not have the relevant physical concepts right now, at this point in history. The principle stated in this premise merely rests on the idea whereby in order for a given entity to be physical, the nature of that entity must at least be graspable in principle by some physical concept, that is, there must be a physical concept C_{phis} such that *if we had* C_{phis} we would be in a position to understand what it is for that entity to be part of reality. But this does not imply that for every phenomenal property the concept or set of concepts in question must be already available: the relevant physical theories might still have to be developed.

The latter point concludes this preliminary exposition of the argument from revelation, as nothing needs to be said, I think, neither about the fifth premise nor about the conclusion.

Chapter 2

The conceivability argument is not an argument about conceivability

2.1 Introduction

The aim of this chapter is to show that Chalmers' (1996; 1999; 2002; 2003_{b}; 2004; 2006_{a}; 2009; $2010_{a,b}$) renowned *two–dimensional argument* – also known as the *conceivability argument* or the *zombie argument* – against physicalism may be reduced to the argument from revelation as presented and described in 1.8. I aim to show that the conceivability argument is *not* an argument about conceivability: rather, it's an argument about revelation. That is to say, I contend that no matter whether phenomenal zombies (from now on just zombies for brevity) are taken to be ideally primarily negatively conceivable or not, in order to avoid the threat posed by Chalmers' overall two–dimensional semantic framework the physicalist is committed to deny both phenomenal transparency and phenomenal translucency as defined in 1.8.1. More specifically I am going to make the case that, assumed Chalmers two–dimensional semantic apparatus, which I will describe presently, in order to avoid the threat posed by the argument physicalists are committed to claim that phenomenal concepts are mildly opaque, meaning that they do not reveal any essential properties of their referents, phenomenal properties, but only reveal some property/ies

which uniquely identify their referents only in the actual world. Notoriously, in its most general formulation, the conceivability argument goes as follows:

[1]It is conceivable that there be phenomenal zombies, namely entities physically *indistinguishable* from us which, however, lack phenomenal consciousness.

[2]If it is conceivable that there be zombies, it is possible that there be zombies.

[3]If it is possible that there be zombies, then physicalism is false.

Conclusion: physicalism is false.

Thus far, most of the major physicalist attempts to rebut the argument have attacked its second premise, namely the one whereby primary ideal negative conceivability implies primary or epistemic possibility (more on conceivability and possibility presently) (see Hill and McLaughlin, 1999: 446; Hill, 1997; Loar, 1990/1997; 1999; 2002; McLaughlin 2005; Webster 2006; Block and Stalnaker 1999; Yablo 1993; Stoljar, 2005; Goff 2010; Hawthorne 2002; Braddon–Mitchell 2003). From now on, I will refer to this thesis just the *conceivability–possibility thesis*, or CP for brevity. Here I will suggest that the latter is not a profitable strategy for the physicalist, provided Chalmers' overall two–dimensional semantic apparatus and certain other assumptions – most notably, either full–revelation/phenomenal–transparency or partial–revelation/phenomenal–translucency – are in place.

What is more, it is often thought that the conceivability argument threatens physicalism insofar as it promises to confute its positive *desideratum*, meaning that it promises to establish that there might be metaphysically possible worlds which are minimal exact physical duplicates of the actual world but not duplicates of the actual world with respect to consciousness. I will suggest, instead, that what Chalmers' two–dimensional framework really threatens, provided either full–revelation/phenomenal–transparency or partial–revelation/phenomenal–translucency is assumed, is the negative *desideratum* of conventional physicalism, namely the one whereby in order for physicalism to be true there cannot be fundamental experience–involving entities.

Here is how I will proceed throughout the chapter. First, I will describe the – indeed, very sophisticated – two–dimensional semantic apparatus Chalmers deploys to draw the conceivability argument. Then I will show that, given how certain key–notions are defined by Chalmers, primary ideal negative conceivability and primary possibility are arguably logically equivalent, meaning that it cannot be the case that a statement is primarily ideally negatively conceivable and not primarily possible or primarily possible and not primarily ideally negatively conceivable: either a statement is both primarily ideally negative conceivable and primarily possible, or it is neither primarily ideally conceivable nor primarily possible. Hence, the only way to rebut the argument via attacking the CP thesis is to deny that phenomenal zombies are primarily ideally negatively conceivable in the very first place. Then I will argue that the primary ideal negative conceivability of zombies is logically equivalent with the cognitive independence of physical and phenomenal concepts as defined in 1.8.2, meaning that the two are either both true or both false. Then I will make a case that even if the cognitive independence of physical and phenomenal concepts – and hence the primary ideal negative conceivability of zombies – happened to be denied, physicalists would still be committed to deny both full–revelation/phenomenal–transparency and partial–revelation/phenomenal–translucency in order to avoid the threat posed by Chalmers' two–dimensional semantics. I will conclude the chapter by arguing that even if the cognitive independence of physical and phenomenal concepts – hence the primary ideal negative conceivability of zombies – happened to be restored, both full–revelation/phenomenal–transparency and partial–revelation/phenomenal–translucency would still threaten physicalism. Hence, the conceivability argument is not really an argument about conceivability: it's an argument about revelation.

2.2 Chalmers' two–dimensional semantics: a brief exposition

A brief exposition of Chalmers' two–dimensional semantic apparatus is now in order.

2.2.1 Conceivability

Let us start with conceivability. Consider a statement such as 'water is not H_2O'. There is a sense whereby such statement is conceivable, as it contains no internal inconsistencies: it is not intrinsically contradictory to entertain the thought that water is not H_2O. Yet there is also a sense whereby, given what we know about the world, it is *not* really conceivable that water might not be H_2O. We will say that 'water is not H_2O' is *primarily* (or *epistemically*) conceivable but not *secondarily* (or *subjunctively*) conceivable (see Chalmers, 2002: 157). The distinction between primary and secondary conceivability stems from Kripke's (1980) discussion of the necessary *a posteriori* (but see also Putnam, 1975). In primary conceivability a given hypothetical possibility – e.g., that water is not H_2O – is considered as *actual*, *i.e.*, as a way the world might *actually* turn out to be. Primary conceivability rests on the idea whereby, for all we know *a priori*, there are several ways the world might turn out to be (Chalmers, 2002: 156–9; 2010_a: 143–4):

> *First definition of primary conceivability*
>
> A statement S is primarily conceivable iff S might actually be the case.

In wondering whether a given statement is primarily conceivable the true character of the world is irrelevant, as any empirical knowledge about the world is suspended. Hence, primary conceivability is always in the *a priori* domain (Chalmers, *Ibid.*). Thus, primary conceivability may be equivalently defined as follows:

> *Second, equivalent definition of primary conceivability*

A statement S is primarily conceivable iff it cannot be *a priori* ruled out by a subject.

By contrast, in secondary conceivability given hypothetical situations are considered as *counterfactual*, *i.e.*, as ways the world *might have been but is not*: to say that S is secondarily conceivable is to say that S might conceivably have been the case. When we wonder whether S is secondarily conceivable, we keep the character of the world as fixed and we ask whether S would have been the case if a certain situation had obtained:

Definition of secondary conceivability

A statement S is secondarily conceivable iff S might conceivably have been the case.

Hence, while primary conceivability is always *a priori*, secondary conceivability is often *a posteriori* (*Ibid.*). What is more, Chalmers (*Ibid.*) draws the following distinction:

Definition of prima facie conceivability

A statement S is *prima facie* conceivable iff it is conceivable by a subject on initial consideration.

Whereas the notion of *ideal* conceivability abstracts away from any contingent cognitive limitation of subjects (see, e.g., Chalmers, e.g., 2010_{a}: 143):

Definition of ideal conceivability

A statement is ideally conceivable iff it is conceivable on ideal rational reflection.

Finally, Chalmers draws a distinction between *positive* and *negative* conceivability:

Definition of positive conceivability

A statement S is said to positively conceivable iff a subject can coherently – without contradicting herself – form a conception or conceive a situation in which S is the case.

Definition of negative conceivability

A statement S is said to be negatively conceivable iff it cannot be ruled out by a subject.

Let now 'P' be a statement which conjoins all the physical truths about the universe plus a "that's all" clause. For brevity, we might understand P as stating that "there is the entire physical world and nothing else". Let 'Q' be an arbitrary phenomenal statement. Suppose we pick 'someone is conscious' as our relevant Q – but for the argument to work whatever other phenomenal statement, say 'someone is in pain', or whatever else, would be fine. The negation of Q, ¬Q, will be true in a given possible world iff in that possible world no one is phenomenally conscious: there is no conscious experience whatsoever. Put formally, the first premise of the zombie argument , namely 'it is conceivable that there be zombies', is meant to be read as 'P∧¬Q is primarily ideally negatively conceivable'. Thus, in what follows, unless otherwise specified I will mostly use just 'conceivability' for brevity to refer to *primary ideal negative* conceivability.

2.2.2 Scenarios

Let us now turn to scenarios. Basically, we might say scenarios stand to primary or epistemic possibility (to be defined presently) just as metaphysically possible worlds stand to secondary or metaphysical possibility (see Chalmers, 2004: 177). A scenario is said by Chalmers (e.g., 2010_a: 168) to correspond to a *maximal a priori coherent hypothesis*, namely, a hypothesis complete in every detail about how the actual world might turn out to be. Consider the statement 'water is XYZ'. Suppose such a statement is conceivable, meaning that it cannot be ruled out *a priori*, not even on ideal rational reflection. Suppose now we "enlarge" our original hypothesis, so to say, until we obtain – for instance, by adding a possibly infinite number of disjunctions[1] – an hypothesis complete in every detail about how the actual world

[1]Of course, it is unlikely that ordinary languages have the potential to express such a kind of hypotheses. For the purposes of two–dimensional semantics, one needs to assume an idealized language capable of expressing infinitary conjunctions and a great number of concepts (see Chalmers,

might be which does not contradict the original one (see, e.g., Tomasetta, 2012: 42). The result thus obtained will be a maximal *a priori* coherent hypothesis. To every maximal *a priori* coherent hypothesis, then, will correspond a scenario, and *vice versa*. Even though Chalmers is not completely explicit on this point, it seems clear from what he writes that that notion of correspondence should not be understood as equivalent to that of identity in this context. For instance, he (2010a: 268) speaks of scenarios being *described* by epistemically complete hypotheses. In (2010b: 549) he says that scenarios *represent* maximal *a priori* coherent hypotheses. It is also clear, though, that he treats the class of scenarios and that of maximal *a priori* coherent hypotheses at least as equinumerous. At any rate, the relation of correspondence between scenarios and maximal *a priori* coherent hypotheses brings us to the first definition of primary or epistemic possibility:

> *First definition of primary possibility*
>
> A statement S is said to primarily possible iff it is contained in at least one maximal *a priori* coherent hypothesis.

'Water is XYZ' is an example of a primarily possible statement. Given that to every maximal *a priori* coherent hypothesis corresponds a scenario, primary possibility may also be defined as follows:

> *Second, equivalent definition of primary possibility*
>
> A statement S is said to primarily possible iff it is verified by at least one scenario.

Now, here comes a crucial point. When Chalmers says that if P∧¬Q is conceivable then P∧¬Q is possible, it has *this* sense of possibility in mind. *Contra* what some authors suggest – e.g., Frankish (2007: 652) – the CP thesis is *not* characterized, in Chalmers (see, e.g., his 2010_a, 147–8), as the claim that conceivability implies *metaphysical* or *secondary* possibility (more on metaphysical possibility momentarily). Rather, Chalmers (*Ibid.*) defines CP as the claim that conceivability implies *epistemic* or *primary* possibility.

2004: 188).

At any rate, there are two ways to construe scenarios, namely, a metaphysical way and an epistemic one. Under a metaphysical approach to scenarios, we have that (a) to every centered[2] metaphysically possible world corresponds a scenario and (b) to every scenario corresponds a centered metaphysically possible world.

Now, (a) is relatively uncontroversial: in fact, arguably given an arbitrary centered metaphysically possible world W, one cannot exclude *a priori* that W is the actual world: at best, one can do so *a posteriori*. Thus, to every centered metaphysically possible world W will correspond a maximal *a priori* coherent hypothesis consisting in a description complete in every detail of W, and since to every maximal *a priori* coherent hypothesis corresponds a scenario, we have that to every centered metaphysically possible world corresponds a scenario. The question is whether the more controversial (b) is true, that is, whether there can be scenarios to which it does not correspond any centered metaphysically possible world or whether, instead, the class of scenarios is equinumerous with that of centered metaphysically possible worlds.

One that wanted to avoid substantive commitments about the relation between scenarios and metaphysically possible worlds might opt for an epistemic approach to scenarios. Here the story goes roughly as follows (see Chalmers, 2004: 188–9). Given an (idealized) language L and two sentences D and H in L, we say that H is compatible with D iff D∧H is primarily possible in the sense specified above. We then say that D is complete when it does not leave any question open, so to say, that is, when there is no H in L such that both H and ¬H are compatible with D – or: for every H in L, there is an *a priori* entailment either from D to H or from D to ¬H. Under this approach, a scenario is characterized as a class of equivalent epistemically complete hypotheses/sentences in L – where two sentences S and T are equivalent iff S implies T and *vice versa*.

[2]The center of a metaphysically possible world W is an ordered pair of an individual and a time in W (see Chalmers, e.g., 2004: 160). Adding a center allows to handle indexical hypotheses such as 'I am a Philosopher' or 'I am in Rome now' whereas the "uncentered" part correspond to the conjunction of all the hypotheses about the objective features of W. A purely objective description of w would leave leave several indexical hypotheses still open, and thus would not be *maximal* (Chalmers, *ivi:* 185).

2.2.3 Primary and secondary intensions

Now that we have a characterization of scenarios at hand, let us turn to Chalmers' distinction between primary and secondary intensions. Basically, for every concept C the primary intension of C can be understood as a function from scenarios to extensions of C in those scenarios. Under an epistemic approach to scenarios as I have described it above, for every scenario S the primary intension of a concept/term C will pick the referent of C in S considered as actual, that is, the thing to which C would refer if an arbitrary one of the epistemically complete sentences/hypotheses which belong to the equivalence class of S were to turn as an accurate description of the actual world, on the assumption that the primary intension of C in s is not empty in S.

Under a metaphysical approach to scenarios, given a concept C and a scenario S the primary intension of C will pick whatever thing C would refer to if the centered metaphysically possible world to which S corresponds were to turn as being the actual world provided, again, the primary intension of C is not empty in S. Thus, for instance, the primary intension of the concept <WATER> will pick the "watery stuff" (Chalmers, e.g. 1996) – *i.e.*, the liquid transparent stuff which fills the oceans *et cetera* – in every scenario in which it is not empty: it will pick XYZ in the twin–earth–scenario, H_2O in the actual–world–scenario, and so on. Given a scenario S and a composite expression E, the truth–value of the primary intension of E in S will be obtained by composition from that of the singular terms that constitute it: the primary intension of the statement 'water is not H_2O' will return 'true' as its truth–value in a scenario S iff in s the primary intension of <WATER> belongs to the class that constitute the primary intension of <EVERYTHING WHICH IS NOT H_2O> (see Tomasetta, 2012: 50). The primary intension of a concept is independent from empirical factors: since it determines the way in which the actual world should turn out to be in order for that concept to have a given extension, it is not itself dependent upon the true nature of the actual world (Chalmers, 1996: 52).

Also, the primary intension of a notion fixes an *explanandum*: If I were to

ask someone what water is, what I would be asking is to explain what the liquid transparent thing, which fills the oceans *et cetera* is (see Chalmers, 1996). Since, as we have seen, a primarily possible claim is defined as a claim that is verified by at least one scenario, in light of how primary intensions are characterized, we may also offer this further, equivalent definition of primary possibility:

> *Third, equivalent, definition of primary possibility*
>
> A statement S is said to primarily possible iff its primary intension returns 'true' as its value in at least one scenario.

As we have seen, to every maximal *a priori* coherent hypothesis corresponds a scenario and *vice versa*. In particular, Chalmers (see his 2010_b: 546, among several other places) mentions the following principle as one of the key–principles of two–dimensional semantics. For brevity, from now on I will refer to it as the PI–AP principle.

> *Primary–Intensions–A–Piori (PI–AP) Principle*
>
> a sentence D is *a priori* iff for any arbitrary scenario S the primary intension of D returns true as its value in S, *i.e.*, iff it is verified by any scenario in which its primary intension is not empty.

The main idea behind this principle can be spelled out as follows (see Tomasetta, 2012: 145, fn. 6). Consider a concept like <WATER>. What Chalmers wants to say, is that a statement like 'water is liquid' is *a priori* iff the primary intension of <LIQUID> has something that is shared by by the primary intension of <WATER> as its value in every scenario in which it is not empty. In the case of water, this is actually the case, given that the primary intension of <WATER> returns the "watery stuff" in every scenario in which it is not empty, where for something to be a "watery stuff" is for it to exhibit certain superficial properties: being liquid, being transparent, being thirst–quenching, filling the oceans, *et cetera*. Chalmers (see, e.g., 2010_b: fn. 16) also agrees, as I do, to the general principle whereby a given piece of knowledge is *a priori* iff experience plays only an enabling role, but not a justificatory role regarding that piece of knowledge (see 1.8.1).

In light of all this, it seems reasonable to infer from what Chalmers says (as per Tomasetta, 2012: 51–2, but very similar remarks can also be found in Goff, 2011) that the primary intension of any arbitrary concept C can be evaluated *a priori*, meaning that a competent speaker, one who possesses a concept C and has mastery in its use, will be endowed with a certain amount of *a priori* knowledge which will allow her to pick the primary intension of C in any scenario in which it is not empty: whoever has the concept <WATER>, say, and has mastery in its use will be put on that basis alone in a position to verify whether in a given scenario there is a liquid, thirst–quenching, transparent *etc.* stuff. As we shall see, this general principle will play a pivotal role in showing why the conceivability argument rests upon revelation, meaning that the physicalist is committed to deny both full–revelation/phenomenal–transparency and partial–revelation/phenomenal–translucency in order to avoid the threat posed by Chalmers' two–dimensional semantics, regardless of whether phenomenal zombies are taken to be conceivable or not.

Let us now turn to secondary intensions. Chalmers accepts the Kripkean (1980) *dictum* whereby several of our concepts function as rigid designators[3]. In light of this, The secondary intension of a concept C can be understood as a function that for every pair C/metaphysically–possible–world returns an extension of C (Nida–Rümelin, 2007: 316–19). Here the relevant metaphysically possible worlds are considered as counterfactual (See 2.2.1; see also Nida–Rümelin, *Ibid.*; Chalmers, e.g., 2004: 159). That is to say, they are considered as ways the world might have been but is not – in the case of any other metaphysically possible world than the actual world – or as the way the world really is – in the case of the actual world itself – where the knowledge that the world is or is not in a certain way may very well be – indeed is, at least in most cases – *a posteriori* – again, as per Kripke (1980). The secondary intension of <WATER>, thus, will have what water essentially is in the actual world, namely H_2O, as its value.

Moreover, secondary intensions have a stable value: since <WATER> is a rigid

[3]Every now and then, e.g., in his (2010_b: 543), Chalmers refers to secondary intensions as *Kripkean intensions*.

designator, the secondary intension of <WATER> will return 'H_2O' as its value in every metaphysically possible world in which it is not empty. Had our world had a different metaphysical structure from the one it has, the secondary intension of <WATER> would have had something else, say XYZ, as its value. But we know *a posteriori* the latter not to be the case. As is the case with primary intensions, the secondary intensions of complex statements is obtained by composition from that of their constituents. A metaphysically possible statement, thus, will be defined as follows:

> *First definition of secondary possibility*
>
> A statement S is secondarily possible iff it is satisfied by at least one metaphysically possible world

> *Second, equivalent definition of secondary possibility*
>
> A statement S is secondarily possible iff its secondary intension returns 'true' as its value in at least one metaphysically possible world.

The statement 'water is not H_2O' is not secondarily or metaphysically possible, since given what we know about the nature of the actual world and given that both <WATER> and <H_2O> are rigid designators, there is no metaphysically possible world in which water is not H_2O.

2.3 The CP principle under an epistemic and under a metaphysical approach to scenarios

Chalmers (e.g., 2010_a 145–6) mostly merely limits himself to argue that all the typical purported counterexamples to the CP thesis that have been provided[4] can

[4]It is often said that both Goldbach's conjecture and its negation are conceivable – in that neither of them can be ruled out *a priori* – hence are contained in at least one maximal *a priori* coherent hypothesis, hence are primarily possible, for to every maximal *a priori* coherent hypothesis corresponds a scenario. According to Chalmers (2010_a: 145) This purported counterexample to the CP thesis conflates *prima facie* conceivability with ideal conceivability. There is no reason to believe that both Goldbach's conjecture and its negation are *ideally* conceivable. Moreover, as I have already underlined, the CP thesis is not characterized, in Chalmers, as the thesis whereby (primary ideal

be rebutted, and simply assumes CP as construed on a metaphysical approach to scenarios in any of the four versions which may be offered of the argument[5]. Here I whish to say something on how the CP could be established and would look like under a weaker, epistemic approach to scenarios as characterized above and under a stronger, metaphysical approach, respectively.

2.3.1 The CP principle under an epistemic approach to scenarios

Under an epistemic approach to scenarios, the general principle whereby a given statement is primarily possible iff it is verified by at least one scenario – as per the second definition of primary possibility above – is maintained. However, we haven't got that to every scenario corresponds a metaphysically possible world.

Now, as Chalmers (2010_a: 169) himself says, that the CP principle as construed under this weaker approach to scenarios holds is very straightforward and close to being trivially true. In fact, under this approach to say that a given statement is primarily possible – *i.e.*, verified by at least one scenario – just is to say that it is primarily ideally negatively conceivable and *vice versa*: primary ideal negative conceivability and primary possibility are logically equivalent. To see why this is the case, take the statement 'water is not H_2O' again. As we saw, such statement is conceivable, in that it cannot be ruled out *a priori*, not even on ideal rational

negative) conceivability implies metaphysical or secondary possibility. Instead, CP is characterized as the claim that (primary ideal negative) conceivability implies epistemic or primary possibility. As I am about to argue, as far as I can see primary ideal negative conceivability and primary possibility are logically equivalent on Chalmers' two–dimensional semantic apparatus. Kripkean (1980) necessary *a posteriori* judgments are also usually invoked as alleged counterexamples to CP. This further purported counterexample to CP, however, conflates primary conceivability with secondary conceivability. It is primary ideal negative conceivability which, according to Chalmers, implies primary possibility.

[5]Tomasetta (2012: Ch. 5) provides an elaborated defense of CP. As I am about to argue, though, in light of how certain key–notions are defined under Chalmers' two–dimensional semantic apparatus, as far as I can see primary ideal negative conceivability and primary possibility are logically equivalent both under an epistemic approach to scenarios and under a metaphysical approach, meaning that it cannot be the case that an arbitrary statement is conceivable but not primarily possible or *vice versa*. Hence, provided one accepts Chalmers' two–dimensional semantics in the very first place, the primary ideal negative conceivability of an arbitrary statement *eo ipso* entails its primary possibility and *vice versa*. So, strictly speaking, a defense of CP would not actually be needed.

reflection. For all we know *a priori*, it might actually be the case that water is not H_2O. And now all we have to do is enlarging the hypothesis that water is not H_2O in a way that does not contradict it until we obtain a maximal *a priori* coherent hypothesis which contains the statement 'water is not H_2O'. Given that to every maximal *a priori* coherent hypothesis corresponds a scenario, we have that 'water is not H_2O' is verified by at least one scenario – *i.e.*, primarily possible (see Tomasetta, 2012: 60). Hence, under an epistemic approach to scenarios if a statement is conceivable, it is also primarily possible.

As for the converse relation of implication – *i.e.*, if a statement is primarily possible, then it is conceivable – a demonstration might go as follows (I will borrow this demonstration from Tomasetta, 2012: 61). Suppose (1) that a given statement K is primarily possible, hence there is at least one scenario S which verifies K. Also, suppose for the sake of argument (2) that K is not conceivable. From (1) follows (3) that there is at least one scenario in which the negation of K, ¬K, is not verified. Now, as we saw above, Chalmers claims that (4) if a given statement is *a priori* – *i.e.*, its truth can be known *a priori*, at least on ideal rational reflection – then it is verified by any scenario, or at least by any scenario in which the referents of the concepts that are deployed in the statement exist – *i.e.*, by any scenario in which the primary intension of the concepts deployed in the statement is not empty. From the conjunction of (3) and (4) follows (5) that ¬K is not *a priori*. From (2) follows (6) that one could rule out by ideal rational reflection that K is true, which implies that (7) one could know *a priori*, at least on ideal rational reflection that ¬K is true. But (7) implies (8) that ¬K is *a priori*. (8) and (5) are in contradiction. Give that combining (1) and (2) leads to a contradiction, we can conclude that the truth of (1) – *i.e.*, that an arbitrary statement is primarily possible – implies the falsity of (2) – namely that the statement is conceivable.

To sum up, we've got that under an epistemic approach to scenarios if a given statement is conceivable then it is primarily possible, and if it is primarily possible then it is conceivable: the two notions are logically equivalent.

2.3.2 The CP principle under a metaphysical approach to scenarios

As we saw, a metaphysical approach to scenarios assumes (a) that to every metaphysically possible world corresponds a scenario and (more controversially) (b) that to every scenario corresponds a metaphysically possible world. Thus, under a metaphysical approach to scenarios the following possible further definition of primary principle applies:

> *Primary possibility under a metaphysical approach to scenarios*
>
> Under a metaphysical approach to scenarios, a statement S is said to primarily possible iff it is verified by at least one metaphysically possible world.

A metaphysical approach to scenarios is clearly the one Chalmers endorses. However, he doesn't say much on how the stronger version of the CP thesis that is implied under such an approach may be established. Here is how this could be done (see Tomasetta, 2012: 92–4).

Take an arbitrary statement D and suppose it is conceivable, hence it could not be ruled out *a priori*, not even on ideal rational reflection. As usual, suppose now we "enlarge" D until we obtain a maximal *a priopri* coherent hypothesis which includes it. Since to every maximal *a priori* coherent hypothesis corresponds a scenario, D will be verified by at least one scenario. Since under this approach it is assumed that to every scenario corresponds a metaphysically possible world, D will be verified by at least one metaphysically possible world – namely, the one which corresponds to the scenario which verifies it. In virtue of the fact that under a metaphysical approach to scenarios it is true that a given statement is primarily possible iff it is verified by at least one metaphysically possible world, it has therefore been established that under such an approach if a statement is conceivable, it is primarily possible.

What about the converse relation? Does it hold, under a metaphysical approach to scenarios, that if a given statement is primarily possible, hence verified by at

least one metaphysically possible world, then it is conceivable? Recall that both under an epistemic approach to scenario and under a metaphysical one – and indeed rather straightforwardly (see 2.2.2) – it is true that to every metaphysically possible world corresponds a maximal *a priori* coherent hypothesis consisting in a description complete in every detail of W. Take now that an arbitrary statement T that is primarily possible, *i.e.*, under this approach, verified by at least one metaphysically possible world. In virtue of the principle whereby to every metaphysically possible world corresponds a maximal *a priori* coherent hypothesis, we have that T will be contained in at least one maximal *a priori* coherent hypothesis, namely, the one which corresponds to the metaphysically possible world which verifies it. But it is clear that if a statement is contained in a maximal *a priori* coherent hypothesis, then it must be conceivable, for by definition maximal *a priori* coherent hypotheses cannot contain self–contradictory claims, so if a statement is contained in one of such hypotheses it cannot be self–contradictory. But a statement that is not self–contradictory is conceivable. Hence, we have that also under a metaphysical approach to scenarios if a statement is is primarily possible, it is conceivable, and *vice versa*: again, the two notions are logically equivalent.

2.4 Primary ideal negative conceivability and the cognitive independence of physical and phenomenal concepts

To take stock, given the way in which maximal *a priori* coherent hypotheses, scenarios, primary ideal negative conceivability and primary possibility are defined by Chalmers, both under an epistemic approach to scenarios and under a metaphysical one there is a relation of necessary co–implication, hence of logical equivalence, between primary ideal negative conceivability and primary possibility. If it is true that a statement is conceivable, then it must be true that it is primarily possible and *vice versa*. As I have said, thus far most of the major physicalist attempts to rebut the zombie argument

have focused on the CP principle, attempting to claim that there might be statements that are conceivable but not primarily possible. But if I'm right that primarily ideal negative conceivability and primary possibility are logically equivalent under both the possible approaches to scenarios that are available, the only way for a physicalist to attack CP is to deny that P∧¬Q is (primarily ideally negatively) conceivable in the very first place where, recall, P∧¬Q consists in the statement that 'there is the entire physical world and nothing else, and no one is conscious'. Recall, also, that even though, in order to stick to the latter of Chalmers, I've chosen 'someone is conscious' as my relevant phenomenal statement, any arbitrary phenomenal statement – say, 'someone is in pain' – would be sufficient for the argument to work.

What I now whish to do, is to show that in order to deny the conceivability of P∧¬Q the physicalist is committed to deny the cognitive independence of physical and phenomenal concepts as stated in the second premise of the argument from revelation – 1.8.2 – in that that the denial of the former implies the denial of the latter and *vice versa*.

The second premise of the argument from revelation states that given an arbitrary phenomenal concept C_{phen} there is no physical concept C_{phis} such that a subject having both C_{phen} and C_{phis} and arbitrary background physical knowledge would be on that basis alone in a position to rationally judge that C_{phen} and C_{phis} are necessarily coextensive.

Consider the phenomenal concept <PHENOMENAL PAIN> and suppose we take 'someone is in pain' as our 'Q' in P∧¬Q. Suppose now for the sake of argument that there is at least one arbitrary physical concept, say <C-fibers FIRING>, such that it is not cognitively independent from <PHENOMENAL PAIN>: that is to say, a subject mastering both <PHENOMENAL PAIN> and <C-fibers FIRING> and accepting the relevant amount of background physical knowledge as true would on that basis alone be in a position to rationally judge that the two concepts are necessarily coextensive. As far as I can see, such a subject would be able to *a priori* infer Q – namely, that someone is in pain – from any physical statement deploying

<C-fibers FIRING>. That is to say, she would be able to infer that necessarily, whenever there is C-fibers firing in a given possible world there is pain in that world. Now recall that the statement 'P' in P∧¬Q consists in a *complete* physical description of our world plus a "that's all" clause. That is to say, P states that there is the *entire* physical world *and nothing else*. In light of this, given that <C-fibers FIRING> is a physical concept, namely a concept referring to a physical property by assumption, P must deploy either <C-fibers FIRING> or at least another concept which, in turn, is cognitive inter–dependent – *i.e.*, not cognitively independent – from <C-fibers FIRING> itself. Put in other words, given that C-fibers–firing is a physical property by stipulation, if it is true there is the entire physical world and nothing else, then it must be true that there is C-fibers firing as well. What is more, given that by stipulation P consists in a complete description of the physical world conjoining all the physical truths about the universe, whatever the amount of background physical knowledge that is required in order for a subject to be put in a position to rationally judge that <PHENOMENAL PAIN> and <C-fibers FIRING> are necessarily coextensive, *a fortiori* such amount of knowledge will be included in P.

Given all this, we have that a subject who accepted P as true and mastered <PHENOMENAL PAIN> and <C-fibers FIRING> would be able to infer 'Q', namely that someone is in pain, *a priori*. That is, she would be able to infer that necessarily, whenever there's the entire physical world, and hence also C-fibers firing, and nothing else, there must be pain. From this follows that P∧¬Q would not be conceivable by such a subject. The very same line of reasoning may be reiterated for any arbitrary phenomenal concept – say, <PHENOMENAL RED>, <MELANCHOLY> or whatever.

So, taken any arbitrary phenomenal concept C_{phen}, denying that C_{phen} is cognitively independent from any arbitrary physical concept, and so accepting that there is at least one physical concept C_{phis} that is cognitively inter–dependent from C_{phen}, implies the denial that P∧¬Q is conceivable where, for any arbitrary C_{phen},

the relevant 'Q' is taken to be, roughly, the statement that the referent of C_{phen} is instantiated. Hence, by *modus tollens* the conceivability of P∧¬Q implies the cognitive independence of physical and phenomenal concepts.

Does the converse relation of implication hold? That is to say: does the cognitive independence of physical and phenomenal concepts imply the conceivability of P∧¬Q? Well, suppose, once again, we take 'someone is in pain' as our relevant 'Q' in P∧¬Q. Moreover, assume for the sake of argument that P∧¬Q is not conceivable, so it can be ruled out *a priori*, at least on ideal rational reflection. This means that an ideal rational subject would judge that whenever there's the physical world and nothing else, *eo ipso* there must be pain: there could *not* not be pain. Given that P states that there's the physical world *and nothing else*, as far as I can see this implies that such a subject would judge that one of the physical properties described in 'P' *just is* pain, which in turn implies that our subject would rationally judge that one of the physical concepts deployed in P is necessarily coextensive with <PHENOMENAL PAIN>.

Hence, we have that denying that P∧¬Q is conceivable implies the denial of the cognitive independence of physical and phenomenal concepts, thus by *modus tollens* the cognitive independence of physical and phenomenal concepts implies the conceivability of P∧¬Q.

Provided the arguments I've drawn so far are correct, the primary ideal negative conceivability of P∧¬Q and the cognitive independence of physical and phenomenal concepts are logically equivalent, as the former implies the latter and *vice versa*. Hence, denying that P∧¬Q is conceivable implies denying that physical and phenomenal concepts are cognitively independent[6].

[6]By the way, as far as I can see the line of reasoning just exposed holds also the other way around. Suppose Q∧¬P was taken to be conceivable. This implies that it would not be necessary, for an ideal rational subject, that whenever there is consciousness there must be the entire physical world and nothing else. But this, in turn, entails that, taken an arbitrary phenomenal concept C_{phen}, C_{phen} would be cognitively independent from any arbitrary one of the physical concepts deployed in P. For if there was a physical concept C_{phis} such that, by mastering C_{phis} and C_{phen} and accepting P as true our subject would be in a position to infer that C_{phis} and C_{phen} were necessarily coextensive, she would not be in a position to conceive the referent of C_{phen} to be instantiated without *eo ipso* considering the referent of C_{phis} to be instantiated as well. But then Q∧¬P would not be conceivable anymore for her. So, we have that, plausibly, the conceivability of P∧¬Q is logically

2.5 The anti–zombie argument

Clearly, given two arbitrary concepts C_1 and C_2 they cannot be both cognitively independent and cognitively inter–dependent: if a subject is in a position to rationally judge that two concepts are necessarily coextensive, she cannot not be in a position to rationally judge that they are necessarily coextensive. Put in order words, provided the cognitive independence of physical and phenomenal concepts is logically equivalent with the (ideal primary negative) conceivability of $P\wedge\neg Q$, $P\wedge\neg Q$ cannot, obviously, be both conceivable and inconceivable. Now suppose $P\wedge\neg Q$ is not conceivable, hence physical and phenomenal concepts are not cognitive independent. This implies that taken an arbitrary phenomenal concept C_{phen}, there is at least one physical concept C_{phis} such that a subject mastering C_{phen} and C_{phis} and accepting the 'P' in $P\wedge\neg Q$ as true, which by definition contains a complete physical description of the world, would on that basis alone be in a position to rationally judge that C_{phen} and C_{phis} are necessarily coextensive. That is to say, such a subject would be in a position to rationally judge that whenever in a given possible world there is, say, phenomenal pain, *eo ipso* there is C-fibers firing. Hence, taken any arbitrary 'Q' as one's relevant phenomenal statement, our subject would be in a position to *a priori* infer 'Q' from

equivalent with the cognitive independence of physical and phenomenal concepts which, in turn, is logically equivalent with the conceivability of $Q\wedge\neg P$. In other words, by denying that zombies are conceivable the conceivability of *ghosts* – creatures phenomenally indistinguishable from us which, however, lack any physical basis – is denied as well. This is so because because 'P' states that there is the entire physical world *and nothing else*. But from the fact that it is inconceivable that there may be the entire physical world *and nothing else* without *eo ipso* there being consciousness follows that every phenomenal *just is* a physical property. But if every phenomenal property *just is* a physical property, then trivially it may not be instantiated without the relevant physical property being instantiated as well. More in general, as Goff (2010) argues cogently, physicalists have to fear ghosts just as much as they fear zombies. In fact (see Goff, 2010: 8), physicalists are committed to say that conscious phenomena are essentially higher–order, *i.e.*, realized by some more fundamental physical goings on. But the possibility of ghosts implies that conscious phenomena are not essentially higher–order, for they occur even in the absence of of the more fundamental physical goings–on. Hence, physicalists must deny that both zombies and ghosts are possible and possibly also that they are conceivable (at least on type–A physicalism). Put in other words, physicalists are committed to say that only fundamental physical properties are *sparse*, where Lewis (1986: 60) defines sparse property as follows: «sharing of them makes for qualitative similarity, they carve nature at the joints, they are intrinsic, they are highly specific, the sets of their instances are ipso facto not entirely miscellaneous, there are only just enough of them to characterise things completely and without redundancy». The possibility of ghosts, by contrast, may make phenomenal properties count as sparse properties of our world (see Goff, 2010: 13). I cannot afford here to get into the detail of what the relation between sparse properties and fundamental properties is.

'P'. But it is clear that if a subject is in a position to infer 'Q' from 'P' *a priori*, then obviously P∧Q must be conceivable. So, if P∧¬Q is not conceivable, P∧Q must be conceivable. Hence, by *modus tollens* if P∧Q is not conceivable, P∧¬Q must be conceivable: P∧¬Q and P∧Q cannot be both conceivable or both inconceivable.

Chalmers' zombie argument seeks to establish the falsity of materialism completely *a priori*, that is to say, without any appeal to empirical results. Yet some philosophers have questioned this approach. In particular, the idea that the falsity of physicalism can be established by purely conceptual means has been challenged in two papers by Keith Frankish (2007) and Richard Brown (2010). Their strategy is to drawn an argument which mirrors the zombie argument but, crucially, is premised on the conceivability of *anti–zombies* (Frankish, 2007) or *shombies* (Brown, 2010). Anti–zombies are creatures which are physically indistinguishable from us, do have phenomenal consciousness, and have no additional non–physical property whatsoever. In its most general formulation, the anti–zombie argument goes as follows:

[1]Anti–zombies are conceivable.

[2]If anti–zombies are conceivable, then they are possible

[3]If anti–zombies are possible, dualism is false

Conclusion: dualism is false.

We may refer, As Frankish (*Ibid.*) does, to physicalists who want to appeal to the anti–zombie case as "anti–zombists" and to anti–physicalists who want to appeal to the zombie argument as "zombists". Notice that while for the zombie argument to work any arbitrary phenomenal statement Q in P∧¬Q is fine, anti–zombists are committed to pick 'everyone is conscious' as their 'Q'. For in order for physicalism to be true any phenomenal property whatsoever must be supervenient upon the physical, which entails that it is sufficient for the anti–physicalist to prove that there is just one phenomenal property which is not supervenient upon the physical. At any rate: suppose we do pick 'everyone is conscious' as our 'Q'. The moral of the story should be easy to be grasped. At least *prima facie*, there is no intrinsic

contradiction involved in thinking that there is the entire physical world and nothing else and everyone is conscious. Yet if the argument just provided is correct, if P∧¬Q is ideally conceivable, P∧Q cannot be, and *vice versa*. Either P∧¬Q or P∧Q must be ideally inconceivable and hence, given all the other premises of (the four version of: see below) the anti–zombie argument or of the zombie one, both epistemically and metaphysically impossible. In virtue of the resulting stalemate, Frankish and Brown conclude that matters of conceivability do not play any positive role in the debate on mind–body metaphysics which, therefore, can only be settled *a posteriori*. Notice, moreover, that as Frankish (2007: 656–7) argues, while, assumed a dispositionalist conception of theoretical physics – on which I've said something in 1.6 and I will elaborate more extensively in 3.5 – the zombie argument leaves open the possibility of *a posteriori* necessities linking the dispositional/causal/functional profile of the physical world with consciousness, the anti–zombie argument does not suffer from this "weakness", in that saying that the functional/structural/causal profile of the physical world suffices to metaphysically necessitate consciousness is already enough to prove the truth of physicalism. Thus, apparently we really face a stalemate, and one which might even favor anti–zombists. Later on I will cast doubt on whether this is really the case.

Yet I do agree with with Frankish and Brown that conceivability is not what really matters here: what really matters here is revelation. Here is how I will proceed to make my case. First I will present the four distinct versions which may be given of the zombie argument and elaborate on how, according to Chalmers, each premise of each version of the argument may be justified. The first version of the argument assumes that both phenomenal concepts and physical ones have identical primary intension and secondary intension; the second version of the argument assumes that while phenomenal concepts have identical primary and secondary intension, the primary and secondary intension of physical concepts differ; the third version of the argument assumes that while the primary and secondary intension of phenomenal concepts differ, physical concepts have identical primary and secondary intension; finally, the

fourth version of the argument assumes that the primary and secondary intension of both physical and phenomenal concepts differ. I will then proceed to show that the claim that phenomenal concepts have identical primary and secondary intension is arguably logically equivalent with full–revelation/phenomenal–transparency provided some minimal and rather plausible assumption are in place. I will then argue that the combination of full–revelation/phenomenal transparency with the denial of the cognitive independence of physical and phenomenal concepts is incompatible with physicalism. Four specular version of the anti–zombie argument may be offered, which my marked with a *. Provided my arguments are correct, the first* version of the anti–zombie argument and the second* version must be ruled out, for they assume that P∧Q is conceivable, hence P∧¬Q is not conceivable, hence phenomenal and physical concepts are not cognitively independent, and that phenomenal concepts have identical primary and secondary intension – which, I argue, is logically equivalent with full–revelation/phenomenal–transparency, provided certain very plausible assumptions are in place. This leaves us with the third* version of the anti–zombie argument and the fourth* one. Following Goff (e.g., 2011) and Tomasetta (2012) I will then suggest that Chalmers' two–dimensional semantics is incompatible with the existence of radically opaque concepts. We are therefore left with two options: either in the third* version of the anti–zombie argument and in the fourth* one phenomenal concepts are assumed to be translucent or they are assumed to be mildly opaque. I will argue that the combination of partial–revelation/phenomenal–translucency with the denial of the cognitive independence of physical and phenomenal concepts is incompatible with physicalism. Thus, the only option left available to physicalists is to claim that phenomenal concepts are mildly opaque, fixing some property/ies which is/are not essential to their referents but uniquely identify them in the actual world. I will conclude the chapter by arguing that even if the cognitive independence of physical and phenomenal concepts was restored – that is, even if zombies were taken to be conceivable, and hence physical and phenomenal concepts were taken to be cognitively independent – physicalists

would still be committed to deny both full–revelation/phenomenal–transparency and partial–revelation/phenomenal–translucency. Hence, the conceivability argument is an argument about revelation, rather than being an argument about conceivability.

2.6 The four versions of the zombie argument

As a matter of fact, four distinct versions of the conceivability argument may be offered, which vary in their assumptions as per the following schema[7] The first version of the argument assumes that both phenomenal concepts and physical ones have identical primary intension and secondary intension; The second verssion assumes that while phenomenal concepts have identical primary and secondary intension, the primary and secondary intension of physical concepts differ; The third version assumes that while the primary and secondary intension of phenomenal concepts differ, physical concepts have identical primary and secondary intension; finally, The fourth version assumes that the primary and secondary intension of both physical and phenomenal concepts differ. Moreover, all the versions of the argument have the CP thesis as construed under a metaphysical approach to scenarios among their assumptions. That is to say, they assume that if a statement is conceivable, then it is primarily possible and hence, since a metaphysical approach to scenarios is assumed, verified by at least one metaphysically possible world. Finally, each version of the argument assumes that both physical and phenomenal concepts function as Kripkean rigid designators. I will now proceed to present each version of the argument and how Chalmers claims each premise (thus the conclusion) of each version may be justified.

[7]The third and fourth versions of the argument are only mentioned *en passant* by Chalmers (e.g., 2010_a) but are not discussed in depth by him. I've discussing them just for completeness. An in depth discussion of each of the four versions can be found in Tomasetta (2012).

2.6.1 Version I

The first version of the argument assumes that both physical and phenomenal concepts have identical primary and secondary intension. This version looks like this:

[1]P∧¬Q is conceivable.

[2]If P∧¬Q is conceivable, it is primarily possible.

[3]If P∧¬Q is primarily possible, then it is secondarily possible.

[4]If P∧¬Q is secondarily possible, then physicalism is false.

Conclusion: physicalism is false.

As far as I am aware, Chalmers does not goes much beyond saying that the primary ideal negative conceivability of P∧¬Q is a very plausible idea. The second premise is justified in virtue of the CP principle, which states that if a statement is conceivable, then it is primarily possible. The third premise is justified in virtue of the assumption whereby both physical and phenomenal concepts have identical primary and secondary intension, plus the general principle whereby both the primary and the secondary intension of composite statements are obtained by composition from the primary/secondary intension of their constituents. Let me elaborate. For a statement to be primarily possible is for its primary intension to be true in at least one scenario. But given that in any version of the argument it is assumed, as per a metaphysical approach to scenarios, that to every scenario corresponds a metaphysically possible world, it follows that the primary intension of the statement in question will be true in at least one possible world[8]. Given that it is assumed that both physical and phenomenal concepts have identical primary and secondary

[8]Given how the third premise of this and the other versions of the argument is justified, some might feel that the argument has to assume a metaphysical approach to scenario in order to work against physicalism. As far as I can see, though, nothing of what I will say below to make the case that the conceivability argument is actually an argument about revelation/phenomenal–transparency hinges upon this. It seems to me that even if one did not assume, as per a metaphysical approach to scenarios, that to every scenario corresponds a metaphysically possible world, both phenomenal transparency and phenomenal translucency would be sufficient to threaten physicalism, both under the assumption that phenomenal and physical concepts are cognitively independent and under the assumption that they are not so.

intension, if the primary intension of P∧¬Q is true in a given metaphysically possible world then its secondary intension will also have to be true in that metaphysically possible world. But if the secondary intension of a statement is true in at least one metaphysically possible world, then that statement is secondarily or metaphysically possible. Now, saying that P∧¬Q is metaphysically possible amounts to saying that there is at least one metaphysically possible world that is a minimal exact physical duplicate of the actual world but not a C–duplicate of the actual world, namely a duplicate of the actual world with respect to consciousness. This in turn amounts to the denial of the positive *desideratum* of physicalism as stated in 1.2, namely the metaphysical supervenience of phenomenal properties upon physical ones, and hence implies the falsity of conventional physicalism, as the fourth premise of the argument states. The conclusion that physicalism is false follows accordingly.

2.6.2 Version II

What distinguishes version II. from version I. is that in this version it is assumed that while phenomenal concepts have identical primary and secondary intension, the primary and secondary intension of physical concepts differ. More specifically, in line with a dispositionalist conception of theoretical physics Chalmers (e.g., 2010_{a}: 150) says that the primary intension of concepts like <ELECTRON> or <MASS> has certain specific structural/causal/functional roles as its value, whereas the secondary intension of the same concepts has whatever really realizes the relevant structural/causal/functional roles in the actual world as its value. This picture allows for the possibility of there being two (or more) metaphysically possible worlds which share the same structural/causal/dispositional/functional profile but have a different intrinsic profile or, which is the same, two (or more) metaphysically possible worlds in which the same structural/causal/dispositional/functional roles are realized but the "things" which realize those roles are different. Let us now see how this version of the argument looks:

[1]P∧¬Q is conceivable.

[2]If P∧¬Q is conceivable, then it is primarily possible.

[3]If P∧¬Q is primarily possible, then P∧¬Q is secondarily possible or Russelian monism is true.

[4]If P∧¬Q is secondarily possible then physicalism is false.

Conclusion: Physicalism is false or Russelian monism is true.

Here is how, according to Chalmers, each premise of the argument may be justified. The second premise is justified in virtue of the CP principle[9]. Here it is worth saying something on the third premise. Given that in this version physical concepts are assumed to have distinct primary and secondary intensions, from the fact that P∧¬Q is primarily possible does not immediately follows that it is secondarily or metaphysically possible, as was the case in the first version of the argument. We are therefore left at a crossroad: either P∧¬Q is both primarily possible and secondarily possible, which is incompatible with the truth of physicalism, for it implies the denial of the positive *desideratum* of it, or P∧¬Q is primarily possible but not secondarily possible. As Chalmers says (e.g., in his 2010_a: 151), assumed as dispositionalist conception of theoretical physics saying that P∧¬Q is primarily possible but not secondarily possible implies that while the structural/causal/dispositional/functional profile of the actual world alone does not necessitate the existence of phenomenal consciousness, the combination of the structural/causal/dispositional/functional profile of the actual world and of its intrinsic profile does necessitate the existence of phenomenal consciousness. That is to say: while a given metaphysically possible world may share its structural/causal/dispositional/functional profile but not its intrinsic profile with the actual world and not be a C–duplicate of the actual world itself, whenever a metaphysically possible world shares both its structural/causal/dispositional/functional profiles and its intrinsic one with the actual world it *eo ipso* is a C–duplicate of the actual world. According to Chalmers (e.g., *Ibid.*) the only plausible explanation for

[9]From now on I will skip the first premise. As I have said, Chalmers does not go very much beyond saying that it is indeed very plausible. What is more, later on – 2.10 and 2.13 – I will show that even if the ideal primary negative conceivability of P∧¬Q happened to be denied, provided certain other assumptions are in place physicalism would still be threatened.

this fact is that phenomenal properties or at the very least something very close to them are the realizers of the structural/causal/dispositional/functional roles that are expressed by physical concepts. The latter claim, though, amounts to a panpsychist version of Russelian monism. Panpsychist Russelian monists assume a dispositionalist conception of theoretical physics and posit that the intrinsic properties underlying the spatiotemporal structure and dynamics that physics describes are either in themselves necessarily conscious or somehow "proto–conscious", namely necessary to constitute consciousness (see Alter and Pereboom, 2023)[10]. The truth of Panpsichist Russelian monism is inconsistent with the negative *desideratum* of (almost) any coherent form of materialism, whereby no experience–involving entity is fundamental, at least under most standard interpretations. Therefore, it would arguably be sufficient to say that materialism is false without adding any further specification. However, there are authors (e.g., Strawson, 2008_{a}) who conceive panpsychist Russelian monism as a genuine kind of materialism, if not even as the only acceptable form of materialism. Indeed, although skeptical, Chalmers himself (e.g., 1996: 114) seems at least open to the idea of conceiving panpsychist Russelian monism as a very peculiar but still genuine form of materialism, even though he (2002a: 265) acknowledges that Russelian monism «fits the letter of materialism» but «shares the spirit of anti–materialism». That is why, along with Chalmers himself (e.g., 2010_{a}: 151) and Tomasetta (2012), among others, I formulated both the third premise and the conclusion of the argument as a disjunction stating that either physicalism is false or Russelian monism is true.

2.6.3 Version III

On this version it is assumed that while the primary and secondary intension of phenomenal concepts differ, physical concepts have identical primary and secondary intension. As in any other version, all the other assumptions – *i.e.*, that both physical

[10]What distinguishes panpsychists versions of Russelian monism from so called neutral ones, is that the latter claim that the categorical basis of the universe is formed by properties that are neutral, *i.e.*, neither physical nor phenomenal.

and phenomenal concepts are rigid designators, and the CP principle as construed under a metaphysical approach to scenarios – are in place. Here is how this version of the argument looks:

[1]P∧¬Q is conceivable.

[1]If P∧¬Q is conceivable then it is primarily possible.

[3]If P∧¬Q is primarily possible, then there is at least one metaphysically possible world which is a minimal exact physical duplicate of the actual world but not a duplicate of the actual world with respect to consciousness' mode of appearance.

[4]If there is at least one metaphysically possible world which is a minimal exact physical duplicate of the actual world but not a duplicate of the actual world with respect to consciousness' mode of appearance, materialism about consciousness' mode of appearance is false.

Conclusion: materialism about consciousness mode of appearance is false.

Now here is how each premise of the argument may be justified. As in any other version, the second the second premise is justified in virtue of the CP thesis. Of course, what is peculiar of this version of the argument are its last two premises. We have seen that while the primary intension of a given concept C has C's referent's superficial properties as its value in any scenario in which it is not empty – e.g., the being liquid, thirst quenching *etc.* of water – the secondary intension of C returns the real essential nature of the referent of C – H_2O in the case of water – as its value in every scenario/metaphysically–possible–world in which it is not empty. Now, given that on this version physical concepts are assumed to have identical primary and secondary intension, and given that a statement is primarily possible iff its primary intension returns 'true' as its value in at least one possible world, P will be primarily possible iff there is at least one scenario, and hence at least one metaphysically possible world, in which there is the real, actual physical world, *i.e.*, the real, actual realizers of certain structural/functional roles, and nothing else. By contrast, since it

is assumed that the primary and secondary intension of phenomenal concepts differ, ¬Q will be primarily possible iff there is at least one scenario, and hence at least one metaphysically possible world, in which nothing has consciousness' superficial properties or its "mode of appearance", so to say, which does not imply, however, given what this version of the argument assumes, that in such a metaphysically possible world there isn't the real, actual consciousness. In virtue of the principle whereby both the primary and the secondary intension of composite statements are obtained from those of their constituents, ¬Q will be primarily possible iff there is at least one scenario, and hence, assumed a metaphysical approach to scenarios, one metaphysically possible world, in which there is the real physical world and nothing else, and nothing has consciousness' mode of appearance. This implies that there is at least one metaphysically possible world which is an exact minimal physical duplicate of the actual world but not a duplicate of it with respect to consciousness' mode of appearance, hence consciousness' mode of appearance does not supervene on physical properties, hence physicalism about consciousness' mode of appearance is false, as the conclusion of the argument states.

2.6.4 Version IV

The fourth version of the argument assumes that the primary and secondary intension of both physical and phenomenal concepts differ. Here is how this version goes:

[1]P∧¬Q is conceivable.

[2]If P∧¬Q is conceivable, then it is primarily possible

[3]If P∧¬Q is primarily possible, then there is at least one metaphysically possible world which is a minimal exact physical duplicate of the actual world but not a duplicate of the actual world with respect to consciousness' mode of appearance or Russelian monism about consciousness' mode of appearance is true.

[4]If there is at least one metaphysically possible world which is a minimal exact physical duplicate of the actual world but not a duplicate of the actual

world with respect to consciousness' mode of appearance, materialism about consciousness' mode of appearance is false.

Conclusion: Materialism about consciousness mode of appearance is false or Russelian monism about consciousness' mode of appearance is true.

This version of the argument is defended in virtue of the very same principles and lines of reasoning I have elaborated on so far. Hence, readers who have followed me up to this point should be able to see rather straightforwardly how each premise (hence the conclusion) of this version may be justified. Since I don't want to be repetitive, I won't say anything more than I've already said. The only thing worth mentioning, is that according to Chalmers (e.g., 2010_{a}: 153) Russelian monism about consciousness' mode of appearance still counts as a form of Russelian monism, although distinct from Russelian monism about consciousness *tout court*. This is so because in the former kind of view the combination of the intrinsic profile of the actual world and of its structural/causal/dispositional/functional profile necessitates consciousness' mode of appearance, whereas the structural/causal/dispositional/functional profile of the actual world alone does not.

2.7 Tacking stock (I)

As I argued, provided in Chalmers' two–dimensional semantics primary ideal negative conceivability and primary possibility are logically equivalent, meaning that if a statement is primarily ideally negatively conceivable it has to be primarily possible and *vice versa*, the only option available to the physicalist who wanted to attack the CP principle is to deny that $P\wedge\neg Q$ is conceivable in the very first place. The conceivability of $P\wedge\neg Q$ is logically equivalent with the cognitive independence of physical and phenomenal concepts. Moreover, since if $P\wedge\neg Q$ is not conceivable $P\wedge Q$ must be conceivable and *vice versa*, the physicalist who wanted to pursue this strategy is committed to say that $P\wedge Q$ is conceivable, as per the anti–zombie argument. What I now whish to do, is to show that even if the cognitive independence of physical

and phenomenal concepts happened to be denied – that is to say, even if P∧¬Q was taken to be inconceivable, hence P∧Q was taken to be conceivable – physicalism would still be threatened both by phenomenal transparency and by phenomenal translucency.

2.8 Why there are good reasons to believe that the claim that phenomenal concepts have identical primary and secondary intension and phenomenal transparency are logically equivalent: some hints

The first step of my argument consists in offering some preliminary remarks on why the claim that phenomenal concepts have identical primary and secondary intension arguably entails full-revelation/phenomenal-transparency as characterized in 1.8.1.

As we saw in 2.2.3, it seems reasonable to infer from what Chalmers says (as per Tomasetta, 2012: 51–2, but see also Goff, e.g., 2011) that the primary intension of a concept C can be valuated *a priori*, meaning that a competent speaker, one who possesses a concept C and has mastery in its use, will be endowed with a certain amount of *a priori* knowledge which will allow her to pick the primary intension of C in any scenario in which it is not empty: whoever has the concept <WATER>, say, and has mastery in its use will be put on that basis alone in a position to verify whether in a given scenario there is a liquid, thirst–quenching, transparent stuff. We also saw that the secondary intension of a given concept C has the real, essential nature of C's referent as its value in every metaphysically possible world in which it is not empty.

Hence, we have that, on the assumption that phenomenal concepts have identical primary and secondary intension, whoever mastered a phenomenal concept C_{phen}

would be *eo ipso* endowed with a relevant piece of *a priori* knowledge regarding the real, essential nature of the referent of C_{phen}. But this is just what the thesis of phenomenal transparency states. For recall that Goff (2017: 15, emphasis mine) defines transparent concepts as those concepts such that they reveal the nature of the entity they refer to «in the sense that *it is* a priori *(for someone possessing the concept and in virtue of possessing the concept) what it is for that entity to be part of reality*».

By the way, that phenomenal properties are picked by their essential properties is something that already Kripke (e.g.,1980: 152–3) clearly endorsed, as it can be seen here:

> Pain, on the other hand, is not picked out by one of its accidental properties; rather it is picked out by the property of being pain itself, by its immediate phenomenological quality. Thus pain, unlike heat, is not only rigidly designated by 'pain' but the reference of the designator is determined by an essential property of the referent [. . .]If any phenomenon is picked out in exactly the same way that we pick out pain, then that phenomenon *is* pain.

2.9 More on why the claim that phenomenal concepts have identical primary and secondary intension and phenomenal transparency are logically equivalent: the NARD principle

The primary intension of a concept C has the superficial features of C's referents as its value. The secondary intension of C has C's referent's real essence as its value. Hence, the claim that phenomenal concepts have identical primary and secondary intensions is equivalent to the claim that the superficial features of phenomenal properties coincide with their real essence: whatever appears as a phenomenal property, *just*

is a *real* phenomenal property. If something has pain's superficial properties – say, it feels like pain, it is painful, it hurts – that thing just is and could not not be a real instance of phenomenal pain. This claim is a version of the general principle Michelle Liu (forth.) has labeled NARD (No–Appearance–Reality–Distinction).

Liu (*ivi:* 5–9) distinguishes between two versions of NARD – namely, NARD1 and NARD2. According to NARD1, «there is no distinction between what phenomenal property an experience appears to the subject to have and what phenomenal property the experience really has» (*ivi:* 5) or, equivalently (*Ibid.*), «Given an experience e and a phenomenal property Q, e has Q iff in having e, it appears to a subject S that e has Q»[11]. Notoriously, philosophers have distinguished between at least two senses of the verb 'to appear', namely an *epistemic* sense and a *phenomenal* one (see Liu, *Ibid.*; Schwitzgebel 2008: 263; Gertler 2012: 106–107). According to an epistemic interpretation of 'to appear', it appears to S that e has Q iff S *believes* that e has Q. According to a phenomenal interpretation of 'to appear', it appears to S that e has Q iff S *experiences* e as having Q. We may therefore formulate two distinct versions of NARD1, namely an epistemic version, and a phenomenal version. According to EPISTEMIC NARD1, given an experience e and a phenomenal property Q, e has Q iff in having e, it epistemically appears to S that e has Q, *i.e.*, iff S *believes* that e has Q. According to PHENOMENAL NARD1, in turn, given an experience e and a phenomenal property Q e has Q if and only if in having e, a subject S *experiences* that e has Q.

Now, as Liu (*Ibid.*) herself notes (but see also Schwitzgebel 2008; 2011) EPISTEMIC NARD1 is indeed very strong, and it is contestable whether such a strong thesis may hold, for it asserts it's impossible for subjects to hold erroneous beliefs about their experiences, a phenomenon known as infallibility.

[11]Liu speaks of subjects having experiences *with* phenomenal properties, which suggests that she's adopting a verison of what Nida–Rümelin (MS, 2018) calls the experience–property framework, whereby experiences instantiate phenomenal properties, rather than being themselves instantiations (by subjects) of phenomenal properties. Since in this work I am adopting a characterization of experiences as instantiations of phenomenal properties rather than as instantiating phenomenal properties, whenever below, just to stick with the lettere of Liu, I will talk of experience *with* phenomenal properties this should be intertpreted as referring to the phenomenal properties whose instantiation the experiences essentially consist in.

Regarding PHENOMENAL NARD1, instead, versions of this claim have been defended by several philosophers (including, Moran 2001: 14; Gertler 2012: 107; Horgan 2012: 406, Liu, *Ibid.*, just to mention some). Some (see, e.g., Schwitzgebel 2008: 263; Gertler 2011: 107, Liu, *Ibid.*) have even described it as being almost indefeasible.

As for NARD2, it states that «there is no distinction between what the essences of phenomenal properties appear to the subject to be and what they really are» or, equivalently, that «all there is to the essence of Q is X if and only if in having an experience–token with Q, the essence of Q appears to S as X and only as X» (Liu, *ivi:* 9). Again, we may possibly distinguish between an epistemic and a phenomenal version of NARD2. However, it seems quite implausible to say that the nature of experiences is such that subjects could not but form the belief that they have such–and–such nature. Hence, it is much more natural to interpret NARD2 just phenomenally. According to PHENOMENAL NARD2, then, «all there is to the essence of Q is X if and only if in having an experience–token with Q, S experiences the essence of Q as X and only as X» (Liu, *Ibid.*).

PHENOMENAL NARD2 was clearly already present in Kripke (1980), as it appears from these passage, among others (*ivi:* 150–3, emphasis mine):

> What was the strategy used above to handle the apparent contingency of certain cases of the necessary *a posteriori*? The strategy was to argue that although the statement itself is necessary, someone could, qualitatively speaking, be in the same epistemic situation as the original, and in such a situation a qualitatively analogous statement could be false. [...] In the appropriate sentient being is it analogously possible that a stimulation of C-fibers should have existed without being felt as pain? If this is possible, then the stimulation of C-fibers can itself exist without pain, since *for it to exist without being felt as pain is for It to exist without there being any pain. [...] To be in the same epistemic situation that would obtain if one had a pain is to have a pain; to be in*

> *the same epistemic situation that would obtain in the absence of a pain is not to have a pain. [...] the notion of an epistemic situation qualitatively identical to one in which the observer had a sensation S simply is one in which the observer had that sensation.* [...]

It is also clear that Chalmers endorses at least PHENOMENAL NARD2, as it can be seen from this passage, among others (2010_a: 149–50):

> As Kripke noted, there does not seem to be the same strong dissociation between appearance and reality in the case of consciousness as in the case[s] of water [...]. While it is not the case that anything that looks like water is water [...] it is plausibly the case that anything that feels like consciousness is consciousness. So it is not clear that the notion of [...] something that satisfies the primary intension of consciousness without being consciousness is coherent [...] so there is a strong case that the primary and secondary intensions of phenomenal terms coincides.

Now, phenomenal transparency claims that phenomenal concepts reveal the essence of the phenomenal properties they refer to, in the sense that it is *a priori*, for someone possessing the concept and in virtue of possessing the concept, what it is for the phenomenal property in question to be part of reality (Goff, 2017: 15). Hence, assumed phenomenal transparency and assumed that having a single token–experience E already puts a subject in a position to acquire a phenomenal concept of the phenomenal property instantiated in E, Liu's (forth: 3) definition of the thesis of revelation follows: «By having an experience–token with phenomenal property Q, S is in a position to know that 'Q is X', where the predicate 'X' captures the essence of Q». As Liu (*ivi:* 10) notes, this definition of the thesis of revelation combines to claim, namely (1) By having an experience–token with phenomenal property Q, S is in a position to know that 'Q is X' and (2) The predicate 'X' captures the essence of Q.

Furthermore, PHENOMENAL NARD2 has been defined consisting in these two claim: (3) all there is to the essence Q is X and (4) in having an experience–token

with Q, S experiences the essence of Q as X and only as X. more specifically, PHENOMENAL NARD2 states that (3) is true iff (4) is true.

Now intuitively – indeed, I would say almost trivially – (2) follows from (3): if the predicate 'X' captures the essence of Q, then all there is to the essence of Q is X. Later on – see 2.10, 2.13 and 2.14 – the fact that (2) implies (3) will play a pivotal role in my arguments to the case that both phenomenal transparency and phenomenal translucency imply the falsity of physicalism. Moreover – and, again, rather straightforwardly – (4) seems to follow from the thesis of revelation – *i.e.*, from the combination of (1) and (2). That is to say: from the fact that by having an experience–token with phenomenal property Q, S is in a position to know that 'Q is X', where the predicate 'X' captures the essence of Q, it seems to follow that in having an experience–token with Q, S experiences the essence of Q as X and only as X.

And now I will follow Liu (*ivi:* 25) in showing that given these assumptions, the thesis of revelation implies PHENOMENAL NARD2.

First, allow me to recapitulate where we are. We've got that the thesis of revelation consists in the combination of two claims: (1) By having an experience–token with phenomenal property Q, S is in a position to know that 'Q is X' and (2) The predicate 'X' captures the essence of Q. Moreover, we've got (plausibly) that (2) implies (3) that all there is to the essence Q is X. Finally, we've got that (again, plausibly), that (1)∧(2) – *i.e.*, the thesis of revelation – implies (4) that in having an experience–token with Q, S experiences the essence of Q as X and only as X.

Now, from (1)∧(2) follows (2). Assumed that from (2) follows (3), we've got that (1)∧(2), namely the thesis of revelation, implies (3). That is to say, from the fact that by having an experience–token with phenomenal property Q, S is in a position to know that 'Q is X', where the predicate 'X' captures the essence of Q follows that all there is to the essence Q is X. Moreover, from (1)∧(2) and from the assumption that (1)∧(2) implies (4) we obtain (4). That is to say, assumed that by having an experience–token with phenomenal property Q, S is in a position to know that 'Q is

X', where the predicate 'X' captures the essence of Q, we obtain that in having an experience–token with Q, S experiences the essence of Q as X and only as X.

Hence, assumed the thesis of revelation plus a couple of other rather plausible claims, both (3) and (4) are implicated. That is to say, assumed revelation, it could not be the case that (3) is true and (4) is false and, likewise, it could not be the case that (4) is true and (3) is false: hence, *under the assumption of revelation* (3) and (4) are logically equivalent. But that (3) and (4) are logically equivalent is precisely what PHENOMENAL NARD2 states, for recall that such principle states that all there is to the essence of Q is X if and only if in having an experience–token with Q, S experiences the essence of Q as X and only as X. Thus, the combination of the thesis of revelation with a couple of other rather plausible assumptions implies PHENOMENAL NARD2.

So, we've got that phenomenal transparency as defined in 1.8.1 combined with a plausible assumption – namely, that undergoing a single token experience E already puts a subject in a position to form a phenomenal concept of the phenomenal property instantiated in E – implies the thesis of revelation as defined by Liu. Plus, we have that the thesis of revelation as defined by Liu together with two further plausible assumptions implies PHENOMENAL NARD2. From this follows that phenomenal transparency plausibly implies PHENOMENAL NARD2. Finally, we have that denying that phenomenal concepts have identical primary and secondary intension implies the denial of PHENOMENAL NARD2. It follows that denying that phenomenal concepts have identical primary and secondary intension implying denying phenomenal transparency.

In the next session, I will show that phenomenal transparency is incompatible with physicalism, regardless of whether physical and phenomenal concepts are taken to be cognitively independent or not.

2.10 Why phenomenal transparency plus the denial of the cognitive independence of physical and phenomenal concepts is incompatible with physicalism

We saw that the primary ideal negative conceivability of P∧¬Q is logically equivalent with the cognitive independence of physical and phenomenal concepts. Moreover, in light of the remarks drawn in 2.8 and 2.9, as far as I can see provided certain rather plausible assumptions are in place the claim that phenomenal concepts have identical primary and secondary intension is logically equivalent with the thesis of phenomenal transparency – *i.e.*, the two are either both true or both false – for as we saw in 2.2.3 and 2.8, it seems very reasonable to infer from what Chalmers argues (see Tomasetta, 2012: 51–2; Goff, e.g., 2011) that, given that the primary intension of a concept can be evaluated *a priori*, if it is true that phenomenal concepts have identical primary and secondary intensions then phenomenal transparency has to be true as well. Moreover, as we saw in 2.9, if it is false that phenomenal concepts have identical primary and secondary intension, then phenomenal transparency has to be false as well. Thus, given certain very plausible assumptions phenomenal transparency and the claim that physical and phenomenal concepts have identical primary and secondary intension are arguably either both true or both false, hence logically equivalent.

Now, assume phenomenal transparency. That is, assume that in virtue of having a given phenomenal concept C_{phen} and just in virtue of having C_{phen}, a subject is put in a position to know the *entire* nature of the phenomenal property C_{phen} refers to – *i.e.*, all its essential properties. Assume, moreover, that physical and phenomenal concepts are not cognitively independent, hence, P∧¬Q is not conceivable, hence P∧Q is conceivable – *i.e.*, for any arbitrary phenomenal concept C_{phen} there is at least one physical concept C_{phis} such that a subject mastering both C_{phen} and C_{phis}

and accepting the relevant piece of background physical knowledge as true would on that basis alone be in a position to infer that C_{phen} and C_{phis} are necessarily coextensive.

I think (some) physicalists do not take the thesis of revelation seriously enough: they do not take it at face value. Let me explain.

I take it that if phenomenal concepts are transparent, then the nature of phenomenal properties is essentially entirely phenomenal. This follows from the very definition of what a transparent concept is. If mastering a phenomenal concepts is sufficient *per se* for a subject to grasp the *entire* nature of the phenomenal property the concept refers to, then as far as I can see, necessarily *all there is* to the nature of phenomenal properties is their being such that whenever one of them is instantiated, it is like W for a subject to live through that moment. To use a paradigmatic example of a transparent concept (see Goff, e.g., 2011), consider the concept <SPHERE>: arguably, necessarily all there is to the nature of spheres is their being sets of points that are all at the same distance r from a given point in three–dimensional space. As Nida–Rümelin (2007: 32) puts it: «to grasp a property is to be cognitively presented with what is essential for having that property. There is no room for any hidden nature behind the property that we are grasping».

Notice, crucially, that this does not imply *per se* that the nature of one and the very same entity may be grasped in two or more conceptually distinct ways. This point will be of some importance when it will come to discussing an argument for physicalism which draws upon the so called thesis of dual carving, namely, the idea whereby two or more concepts may be both transparent with respect to the same entity and cognitive independent.

Assumed phenomenal transparency and assumed that physical and phenomenal concepts are not cognitive independent, there are four options available: under (option 1) phenomenal concepts are transparent and physical concepts are transparent; under (option 2) phenomenal concepts are transparent and physical concepts are translucent; under (option 3) phenomenal concepts are transparent and physical concepts are

mildly opaque; finally, under (option 4) phenomenal concepts are transparent and physical concepts are radically opaque. My aim is to show that none of these options is compatible with physicalism. I will examine each option in due order. I will start with option 1. Then I will proceed to option 2. Option 3 and option 4 will be discussed together.

2.10.1 Option 1

On this option, both phenomenal and physical concepts are transparent and, as on any other of the options we are considering, phenomenal and physical concepts are not cognitively independent. Now, in light of remarks just drawn regarding phenomenal transparency, I take it that if physical concepts are transparent, then the nature of physical properties is essentially entirely physical. Moreover, in virtue of the denial of the cognitive independence of physical and phenomenal concepts, we have that every phenomenal concept is necessarily coextensive with at least one physical concept, hence every phenomenal property is a physical property. It follows that the nature of "physico–phenomenal" properties is *both essentially entirely phenomenal and essentially entirely physical*. Call this the *PP (physico–phenomenal) view*.

As far as I can see, the PP view puts the physicalist in a rather weird position. This is because most physicalists would accept what Strawson (2008_{a}: 60) calls the NE (Non–Experiential) principle: «physical stuff is, in itself, in its fundamental nature, something wholly and utterly non–experiential». Indeed, as we have seen it is one of the necessary conditions for conventional physicalism to be true that there aren't fundamental experience–involving entities, and several physicalist philosophers merely equate the physical with the essentially non–experiential. Maybe the PP view leads to a form of idealism. At least under most interpretations idealism is characterized as a view whereby the ontologically fundamental entities are essentially mental or, put another way, «something mental (the mind, spirit, reason, will) is the ultimate foundation of all reality, or even exhaustive of reality» (Guyer and Horstmann, 2023). Maybe the PP view is best characterized as a form of panpsychism, the latter being

characterized, as per Chalmers (2013: 3), as «the thesis that everything is (or at least some things are) fundamentally physical and fundamentally mental».

An objector might want to reply that the PP view does not imply, necessarily, that physico–phenomenal entities are fundamental (though it is not incompatible with it): they might be supervenient upon (grounded on/realized by/emergent from/...) some more fundamental entities.

However, as we saw in 1.8 the physicalist would better assume that the nature of phenomenal properties is cognitively accessible via physical concepts, the latter being defined as those concepts that are deployed in our best physical theories. And as we saw in 1.2, it seems a reasonable constraint to impose on physicalist theories that in order for them to be true at least some of the entities physics deals with should be fundamental. Indeed, as we shall see momentarily, it is reasonable to suppose that according to an overall physicalist worldview the entire world should be "made up", so to say, of fundamental physical properties. Given this, the PP leads quite naturally, it seems to me, to the conclusion that at least some of the fundamental entities physics deals with are to be essentially experience–involving, which violates the second, negative *desideratum* of conventional physicalism, namely that there are no fundamental experience–involving entities.

A possible reply might go as follows: although several physical concepts refer to some fundamental entities, at least some of them refer to some entities that, although not fundamental are metaphysically necessitated by (grounded on/supervenient upon/realized by/...) the fundamental ones. So, proponents of this objection would contend, one might say that the nature of phenomenal properties is revealed by those and only those physical concepts which do not refer to some fundamental entity. I find such a view very implausible and hard to defend. But even if I was wrong and there was a way to defend it, it would still violate the NE principle as stated above, and I don't think many conventional physicalists would be happy with this. What is more, consider a distinction Elpidorou (2015), among others, has drawn between senses in which something might be taken to be physical. As Elpidorou (*Ibid.*) says

– and indeed, in line with a mainstream approach (see 1.2) – an entity is physical in a restricted sense if it is posited by some of our best physical theories. Entities like protons, quarks, *et cetera* are physical in this sense. An entity is physical in the broad sense if it is realized by, necessitated by, or metaphysically supervenient upon some entities that are physical in the restricted sense. Ordinary objects like tables, chairs and stones are physical in this broad sense. But now consider the property of 'being liquid' as an example. Such a property is arguably physical in the broad sense: no physical – in the restricted sense – knowledge seems to be needed to entertain the thought, say, that something is liquid. Yet as Strawson (2008_{a}: 61) puts it:

> The emergent character of liquidity relative to its non–liquid constituents does indeed seem shiningly easy to grasp. We can easily make intuitive sense of the idea that certain sorts of molecules are so constituted that they don't bind together in a tight lattice but slide past or off each other (in accordance with Van De Waals molecular interaction laws) in a way that gives rise to – is – the phenomenon of liquidity [...] we move in a small set of conceptually homogeneous shape–size–mass–charge–number–position–motion–involving–physics notions with no sense of puzzlement. Using the notion of reduction in a familiar loose way, we can say that the phenomena of liquidity reduce without remainder to shape–size–mass–charge–etc. Phenomena [...] We can see that the phenomenon of liquidity arises naturally out of, is wholly dependent on, phenomena that do not in themselves involve liquidity at all.

If Strawson is right, and I think he is, then physical concepts in the broad sense are not cognitive independent from physical concepts in the restricted sense.

More in general, it seems fair to say that under a physicalist worldview only fundamental physical entities carve nature at its joints. In other words, under a physicalist worldview only physical properties are *sparse*, where Lewis (1986: 60) defines sparse properties as follows (see also Goff, 2010; O'Connail, MS):

> Sharing of them makes for qualitative similarity, they carve nature at the joints, they are intrinsic, they are highly specific, the sets of their instances are ipso facto not entirely miscellaneous, there are only just enough of them to characterise things completely and without redundancy.

If this is right, then in order for physicalism to be true *any concept whatsoever* should be ultimately cognitive inter–dependent – not cognitively independent – with some *fundamental* physical one. Or, assumed physicalism, if the latter weren't the case – that is, if not every concept turned out as ultimately cognitively inter–dependent from some fundamental physical one – then this would be due to some intrinsic limitation of our cognitive architecture, that is, to our intrinsic incapacity to get to see that only fundamental physical properties carve nature at its joints. But the latter would nonetheless have to be true in order for physicalism to be true.

Thus, arguably, even if phenomenal concepts were taken to be cognitively inter–dependent with only those physical concepts which do not refer to some fundamental entity, these latter concepts, in turn, would most likely ultimately turn out to be cognitive inter–dependent with some physical concepts referring to some fundamental entity. And since the relation of cognitive inter–dependence is transitive – meaning that if the concept C is cognitive inter–dependent with the concept C, and C_1 is cognitive inter–dependent with C_2, C is cognitive inter–dependent with C_2 – from this would still follow that any phenomenal property is necessarily a fundamental physical property, contra the negative *desideratum* of physicalism. Or, again, assumed physicalism if the latter weren't the case this would only be due to some intrinsic limitation of our cognitive architecture.

An objector could now urge that sponsors of the PP view need not say that any property whatsoever has to be essentially wholly phenomenal and essentially wholly physical: they just have to say that some properties – phenomenal ones – are of this sort. But provided the remarks I have drawn regarding fundamentality, physicalism and the transitivity of the relation of cognitive inter–dependence are right, this view would still lead to the conclusion that some fundamental entities are essentially

wholly physical and essentially wholly phenomenal, *contra* the negative *desideratum* of physicalism. That is to say, although such a view would not count as a form of *pan*psychism, it would still count as a form of *micro*physichism, as Strawson (2008_a) calls it, namely a view on which although not all fundamental physical entities are essentially experience–involving, at least some of them are of this sort.

I won't take a stance here on how the details of the PP view would better be fleshed out. Although I happen to believe that something close to the PP view may be true, I find such a view blatantly at odds with physicalism. That is because as far as I can see the NE principle or at least something very close to it lies at the very core of the physicalist worldview. As Braddon–Mitchell and Jackson (2007:20) emphatically write:

> There is a clear idea behind the materialist's thought that the mental is physical through and through. The idea is that the mental is a very complex arrangement of items having the kinds of properties [...] that non–sentient non–minded items have. The idea is that the building blocks of mind and feelings are items that lack mind and feeling. Just as a powerful car is made up of bits that are not powerful – it is the mode of composition that creates the power – so we minded and sentient beings are made up of suitably arranged bits that are neither minded nor sentient.

2.10.2 Option 2

Reader who have managed to follow me up to this point, should be able to see rather straightforwardly why it is the case that the combination of phenomenal transparency and physical translucency plus the denial of the cognitive independence of physical and phenomenal concepts is most likely incompatible with physicalism. From phenomenal transparency follows that phenomenal properties are essentially entirely phenomenal. From physical translucency follows that physical properties are partially essentially physical. Assumed a dispositionalist conception of theoretical physics this amounts,

roughly, to the claim that part of the nature of physical properties is that they play such–and–such dispositional/functional/causal roles, even though the nature of (fundamental) physical properties is not exhausted by their playing those roles. From the denial of the cognitive independence of physical and phenomenal concepts and from the remarks just drawn in 2.10. regarding physical concepts, fundamentality and the transitivity of the relation of cognitive inter–dependence follows at least that some fundamental physical entities are essentially entirely phenomenal and hence essentially experience–involving, *contra* the negative *desideratum* of physicalism, and essentially partially physical.

2.10.3 Options 3 and 4

As on any other of the options we are considering, both on option 3 and on option 4 it is assumed that physical and phenomenal concepts are not cognitively independent, and that phenomenal concepts are transparent. Option 3 assumes that physical concepts are mildly opaque, whereas option 4 assumes that physical concepts are radically opaque.

Now, as I have already said – 1.8.4 – there are very strong *prima facie* reasons to be skeptical as to whether physicalists may be happy with the possibility of physical concepts being either mildly opaque or radically opaque, for the latter view implies that it is *only contingent* that phenomenally conscious mental states have whatever feature or combination of features physics ascribes to them.

Moreover, both option 3 and option 4 would entail an extremely radical revision of what counts as physical which, allow me to say, is something I could potentially be on board with, but this is beyond the point here. For on both option 3 and option 4 while on the one hand physics tells us either nothing at all or very little of what it is for physical properties to be part of reality, introspection, via phenomenal concepts, tells us *everything* about what it is for physical properties, or at least for physico–phenomenal ones, to be part of reality – which, again, is not *per se* incompatible with the possibility of the essence of one and the very same entity being

grasped in two or more conceptually distinct ways. Moreover, in light of phenomenal transparency, on both option 3 and option 4 – as well as on option 1 and option 2, by the way – the physical is construed as such that its nature is revealed to subjects in introspection *simply by being experienced* (see Nida–Rümelin, 2007: 332). Although this might be something I would be on board with (but again, this is way beyond the point here), to my knowledge, no one who declares herself as physicalist without further specification would deem this as an acceptable view to hold.

2.11 Tacking stock (II)

We saw that, since under Chalmers' two–dimensional semantic apparatus conceivability is arguably logically equivalent with primary possibility both under an epistemic approach to scenario and under a metaphysical one – meaning that if a statement is conceivable it has to be primarily possible and *vice versa* – the only way to attack the CP principle is to deny that P∧¬Q is conceivable in the very first place where, recall, P∧¬Q consists in the statement that there is the physical world and nothing else, and no one is conscious. We also saw that the primary ideal negative conceivability of P∧¬Q is arguably logically equivalent with the cognitive independence of physical and phenomenal concepts as defined in 1.8.2 – *i.e.*, denying the conceivability of P∧¬Q commits one to say that for any arbitrary phenomenal concept C_{phen} there is at least one physical concept C_{phis} such that a subject having C_{phen} and C_{phis} and accepting the relevant background physical knowledge as true would on that basis alone be in a position to rationally judge that C_{phen} and C_{phis} are necessarily coextensive. Version 1 and version 2 of the conceivability argument assume that phenomenal concepts have identical primary and secondary intension which, provided certain minimal and rather plausible assumptions are in place, implies the thesis of phenomenal transparency. In 2.10 I argued that the combination of the denial of the cognitive independence of physical and phenomenal concepts with phenomenal transparency is incompatible with physicalism, regardless of how one construes physical concepts, namely, regardless of whether one takes physical concepts to

be transparent, translucent, mildly opaque or radically opaque. Now, suppose we assume that P∧¬Q is not conceivable, hence that P∧Q is conceivable, hence that physical and phenomenal concepts are not cognitively independent, thus obtaining four versions of the anti–zombie argument, namely, four updated versions of the original zombie argument that are perfectly specular the original four ones but are premised on the conceivability of P∧Q. Provided I'm right that that on Chalmers' two–dimensional semantic apparatus (primary ideal negative) conceivability and primary possibility are logically equivalent, the conceivability of P∧Q *eo ipso* implies its primary possibility. We might mark each updated version of argument with a * to distinguish them from the original versions.

Now, given all that we've seen thus far, both version I* and version II* of the anti–zombie argument must be ruled out, for they assume phenomenal transparency and the denial of the cognitive independence of physical and phenomenal concepts, and since the latter two, taken together, are incompatible with physicalism, these two versions are incompatible with physicalism as well. This leaves us with version III* and version IV*. These versions assume the primary and secondary intension of phenomenal concepts differ. Hence, phenomenal transparency is not assumed. Here is how I will proceed in what follows. In 2.12 I will argue, along with Goff (2011: 194–5), that the overall two–semantic apparatus Chalmers adopts is arguably incompatible with the existence of radically opaque concepts. This leaves with two options: either in version III* and version IV* of the anti–zombie argument phenomenal concepts are assumed to be translucent, *i.e.*, to reveal part of the nature of the experiences they refer to, but not the entirety of it, or it is assumed that they are mildly opaque, *i.e.*, that they do not reveal any essential property of experiences, but they reveal some property/ies of them which uniquely identify them only in the actual world. I will then show – 2.13 – that the combination of the denial of the cognitive independence of physical and phenomenal concepts with phenomenal translucency is incompatible with physicalism, again, regardless of how one treats physical concepts. I will therefore conclude that the only option left available to the

physicalist who wanted to deny the conceivability of P∧¬Q and thus the cognitive independence of physical and phenomenal concepts is to deny both phenomenal transparency and phenomenal translucency, claiming that phenomenal concepts are mildly opaque.

2.12 Why Chalmers' overall two dimensional framework is incompatible with the existence of radically opaque concepts

The primary intension of a concept C returns a set of superficial properties which, taken all together, at least uniquely identify the referent of C in the actual world – in the case of water, being transparent, being liquid, being thirst–quenching, filling the oceans, *et cetera*. Moreover, in virtue of the PI–AP principle (see Tomasetta, 2012: 51–2, Goff, 2011: 194–5) it is reasonable to say that the primary of intension of C can be evaluated *a priori*, meaning that a subject mastering C is *eo ipso* endowed with a certain amount of *a priori* knowledge which allows her to pick the primary intension of it in every scenario/metaphysically–possible–world in which it is not empty. As Goff argues, and, to my eyes, rightly so, combined together these two claims entail that there cannot be radically opaque concepts, for they imply that whoever mastered an arbitrary concept C would *eo ipso* be endowed with an amount of *a priori* knowledge regarding some properties which at least uniquely identify the referent of C in the actual world. But this entails that any arbitrary concept is at least mildly opaque, for recall that a concept C reveal some properties P, Q, R... of its referent iff it is *a priori*, for someone having C and just in virtue of having C, that there referent of C has P, Q, R, *et cetera*.

I won't take a stance here on whether it is really plausible to say that there cannot be radically opaque concepts, namely concepts that do not reveal anything of their referents, or whether, instead, as Goff seems to suggest, the latter claim is contentious and Chalmers' two–dimensional semantic apparatus should be appropriately amended

so as to allow for the possibility of there being concepts of this sort. For my purposes, it is sufficient to note that as it stands by now Chalmers' two–dimensional apparatus is arguably incompatible with the existence of radically opaque concepts.

By the way, this fact about Chalmers' two–dimensional semantics automatically rules out the fourth of the options seen above – 2.10.3 – namely the option whereby phenomenal concepts are transparent and physical concepts are radically opaque. But as we saw, such an option would most likely be inconsistent with physicalism anyway.

2.13 Why phenomenal translucency plus the denial of the cognitive independence of physical and phenomenal concepts is incompatible with physicalism

Suppose now phenomenal concepts were taken to be translucent, meaning that they reveal only part of the nature of experiences, but not all of it. This entails that phenomenal properties are essentially partially phenomenal. Now, again, there are three options available, which I will mark with a °: on option (1°), phenomenal concepts are translucent and physical concepts are transparent; on option (2°) phenomenal concepts are translucent and physical concepts are translucent; on option (3°) phenomenal concepts are translucent and physical concepts are mildly opaque; I won't consider the option whereby phenomenal concepts are translucent and physical concepts are radically opaque for as we saw this option is most likely inconsistent with Chalmers' two–dimensional semantics, which as we just saw rules out the possibility of there being radically opaque concepts. Such an option would most likely turn out to be at odds with physicalism anyway.

Now, given phenomenal translucency, the denial of the cognitive independence of physical and phenomenal concepts, and the remarks I drawn in 2.10.1 regarding

fundamentality, physical concepts, and the transitivity of the relation of cognitive inter–dependence between concepts, it follows that (some of) the fundamental entities physics deals with are essentially partially phenomenal and hence essentially partially experience–involving. This arguably amounts to a form of Russelian monism, and as such violates the negative *desideratum* of physicalism. This holds on any of the options (1°)–(3°), that is, regardless of wether physical concepts are taken to be transparent, translucent or mildly opaque. Yet I want to spend a few extra–words on each option. On option (1°) we have that fundamental physical entities have at least one essential component which is both essentially physical and essentially phenomenal/experience–involving contra the second *desideratum* of physicalism. On option (2°) we have that both physical and phenomenal concepts are translucent. Now, there are a bunch of sub–options. According to an option which we might call option ($2^{\circ a}$), physical and phenomenal concepts reveal the same essential part of the nature of fundamental physico–phenomenal entities, although in conceptually distinct ways. Again, this entails that part of the nature of fundamental physical entities is both essentially physical and essentially experience–involving. Moreover, note that the part in question could not exhaust the nature of experiences, for otherwise both physical and phenomenal concepts would turn out to be transparent, which is note the case by assumption. In light of this, option ($2^{\circ a}$) splits into two further sub–sub–options. On option ($2^{\circ a1}$) the "missing extra essential ingredient(s)" is/are revealed, or at least may be revealed in principle, by some other concepts which are neither physical nor phenomenal. Frankly, it is quite hard to see what sort of putative "extra–concepts", so to say, could do the job. Moreover, as far as I can see this sub–option would lead either to a form of property dualism, on which there are two ontologically distinct essential components to the nature of phenomenal properties[12] or to a form of pluralism whereby there are several – more than two – ontologically distinct essential aspects or components to the nature of phenomenal properties. On option ($2^{\circ a2}$) the "missing extra essential ingredient(s)" is/are not and

[12]By the way, this point is shared by a number of physicalists such as Lewis (1980: xVIII) and Papineau (2002: 82). See also Goff (2011: 97)

could not, not even in principle, be revealed by any kind of concepts whatsoever. On this sub–sub–option, by assumption no characterization of the universe, no matter the concepts it deploys and no matter how extensive, could ever get us to know the entire nature of experiential properties. As far as I can see, this is a form of mysterianism, and I would say that not many physicalists would be comfortable with it, for intuitively, under a physicalist worldview even very broadly understood a complete physical characterization of the universe would, indeed get us to know the entire nature of phenomenal properties – indeed, presumably, the entire nature of *everything*.

As for option ($2^{\circ b}$), on this sub–option physical and phenomenal concepts are not taken to reveal the same essential part of the nature of experiential properties, but rather two or more distinct essential aspects or features of them. Arguably, again, this amounts at least to a form of property dualism, on which there are at least two ontologically distinct essential aspects of experiential properties – and possibly to a form of pluralism. Property dualism is clearly incompatible with the negative *desideratum* of physicalism. What is more, again there are two sub–sub–options available. Either ($2^{\circ b1}$) the combination of what physical and phenomenal concepts reveal of experiential properties is sufficient to exhaust their nature or ($2^{\circ b2}$) the combination of what physical and phenomenal concepts reveal of experiential properties is not sufficient to exhaust their nature. In turn, option ($2^{\circ b2}$) splits into two sub–sub–sub–options. On option ($2^{\circ b2.1}$) the part of the nature of experiences which is not revealed neither by physical concepts nor by phenomenal ones is revealed, or at least may be revealed in principle, by some putative "extra–concepts", neither physical nor phenomenal. The same remarks drawn regarding option ($2^{\circ a1}$) holds for option ($2^{\circ b2.1}$) as well. On option ($2^{\circ b2.2}$) the missing essential ingredient(s), namely the one(s) that is/are not revealed neither by phenomenal concepts nor by physical ones is/are not revealed nor revealable, not even in principle, by any concept whatsoever. Again, this leads to a form of mysterianism. See the remarks drawn regarding option ($2^{\circ a2}$).

By the way, as far as I can see from the arguments just expressed follows that physical translucency may be incompatible with physicalism in general, that is, no matter the stance one takes as for what regards phenomenal concepts. For if physical concepts are taken to be translucent then by definition they cannot reveal the entirety of the nature of the physical properties they are about. Now either the part of the nature of physical properties that is not revealed by physical concepts is revealed or revealable in principle but some other concepts, or such part is not revealable, not even in principle, by any concept whatsoever. The former option implies either a form of property dualism or a form of pluralism; the latter option leads to a form of mysterianism whereby no combination of concepts could ever get us to know the entire nature of physical properties, and there are reason to doubt that none of these options fits with an overall physicalist worldview. This is yet another reason, besides the ones offered in 1.8.4, for believing that the fourth premise of the argument from revelation – namely the premise whereby in order for physicalism to be true, for every experience there must be a transparent physical concept of it – is most likely true. Moreover, it might lead to the conclusion that the second and the fourth version both of the original zombie argument and of the anti–zombie argument should be ruled out, for they assume that physical concepts have distinct primary and secondary intension, and hence are either translucent or mildly opaque (provided Chalmers' two–dimensional semantics is incompatible with the existence of radically opaque concepts).

Regarding option (3°), under which, recall, physical concepts are taken to be mildly opaque, it still implies that fundamental physical entities have essentially at least one essentially experience–involving component, namely the one(s) revealed by phenomenal concepts, and hence it still violates the negative *desideratum* of physicalism. Moreover, as was the case with the options (3) and (4) seen in 2.10, option (3°) would still imply that it is only contingent that phenomenally conscious mental states have the features physics ascribes to them, and it would require a radical revision of what counts as a physical property(/state/process/...) whereby while, on

the one end, physics tells us very little about the nature of physical properties, on the other hand the latter are such that part of their essential nature is revealed to subjects just in virtue of their being experienced. Again, this is clearly at odds with most mainstream forms of physicalism.

2.14 Why Revelation threatens physicalism also under the assumption that physical and phenomenal concepts are cognitively independent

We saw that, given certain assumptions regarding Chalmers' two–dimensional semantics, primary ideal negative possibility and primary possibility are most likely logically equivalent both under an epistemic approach to scenarios and under a metaphysical one. Hence, the only way to attack the CP principle is to deny that zombies are conceivable in the very first place. This amounts to the claim that phenomenal and physical concepts are not cognitively independent. On the assumption the the arguments I have drawn are correct, both phenomenal transparency and phenomenal translucency are inconsistent with physicalism when combined with the denial of the cognitive independence of physical and phenomenal concepts. Given this and the fact that Chalmers' two–dimensional semantic apparatus is arguably incompatible with the existence of radically opaque concepts, to rebut the conceivability argument physicalists who wanted to attack the CP principle are committed to deny both phenomenal transparency and phenomenal translucency and to argue that phenomenal concepts are mildly opaque, that is, they do not reveal any essential property of their referents, phenomenal properties, but reveal some property(/ies) which uniquely identify their referents only contingently, that is only in the actual world.

But now a question might arise. What about accepting that $P \wedge \neg Q$ is, indeed, conceivable, and hence that physical and phenomenal concepts are cognitively in-

dependent? First of all, if I'm right that primary possibility is logically equivalent with primary ideal negative conceivability on Chalmers' two–dimensional apparatus both under a metaphysical approach to scenarios and under an epistemic approach, then from the fact that P∧¬Q is conceivable *eo ipso* follows that P∧¬Q is primarily possible. Hence, the four original versions of the conceivability argument would be restored. So, the physicalist who wanted to concede that physical and phenomenal concepts are cognitively independent would have the burden of showing hat the four original versions of the argument are faulty or invalid.

More in general, consider the following line of reasoning, which I've already been suggesting throughout the chapter as well as in other places – e.g., 1.8.4. Assume that P∧¬Q is conceivable hence primarily possible, provided conceivability is logically equivalent with primary possibility. By the law of the excluded middle, either (A) P∧¬Q is conceivable/primarily–possible and secondarily possible or (B) P∧¬Q is conceivable/primarily–possible and not secondarily possible. (A) must be immediately ruled out, for it states that since P∧¬Q is secondarily possible, there is at least one metaphysically possible world which is a minimal exact physical duplicate of the actual world but not a C–duplicate of the actual world, namely a metaphysically possible world in which there is the physical world and nothing else and no one is conscious, *contra* the first, positive *desideratum* of physicalism. Therefore, the physicalist is committed to go for (B). In light of the denial of the secondary possibility of P∧¬Q, (B) states that it is not metaphysically possible that there is the physical world *and nothing else* and no one is conscious, that is to say, whenever a metaphysically possible world W is a minimal exact physical duplicate of the actual world, *eo ipso* in W there must be consciousness. Now, combined with the cognitive independence of physical and phenomenal concepts, hence with the conceivability/primary–possibility of P∧¬Q, this option entails that while given an arbitrary phenomenal concept C_{phen}, there is no physical concept C_{phis} such that a subject mastering both C_{phen} and C_{phis} and accepting an arbitrary amount of background physical knowledge as true would be in a position to see

that C_{phen} and C_{phis} are necessarily coextensive, C_{phen} and C_{phis} *must*, indeed, be necessarily coextensive. In other words, (B) implies that although for some reason – say due to our intrinsic cognitive limitations, or to some peculiar feature(s) of phenomenal concepts – we will never be in a position to rationally judge that every phenomenal property is a (fundamental) physical property, it is nonetheless the case that, as a matter of metaphysical necessity, every phenomenal property *is*, indeed, a (fundamental) physical property, for otherwise $P \wedge \neg Q$ would turn out to be secondarily possible.

But now the remarks drawn so far regarding phenomenal transparency and phenomenal translucency may be reiterated, for assumed phenomenal transparency follows that at least some fundamental physical entities are both essentially entirely physical and essentially entirely phenomenal, hence essentially experience–involving, *contra* the negative *desideratum* of physicalism, and assumed phenomenal translucency follows, again, that at least some fundamental physical entities have at least one essential component which is essentially phenomenal thus essentially experience–involving, which arguably amounts to a form of Russelian monism and as such violates the negative *desideratum* of physicalism as well.

Thus, once more, both phenomenal transparency and phenomenal translucency are inconsistent with physicalism whereby, provided Chalmers' two–dimensional semantics is incompatible with the existence of radically opaque concepts, follows that in order to rebut the conceivability argument physicalists must claim that phenomenal concepts are mildly opaque. Moreover, and crucially, provided that in order for physicalism to be true $P \wedge \neg Q$ cannot be metaphysically possible, the latter point holds regardless of whether one concedes that $P \wedge \neg Q$ is primarily ideally negative conceivable and hence primary possible or not, that is, regardless of whether one concedes that physical and phenomenal concepts are cognitively independent, as per type–A physicalism, or not, as per type–B physicalism. One might want to reject Chalmers' two–dimensional semantic apparatus *in toto*. As far as I can see, the only result that would be thus obtained would be to allow for the possibility

of phenomenal concepts being radically opaque. Overall, it seems fair to say that the conceivability argument is not an argument about conceivability: rather, it's an argument about revelation. *Quod erat demonstrandum.*

Chapter 3

The knowledge argument, the modal argument and the argument from revelation

3.1 Introduction

In this chapter, I aim to make the case that Frank Jackson's (1982) renowned knowledge argument against physicalism[1] rests upon the thesis phenomenal transparency

[1]As some readers might already be aware of, Jackson later recanted his anti–physicalist stance and is now a committed physicalist, who has even declared to have sympathies for weak illusionism (see https://youtu.be/tdTSymICyf4). It is also worth underlining that the idea beneath the knowledge argument predates Jackson by a long way. A precursor to the argument can already be found in Locke's *Essay* (1689[1975]), which, however, is centered on the idea of an inverted spectrum. A version of the argument can also be found in Leibniz' *Monadology* (1714[1998]). Moreover, the argument has been foreshadowed in various ways in the recent literature. Typically, precursors to the knowledge argument employ one of two strategies, both of which are reminiscent of Jackson's formulation (see Nida–Rümelin and O'Conaill, 2023). The first strategy relies on what Stoljar and Nagasawa (2004: 2–3) term the *knowledge intuition*. This intuition asserts that having complete physical information concerning certain experiences is insufficient on its own to possess knowledge of what these experiences are actually like. The second strategy involves thought experiments similar to the one about Mary. In these thought experiments, there's a being who possesses comprehensive knowledge of the physical aspects of certain experiences but is claimed to lack knowledge of the actual subjective nature of those experiences (see Nida–Rümelin and O'Conaill, *Ibid.*). Examples of philosophers who deployed and/or discussed the first strategy include Russell (1912: 13–14), Dunne (1929), and Broad (1925), among others (Nida–Rümelin and O'Conaill, *Ibid.*). Examples of philosophers who deployed and/or discussed the second strategy include Feigl (1958), Farrel (1950), Meehl (1966) and Robinson (1982), among others (again, see Nida–Rümelin and O'Conaill, *Ibid.*). However, offering a detailed overview of how the idea underneath the argument has been formulated and/or deployed throughout the history of philosophy is beyond my purposes. The choice to stick with Jackson's version of the argument is ultimately arbitrary. As long as the main idea underneath the argument is clear, as far as I'm concerned readers may choose any version of

as one of its key–assumptions, meaning that it is necessary and possibly necessary and sufficient that physicalists deny both phenomenal transparency and phenomenal translucency as stated in 1.8.1 in order for them to avoid the threat posed by the argument. In the second chapter I argued that both phenomenal transparency and phenomenal translucency are inconsistent with physicalism, both under the assumption that phenomenal and physical concepts are cognitive independent, and under that assumption that they are not. Basically, as far as I can see the very same line of reasoning may be reiterated as for what concerns the knowledge–argument. Still, I whish to offer some further remarks. More specifically, after having briefly recapped my arguments, I will examine a bit more closely a number of possible physicalist responses to the knowledge–argument and argue further that in order for each of them to go through, it is at least necessary and possibly necessary and sufficient that both phenomenal transparency and phenomenal translucency are denied.

The story of Mary is so famous that I guess it wouldn't need to be told yet another time. However, just to make sure everyone is on the same page, I will quote once more the very much quoted passage in which Jackson's (1982:130) formulated the intuition beneath the argument:

> Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room via a black and white television monitor. She specializes in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on. She discovers, for example, just which wavelength combinations from the sky stimulate the retina, and exactly how this produces via the central nervous system the contraction of the vocal chords and expulsion of air from the lungs that results in the uttering of the sentence 'The sky is blue'.... What will happen when Mary is released from her black and white room or is given a color

the latter.

television monitor? Will she learn anything or not?

It might perhaps be helpful to start by examining the relations between this thought–experiment and Chalmers' conceivability–argument, which has already been discussed extensively in the second chapter.

Recall, first, that in order to resist the conclusion of (each the four version of) the conceivability–argument, physicalists are committed to deny that P∧¬Q is secondarily or metaphysically possible, where P is a statement which conjoins all the physical truths about the universe with a "that's all" clause – *i.e.*, roughly, a statement claiming that there is the entire physical world and nothing else – and Q is an arbitrary phenomenal statement. In fact, the metaphysical possibility of P∧¬Q would violate the first, positive *desideratum* of physicalism, namely the one whereby phenomenal properties must metaphysically supervene upon physical ones.

In 2.4 we also saw that, taken an arbitrary phenomenal concept C_{phen} – say <PHENOMENAL PAIN> – the (primary ideal negative: see 2.2) conceivability of P∧¬Q – where Q is taken to be the statement that the referent C_{phen} is instantiated, say something like 'there is pain' or 'someone is in pain' – is logically equivalent with the cognitive independence of C_{phen} with every arbitrary physical concept, meaning that if P∧¬Q is conceivable C_{phen} must be cognitively independent from every arbitrary physical concept, and if, on the contrary, P∧¬Q is not conceivable, then there must be at least one physical concept deployed in P that is not cognitively independent from C_{phen}. Suppose now we take <PHENOMENAL RED> as our relevant phenomenal concept, and hence we take something like 'there is the visual experience of red' or 'someone is seeing red', where the visual experience of red is characterized phenomenally, as our Q. Four options are available:

[1]P∧¬Q – *i.e.*, 'there is the entire physical world and nothing else, and no one is having the visual experience of red' – is conceivable. Hence, the phenomenal concept <PHENOMENAL RED> is cognitively independent from every arbitrary physical concept. P∧¬Q, however, is not metaphysically possible.

[2]P∧¬Q is conceivable and metphysically possible.

[3]P∧¬Q is not conceivable. Hence, there is at least one physical concept that is not cognitively independent from <PHENOMENAL RED>. P∧¬Q is also not metaphysically possible.

[4]P∧¬Q is not conceivable but metaphysically possible.

The second option and the fourth one must be ruled out by the physicalist, for the metaphysical possibility of P∧¬Q is inconsistent with the positive *desideratum* of physicalism, which states that the phenomenal must be metaphysically supervenient on the physical. This leaves us with the first option and the third one. Start with the third one.

In 2.2.1 we saw that a given statement is ideally primarily negatively conceivable iff it cannot be *a priori* ruled out by a subject, not even on ideal rational reflection. To say that P∧¬Q is not conceivable, thus, is to say that it can be *a priori* ruled out by a subject, at least on ideal rational reflection. Now, P states that there's the entire physical world and nothing else. Thus, the inconceivability of P∧¬Q implies that an ideal rational subject would be able to infer Q – namely that there is the visual experience of red, characterized phenomenally – from P: that is, she would be in a position to infer *a priori* that whenever in a given metaphysically possible world there's the entire physical world and nothing else, *eo ipso* there must be the visual experience of red characterized phenomenally. As we saw in 2.4, this implies, in turn, that the phenomenal concept <PHENOMENAL RED> is cognitively inter–dependent with at least one of the physical concepts deployed in P.

By assumption, Mary has complete physical knowledge about color vision, that is, she knows every physical fact, or at least all the physical facts that are relevant for color vision. We may also assume that there is no cognitive, environmental, or whatever other factor impeding Mary's ideal rational reflection. As far as I can see, the assumption that P∧¬Q is not conceivable, together with the way in which the thought–experiment is formulated, entails that Mary does not learn anything upon her release, for the inconceivability of P∧¬Q implies that a ideal rational agent would be in a position to *a priori* infer Q from P, and by stipulation Mary knows P

– or at least a statement, contained in P, which conjoins all and only the physical truths that are relevant for color vision – already before her release.

Moreover, assumed a Kimean (1976) characterization of facts/states/processes/... as exemplifications of properties over spans of time, we might define, indeed, rather straightforwardly, physical facts as those facts that involve the instantiation of at least one physical property, and non–physical facts as those facts that do not involve the instantiation of any physical property. Thus, from P – namely the claim that there is the entire physical world and nothing else, hence all the physical properties are instantiated, and no non–physical property is instantiated – clearly follows that all facts, including those about color vision, are physical facts.

So, I take it that the third of the possible reactions to the conceivability–argument listed above – *i.e.*, P∧¬Q is not conceivable and not metaphysically possible – corresponds to the following possible reaction to the thought–experiment on which the knowledge–argument is based, which I presume is the kind of reaction type–A physicalists (see the introduction; see also 1.8.2 and Chalmers, 2003_b) would offer (see, e.g., Dennett, 1991):

> There is nothing new Mary learns upon her release, and all facts about color vision are physical facts.

Let us now have a look at the first one of the possible reactions to the conceivability–argument listed above, namely the one whereby P∧¬Q is conceivable but not metaphysically possible. Readers who have followed me up to this point should be able to foresee the kind of line of reasoning I am about to make explicit. The conceivability of P∧¬Q implies that such a statement could not be *a priori* ruled out by an ideal rational subject as, we may suppose, is Mary. Hence, Mary could not *a priori* infer Q from P. Moreover, the conceivability of P∧¬Q implies that <PHENOMENAL RED> is cognitively independent from every arbitrary one of the physical concepts deployed in P which, by assumption, Mary masters. Hence, according to this possible reaction to the thought–experiment, there is a clear sense whereby Mary will, indeed, learn something new upon her release. In fact, although

the phenomenal thoughts Mary will or may have – *i.e.*, thoughts like 'THIS is what it is like to see something red', 'I wonder whether it is like THIS for other people to see red', 'I like the way it is like for me to see red' *et cetera* – will still concern a physical state, thus a physical property or a number of physical properties – for otherwise P∧¬Q would be metaphysically possible and hence physicalism would be confuted – due to the cognitive independence of the phenomenal concept <PHENOMENAL RED> with every arbitrary one of the physical concepts Mary masters, these thoughts will have an "aura" of novelty and of contingency for her. Thus, what Mary will learn upon her release is not a new fact but rather a new way of thinking about a fact she already knew. Notoriously, this is the response type–B physicalist offer to the argument. Thus, the first of the possible responses to the conceivability–argument listed above coincides with the following possible reaction to the thought–experiment of Mary:

> There is something new Mary learns upon her release, but all the facts about color vision are physical facts.

Later on – 3.2–3.4 – I will offer some further remarks on how this approach has been developed in the recent literature.

For now, let it be sufficient to say that all we have to do now, given what we have managed to establish up to this point, is to reiterate the arguments drawn in 2.10, 2.13 and 2.14 to the case that both phenomenal transparency and phenomenal translucency are incompatible with physicalism, both under the assumption that physical and phenomenal concepts are cognitively independent, and under the assumption that they are not. Let me briefly recap those arguments.

3.1.1 There is nothing new Mary learns upon her release, and all the facts about color vision are physical facts

I won't go into a lot of details, for I have already elaborated extensively in 2.10 and 2.13. It will be sufficient to recall that (1) Assumed that all the facts are physical

facts, both phenomenal transparency and phenomenal translucency violate the NE principle, namely the principle whereby «physical stuff is, in itself, in its fundamental nature, something wholly and utterly non–experiential» (Strawson, 2008_a: 59); (2) in virtue of the fact that on a physicalist worldview only fundamental physical properties are sparse, both phenomenal transparency and phenomenal translucency violate the negative *desideratum* of physicalism, namely the *desideratum* whereby in order for physicalism to be true, there cannot be fundamental experience–involving entities; (3) depending on how the details of one's view are further spelled out, phenomenal transparency or phenomenal translucency may lead to one of the following views, or to a combination of some of them: (i) a form of idealism, whereby the physical is essentially entirely mental; (ii) panpsychist Russelian monism, *i.e.*, a view whereby every fundamental physical entity is at least partially essentially phenomenal; (iii) micro–psychism, *i.e.*, a view according to which some fundamental physical properties are at least partially essentially phenomenal; (iv) property dualism, namely a view on which there are at least two ontologically distinct aspects to the essence of fundamental physical entities; (v) a form of mysterianism whereby no characterization of the universe, not matter how extensive and no matter the concepts it deploys, could ever get us to know the entire nature of phenomenal properties. (vii) a form of pluralism whereby there are several – more than two – ontologically distinct aspects to the nature of phenomenal properties.

3.1.2 There is something new Mary learns upon her release, but all the facts about color vision are physical facts

Although, on this approach it is conceded that physical and phenomenal concepts are cognitively independent, it is still posited that all facts about color vision – all facts in general – are physical facts, for otherwise the positive *desideratum* of physicalism would be violated. In other words, it is stated that although every phenomenal property is ultimately a physical property, we are not and may never be in a position to rationally judge that this is the case *a priori*, that is only on the basis of mastering

the relevant physical and phenomenal concepts and accepting the relevant amount of background physical knowledge as true. In 2.14 I argued that, assumed phenomenal transparency from this follows at least that some (fundamental) physical entities are essentially entirely phenomenal – which leads at least to a form of micro–psychism and possibly either to a panpsychist version of Russelian monism or to a form of idealism – and assumed phenomenal translucency follows that at least some (fundamental) physical entities have at least one essentially–experience–involving component – which leads either to a form of micropsychism or to a form of panpsychist Russelian monism, plus possibly to property dualism, to a form of pluralism or to a form of mysterianism.

3.2 More on the old–facts–new–ways strategy, plus an interlude on Kripke's modal argument against physicalism

I now wish to delve deeper than I have already done into how the classic type–B physicalist "old–facts–new–ways" response to the knowledge argument may be developed. According to such a possible response to the knowledge–argument, as we saw, what Mary learns upon her release is not a new fact, but rather a new way of thinking about an old fact or, in line with a two–factors post–Fregean approach (see 1.5 and 1.7), a new mode of presentation of one and the very same fact. Advocates of such an approach often appeal to the intensionality of knowledge (see Chalmers, 1996: 125–6). In 1.5 and and 1.7 we saw that on a post–Fregean approach an identity statement 'A is B' cognitively significant – *i.e.*, not trivially true – for a subject iff it is not *a priori*, and that if a statement is cognitively significant for a subject, than the concept 'A' and 'B' must have a different intension or sense. In Chalmers' two–dimensional terminology, we may say that in order for the judgment 'A is B' to be cognitively significant and hence not *a priori*, the concepts 'A' and 'B' must have a different primary intension. For recall (see 2.2.3) that on Chalmers'

two–dimensional approach a given statement is *a priori* iff its primary intension is true in every scenario/metaphysically–possible–world in which it is not empty. But since the primary intension of complex statements is obtained by composition from that of their constituents, it is clear that the judgment 'A is B' can be *a priori* if and only if the primary intension of 'A' is identical with that of 'B'.

Advocates of the old–facts–new–ways approach claim that the judgment 'the experience of seeing something red is such–and–such neuro–physical state' is cognitively significant for Mary – which is why she learns something new – hence *a posteriori*. However, they claim that such a statement is still metaphysically necessary. In Chalmers' two–dimensional terms, this implies that the two concepts Mary deploys have a distinct primary intension but the same secondary intension. So, it is posited that the judgment of Mary is a genuine examples of Kripkean necessary *a posteriori* judgments. Lycan (1995) and Tye (1986) claim that the judgment of Mary' 'the experience of seeing something red is such–and–such neuro–physical state' is akin to the 'water is H2O' one; Churchland (1985) appeals to the gap between knowledge of temperature and knowledge of mean kinetic energy (Chalmers, 1996: 125). Horgan (1984) appeals to the difference between the knowledge of Clark Kent and the knowledge of Superman; McMullen (1985) appeals to the difference between the knowledge of Mark Twain and the knowledge of Samuel Clemens (see Chalmers, *Ibid.*)

Notoriously, Kripke held that necessary *a posteriori* judgments can be true only if at least one of the two concepts deployed provides a contingent characterization of their referents. Indeed, an early type–B physicalists such as Loar (1990/1997; 2002) argues that typical anti–physicalist arguments such as the knowledge argument and Kripke's (1980) modal argument rest upon two assumptions, namely the cognitive independence of physical and phenomenal concepts and what he (2002: 297) calls *the semantic premise*[2](see also Holman, 2013: 809):

[2]While Loar, in line with mainstream type–B physicalism, accepts the cognitive independence of physical and phenomenal concepts, he rejects the semantic premise. What is more, he (2002: 299) writes that «there is no good philosophical reason to deny that, odd though it may sound, the properties these [recognitional–demonstrative concepts] phenomenologically reveal are physical–functional

The semantic premise

A statement of property identity that links conceptually independent concepts is true only if at least one concept picks out the property it refers to by connoting a contingent property of that property.

As far as I can see, Loar's semantic premise is equivalent to the third premise of the argument from revelation, namely the principle of cognitive inter–dependence of transparent co–referential concepts. Let me explain. It seems to me that put schematically, Kripke's modal argument against physicalism may look as something like this. Readers should be able to notice even at a first glance that the similarities between this argument and the argument from revelation presented in 1.8 are striking[3]

[1] *The semantic premise*: A statement of property identity that links cognitive independent concepts is true only if at least one concept picks out the property it refers to by connoting a contingent property of that property.

[2]Phenomenal and physical concepts are cognitive independent.

[3]Phenomenal concepts connote a necessary property of their referents.

[4] In order for (identity) physicalism to be true – that is, in order for experiences to be nothing over and above a kind of physical states – physical concepts must connote a necessary property of their referents.

[5]Either phenomenal concepts connote a contingent property of their referents, or (identity) physicalism is false [in virtue of 1, 2 and 4]

properties—but not of course under physical–functional descriptions», and he acknowledges that while anti–physicalist will urge that if two concepts reveal the nature of the same property then a subject should be able to see *a priori* that the they co–refer, this in fact does not follow. All this might suggest that Loar endorses a version of the thesis of dual carving (on which I will elaborate in 3.4). Dual carving states that the nature of one and the very same entity/property/... may be revealed in two or more conceptually independent ways. However, it not clear whether Loar would subscribe to the thesis of phenomenal transparency. In fact, for him phenomenal concepts do not attach any mode of presentation at all, let alone any *essential* mode of presentation, to their referents, phenomenal properties. I will address Loar's approach momentarily.

[3]Truth be told, Kripke (1980) does not present the matter in terms of phenomenal concepts. It seems very clear to me, however, that it has phenomenal concepts, or at least something very close to them, in mind. Here, for instance, is what he (1980: 152, emphasis mine) writes: «Pain, on the other hand, is not picked out by one of its accidental properties; rather it is picked out *by the property of being pain itself, by its immediate phenomenological quality.*»

Conclusion: (Identity) physicalism is false [in virtue of 3 and 5].

In 1.4 I presented a distinction between a modal and a non–modal/Finean account of essentiality. We saw that according to a modal account of essentiality a given entity as a property P essentially iff it has it necessarily, namely in every possible world, or at least in all the possible worlds in which it exists. Now, Kripke notoriously endorsed a modal account of essentiality. So, to say that phenomenal concepts connote at least one necessary property of their referents just is, on his view, to say that they connote an essential property of their referents. But to deny that phenomenal concepts connote at least one essential property of their referents, as physicalists are committed to do to rebut the argument just presented, just is to deny both phenomenal transparency and phenomenal translucency, for phenomenal translucency claim phenomenal concepts provide cognitive access to part of the essence of their referents *a priori*, whereas phenomenal transparency claims that phenomenal concepts allow to know *a priori* the full nature of their referents. Even if one assumed a non–modal/Finean account of essentiality, as far as I can see my point would hold. As we saw in 1.4, it is widely agreed that if a given property P is essential to an entity in the Finean sense, it must also be essential to that entity it the Kripkean/modal sense, namely possessed by the entity necessarily, whereas whether the converse relation of implication holds is controversial, for it is not clear whether every property that is possessed necessarily by an entity is essential in the Finean sense to that entity. From the fact that if a property is essential in the Finean sense to an entity it must be possessed necessarily by that entity follows, by *modus tollens*, that if a given property is possessed contingently by a certain entity that it cannot be essential to it in the Finean sense, which again implies that denying that phenomenal concepts connote at least one necessary property of their referents implies denying both phenomenal transparency and phenomenal translucency.

Moreover, *contra* what Loar suggests, we saw throughout the past chapter that even if both the semantic premise and the cognitive independence of physical and phenomenal concepts happened to be denied physicalists would still be committed to

deny both phenomenal transparency and phenomenal translucency, for both phenomenal transparency and phenomenal translucency are inconsistent with physicalism no matter whether physical and phenomenal concepts are taken to be cognitively independent or not.

So, again, the knowledge argument – as well as the conceivability argument and, as far as I can see, Kripke's modal argument – rest upon phenomenal transparency or at least phenomenal translucency as one of their key–premises, if not *the* key–premise, meaning that it is necessary and most likely necessary and sufficient that both full–revelation/phenomenal–transparency and partial–revelation/phenomenal–translucency are denied in order for those arguments to be rebutted.

3.3 Phenomenal concepts as indexicals and/or as recognitional/demonstratives

As we just saw, type–B physicalist theories claim that although for every phenomenal concept there is at least a physical one that designate/refer–to the very same property – meaning that every phenomenal concept is necessarily coextensive with at least a physical one – physical and phenomenal concepts nonetheless *indicate*, *introduce* or *connote* distinct properties, meaning that they characterize their referents under different aspects or modes of presentations. We also saw that if two concepts C_A and C_B characterize the same referent differently, then the judgment 'A is B' is cognitively significant – *i.e.* not trivially true or *a posteriori*. So, if a judgment 'A is B' is cognitively insignificant or *a priori* then the two concept deployed, C_A and C_B, must have the same mode of presentation of their referent – in Chalmers' two–dimensional terms, the same primary intension.

Brian Loar (1990/1997; see also his 2002) says something different. He argues that two concept may introduce the same property without a subject being able to see this *a priori*. In light of what we have seen up to this point, it is clear that the only

way for Loar to say something like this is by arguing that one of the two concepts – in the case the concerns us, this would be a phenomenal concept – do not characterize its referent in any way – *i.e.*, does not have any primary intension, or do not present their referent under any meaningful mode of presentation at all. As a matter of fact, this is precisely what Loar states: he argues that phenomenal concepts function as recognitional/demonstratives that allow to recognize or re–identify given kind of experiences *directly*, namely without attaching any meaningful mode of presentation, let alone any *essential* mode of presentation, to them. But a concept which reveal no property of its referents at all is a radically opaque concept. And it is clear if phenomenal concepts are taken to be radically opaque both phenomenal transparency and phenomenal translucency are *eo ipso* denied.

As far as I can see, very similar remarks may be drawn as for what concerns other recognitional/demonstratives and/or indexical accounts of phenomenal concepts (Tye 2000, chap. 2; 2003; Carruthers 2000, 2004; Perry 2001; Levin 2007a,b; 2019; Tye 1995, chap. 6; Lycan 1996, sect. 3.3; Bigelow and Pargetter 1990; McMullen 1985; Papineau 1993) and, possibly, for so called quotational models, according to which phenomenal concepts deploy the very conscious state that one is thinking about without, however, and again, attaching any meaningful mode of presentation to them (Papineau 1993, chap. 4; 2002, chap. 4; 2007; Balog 1999, 2011; 2012; Melnyk 2002; Block 2007)[4].

3.4 On dual carving

I will now examine a sub–version of the Type–B physicalist response to the knowledge argument as well as, possibly, to Kripke's modal argument. On this sub–version of type–B physicalism physical and phenomenal concepts are cognitive independent and both are transparent. I take it that physicalist sponsors of such a view are

[4]This is what Papineau (2002: 85–6) writes: «Materialists should refuse Kripke's invitation. They should say that phenomenal concepts refer directly, and not by description. Phenomenal concepts don't pick out their referents by invoking certain further features of those referents, but in their own right, so to speak». This is just one example among a plethora of others I might have made.

still committed to claim that every phenomenal concept is ultimately necessarily coextensive with at least a physical one even though, in virtue of the cognitive independence of physical and phenomenal concepts, we may never rationally judge that this is the case only on the basis of mastering those concepts and accepting the relevant background physical knowledge as true, say due to some peculiar features of phenomenal concepts, or to some intrinsic cognitive limitations of ours. I've already elaborated extensively on this point, but justo to recap, suppose for the sake of argument there was a phenomenal concept C_{phen} that was not necessarily coextensive with at least a physical one. From this would follow that there may be a possible world in which the referent of every arbitrary physical concept is instantiated without the referent of C_{phen} being instantiated, for otherwise C_{phen} would turn out to be necessarily coextensive with at least one physical concept, *contra* what we are assuming for the sake of argument. But a world in which the referent of every physical concept is instantiated, is an exact physical duplicate of the actual world. So, denying that every arbitrary phenomenal concept is necessarily coextensive with at least a physical concept (no matter whether subject are in a position to rationally judge that the latter is the case on the basis of mastering the relevant concepts or not), and so conceding that there may be a phenomenal concept that is not necessarily coextensive with any physical concept, implies that there may be an exact physical duplicate of our world in which at least one phenomenal property is not instantiated, *i.e.*., an exact physical duplicate of our world which is not a duplicate of our world with respect to phenomenal properties – a C–duplicate of our world. The latter violates the positive *desideratum* of physicalism.

On the view we are considering, thus, every phenomenal concept is necessarily coextensive with at least a physical concept, though, again, we may not be in a position to rationally judge that the latter is the case only on the basis of mastering the relevant concepts and accepting the relevant background physical knowledge as true. Moreover, it is assumed that both physical and phenomenal concepts are transparent. So, we have the the nature of one and the same phenomenal property

is fully revealed by at least two cognitively independent concepts.

The view whereby the nature of one and the same entity may be revealed by two or more cognitively independent concepts has been called *dual carving*. (see Goff 2011, 200–201; 2017: 125–32; Díaz–León 2013; 2016: 1193; Taylor, 2013; Damjanovic 2012, 76–77; Elpidorou, 2015; Levin, 2019, and possibly Loar, 1990/1997; 2002). Dual carving violates the third premise of the argument from revelation, namely the principle of cognitive inter–dependence of transparent co–referential concepts (see 1.8.3), which states that no two concepts can be cognitively independent and both transparent. I will now attempt to show that none of the purported counterexamples that to my knowledge have been offered against the principle of cognitive inter–dependence of transparent co–referential concepts uncontroversially works. I acknowledge that showing that none of the purported counterexamples that have been offered against a certain thesis up to a certain point uncontroversially works does not suffice, *per se*, to show that the thesis is true. However, I think it is fair to say it gives *prima facie* reasons to believe that the thesis, or at least something close to it, holds. At any rate, let me stress that in virtue of the arguments drawn in 2.10, 2.13 and 2.14 and summarized yet again in 3.1 the view we are now considering is most likely inconsistent with physicalism anyway.

Let us start with the case for dual carving as made by Díaz–León (2013; 2016). As a matter of fact, Díaz–León does not put the matter neither in terms of inferential isolation nor in terms of cognitive independence. Instead, she (2013: 7) claims that «we can make sense of the idea that different subjects can know what it takes for a property to be instantiated in numerous conceptually distinct ways». This already marks a point of distinction with respect to the principle of cognitive inter–dependence of transparent co–referential concepts as it is characterized in the third premise of the argument from revelation. In fact, as I have already had the occasion to underline – and as Nida–Rümelin (2007: 327) also acknowledges – such principle does not state that there is just one conceptual way to understand what it is for a given entity to be part of reality, so that if two subjects understand what it is for that entity to be part

of reality, they must have the same concept: indeed, it allows for the possibility of the nature of a property being graspable via two distinct concepts. Rather, it states that if a subject understands what it is for a given entity to be part of reality in two or more conceptually distinct ways, then they must on this basis be in an epistemic position to rationally judge that the concepts deployed are necessarily coextensive, provided they accept the relevant amount of background knowledge K as true. That is, the principle states that two concepts cannot be both transparent with respect to the same entity and cognitive independent.

On this respect – again, as I have already had the occasion to underline in 1.8.3 – I disagree with Goff. Goff (e.g., 2011: 198) labels the claim whereby the essence of a given entity or set of entities may be revealed in two or more conceptually distinct ways as *the thesis of dubious intelligibility* (TDI). I believe that, at least in certain cases, we can, indeed, make sense of TDI. Thus I suppose that Díaz–León (2013; 2016. See also Elpidorou, 2015) is on the right track. Yet TDI does not contradicts the principle of cognitive inter–dependence of transparent co–referential concepts. Indeed, it seems to me that the only way to make TDI intelligible is to combine it with that principle, and I could not find any uncontroversial counterexample to this.

At any rate, Díaz–León does not elaborate much on what it is for two pieces of knowledge to be conceptually distinct. Rather, she (2013: 7) offers the following examples:

[A]For x to be a bachelor is for x to be an unmarried man.

[B]For x to be a bachelor is for x to be an unmarried male Homo Sapiens.

[C]For x to be a bachelor is for x to be an unmarried Homo Sapiens with XY chromosomes.

[D]For x to be a bachelor is for x to be an unmarried male member of such and such clade (insert here a correct description of the clade that corresponds to the species Homo Sapiens, according to the cladistic account of species).

Then she (*Ibid.*) goes on to argue that these propositions are «conceptually distinct, in the familiar sense that a rational subject could believe one but fail to believe the others». Assumed that to know what it is for a given entity to be part of reality is to have a certain amount of conceptual/propositional knowledge, it follows that one could know what it is for a certain entity to be part of reality in two or more conceptually distinct ways.

Now, of course there is an obvious sense whereby the concepts <MALE>, <HOMO SAPIENS> , <X CHROMOSOME>, <Y CHROMOSOME>, < INSERT HERE THE CLADE THAT CORRESPONDS TO THE SPECIES HOMO SAPIENS> *et cetera* are distinct: the experiences needed to coming to form those concepts are distinct. One could have studied genomics but not cladistics and therefore master the concept of chromosome but not the one of a clade. Thus, there is no denying that one could, say, believe [C] but fail to believe [D], provided they master certain concepts but not others. Damnjanovic's (2012: 77) remarks somehow go in the same direction. As he writes:

> [...] it is implausible to read Revelation as entailing knowledge of all essential truths about a phenomenal property: on that exacting standard no–one could ever know the nature of anything without possessing all possible concepts for that thing. It is possible, therefore, that Thomas knows the full nature of the taste of peaches, and yet does not know certain essential truths about it that involve physical concepts, simply because he does not possess those concepts.

What I find way harder to understand, however, is how could it be possible for someone to master the concepts <MALE>, <HOMO SAPIENS>, <Y CHROMOSOME> and <X CHROMOSOME>, accepting the relevant background knowledge K as true – say, know the relevant amount of biology, chemistry, and so on – and yet given an arbitrary possible world believe that X is a male homo sapiens in that world while failing to believe that X is a homo sapiens with XY Chromosomes in that world, or at least while not being, just in virtue of mastering those concepts

and accepting K as true in our world, in an epistemic position to agree to the latter belief but not to the former one or *vice versa*.

This follows from the fact that <MAN>, <MALE>, <HOMO SAPIENS>, <Y CHROMOSOME> and <X CHROMOSOME> are all Kripkean rigid designators. Given the definition of a rigid designator, namely a concept whose reference is fixed in every possible world in which it exists, if two rigid designators are actually coextensive, as <MAN> and <HOMO SAPIENS> are, then they must be necessarily coextensive.

Let me go a bit more slowly and reiterate the point in a slightly different way. Consider the statement 'water is H_2O'[5]. Indeed, it took time and effort to discover the molecular structure of water. Yet, whoever now knows even a minimal amount of chemistry and masters the concepts <WATER> and <H_2O> is on that basis in an epistemic position to rationally judge that the water–H_2O identity statement is necessarily true, namely, that if there is <H_2O> in a given possible world, then *eo ipso* there is water in that world and *vice versa*. Thus, the concepts <WATER> and <H_2O>, even though clearly distinct, are *not* cognitively independent (relatively to the relevant amount of background chemical knowledge).

One may object that it still conceivable that for some reason, say because there

[5]One may question whether the water example is appropriate. In fact, it might be said, it is not clear whether the concept <WATER> is translucent or whether, instead, it is mildly opaque – *i.e.*, only reveal some features (e.g., being thirst quenching, filling the oceans, *etc.*) which, although do not belong essentially to water/H_2O, are sufficient to fix the referent of the concept in the actual world. However, <WATER> is a rigid designator. In the present context, the example is only meant to show that in most cases in which two concepts are both rigid designators and coextensive, they are not cognitive independent. Elpidorou (2015) raises a similar point with respect to his <JELLYFISH> example (see below). The same remarks apply to his case. Later I will consider the case for dual carving as made in Levin (2019) and Damjnanovic (2012). I will concede that there *might* be peculiar cases, namely those involving demonstrative concepts and proper names, which *may* violate the principle whereby if two concepts are both rigid designators and coextensive they cannot be cognitive independent. Yet, as I will argue and as I have already suggested in 3.3 it is not at all clear nor uncontroversial whether demonstratives and proper names can reveal anything at all, let alone any essential property, of their referents, thus being either translucent or transparent, or whether, instead, they merely function as "pinpointers" that designate certain entities directly without attaching any (essential) mode of presentation to them, thus being mildly opaque at best, if not even radically opaque. In light of this, the examples brought by Levin and Damnjanovic still do not confute the principe of cognitive inter–dependence of *transparent* co–referential concepts, or at least it not uncontroversial that they do so. What is more, the arguments provided in 2.10, 2.13 and 2.15 and summarized in 3.1 are neutral with respect to the the principle of cognitive inter–dependence of transparent co–referential concepts. Thus, appealing to these or other purported counterexamples to such principle

is an "hidden element" which cannot be detected by the instruments currently at our disposal, chemists have been deceived all this time, and water is *not*, actually, H_2O but rather, say, H_2OQ. So, the objection would go, there is at least a sense whereby it is conceivable that water might not be H_2O and, in this sense, none of us is actually in a position to exclude that <WATER> and <H_2O> are not necessarily coextensive. To counter this objection it is sufficient to assume an ideal rational agent and take her ideally complete chemical knowledge as the relevant background knowledge K. One could be skeptical as to wether such a kind of knowledge, so complete as to make it inconceivable that it be replaced or confuted by new empirical findings, may be ever achieved. Recall, however, that the second premise of the argument from revelation states that given a phenomenal concept C_{phen} there is no arbitrary physical concept C_{phis} such that a subject understanding C_{phen} and C_{phis} and having *arbitrary* physical knowledge would be in a position rationally judge that they are necessarily coextensive. By the way, as I have already underlined, this is something that all type–B physicalists would accept, for the concede that physical and phenomenal concepts are inferentially isolated, meaning that a phenomenal characterization of certain properties/states/events/... could not be inferred from a physical one *however extensive* (see Sundström, 2011), and inferential isolation clearly implies cognitive independence. Therefore, I do not think the objection here under consideration poses any serious challenge to the picture I am trying to sketch[6].

Elpidorou (2015) raises objections to the argument from revelation, which he labels the 'introspection objection' to type–B physicalism, that are close in spirit to the ones raised by Díaz–León. First, he argues that a given concept may reveal that its referent(s) has/have certain properties without revealing, in turn, that those

[6]Holman (2013) critique to type–B physicalism is close in spirit the kind of critique I am pursuing here. According to Holman, type–B physicalists need to provide an explanation of why physico–phenomenal identity statements, unlike other identity statements involving rigid designators such as 'water is H_2O', although necessarily true will maintain an appearance of contingency no matter how far scientific research progresses. Typically, at this point type–B physicalists appeal to some kind of peculiar if not even unique feature of phenomenal concepts. However, the arguments I have drawn –2,10, 2.13, 2.14 and 3.1 – to the case that both phenomenal transparency and phenomenal translucency are inconsistent with physicalism are neutral as to whether phenomenal concepts are inferentially isolated, thus cognitive independent, from physical ones or not. Thus, this type of strategy on the part of the physicalist won't work.

properties are physical. As an example, he mentions the concept <JELLYFISH>. To apply <JELLYFISH> to a given entity is to think about that entity as having a number of features like having tentacles, having a soft body, *et cetera*. These are physical features. More in general, the property 'being a jellyfish' is a physical one. Yet, mastering the concept <JELLYFISH> *per se* does not provide a subject with any new piece of physical knowledge. That is, although it does reveal that its referents have certain physical features, it does not reveal those features as physical.

However, again as far as I can see the jellyfish case does not confute the principle of cognitive inter–dependence of transparent co–referential concepts. <JELLYFISH>, as all the concepts that designate species and natural kinds such as <TIGER>, <WATER> or <GOLD>, is a rigid designator (see Kripke, 1980). Presumably, it refers to 'all and only the entities with such–and–such genome' or something close, where the concept <INSERT HERE THE GENOME OF JELLYFISHES> is also a rigid designator. And now the story goes as already told: whoever knew the relevant amount of genomics would be in a position to rationally judge that iff there's an entity with such–and–such genome in a given possible world, then there's a jellyfish in that world and *vice versa*.

Elpidorou (*ivi:* 15) also mentions another putative counterexample to the principle of cognitive inter–dependence of transparent co–referential concepts, which relies on the distinction between senses of 'physical' we already saw in 2.10. Recall that an entity or property is said physical in the restricted sense if it is posited by our best physical theories, whereas it is said to be physical in the broad sense if it is realized by, necessitated by, or metaphysically supervenient upon some entities that are physical in the restricted sense.

The fact now is, according to Elpidorou, that phenomenal concepts may not reveal that their referents have certain physical aspects in the restricted sense, but still reveal that they have some physical aspects in the broad sense, and thus are realized by, necessitated by, or metaphysically supervenient upon some physical properties, states or entities in the restricted sense. To strengthen his case, Elpidorou (*ivi:* 17)

suggests that introspection might not be different, on this regard, from perception: Perception does not directly provide us with physical objects in the broad sense sense; we cannot visually perceive particles like quarks or leptons. However, according to Elpidorou, and I agree with him on this, it would be hasty to assume from this that our perception doesn't reveal physical aspects. For instance, I can sense the solidity of the table before me without simultaneously perceiving its micro–physical properties.

However, as I already suggested in 2.10 – and as Strawson (2008_a: 24) makes very vivid – there is a clear sense whereby broad physical concepts are cognitive inter–dependent from narrow ones. Hence, as far as I can see one may very well accept Elpidorou's distinction between senses of 'physical' and still accept the principle of cognitive inter–dependence of transparent co–referential concepts.

I now whish to briefly address Levin's (2019) defense of the possibility of dual carving. Following Loar (1990/1997) Levin regards phenomenal concepts as recognitional–demonstratives. That is, according to Levin these concepts detect properties of our experiences from an introspective standpoint and allow one to recognize, re–identify or classify certain experiences. As is the case for Loar, also for Levin recognitional–demonstrative phenomenal concepts refer rigidly and directly, that is, without the mediation of any (contingent) reference–fixing mode of presentation. When, for instance, I instantiate a given phenomenal property, say the property of being visually presented with red, and I recognize it as "again one of those" or "the property that is shared by all the visual experiences of red", the (presumably) physical property that is causally responsible for such an act of recognition or reidentification would constitute the denotation of a phenomenal concept as construed under such an account.

Now first of all, as I have already underlined in 3.3 it is not at all clear whether there can be any transparent/translucent recognitional–demonstrative concept in general or whether, instead, the latter kind of concepts can only be mildly opaque if not even radically opaque by their very nature, in virtue of the peculiar way in

which they refer.

Levin (2019: 64) believes that there may be nature–revealing recognitional–demonstrative concepts. For instance, she mentions a demonstrative concept of sphericity that denotes its referents as 'that shape'[7], pointing at spheres under normal conditions. Yet she acknowledges that this is indeed a matter of dispute and not uncontroversial. However, she (*ivi:* fn. 18) dismisses the worry by stating that «this may be yet another way that phenomenal recognitional–demonstratives are unique».

Likewise, although she concedes that in most cases in which two concepts reveal the essence of the same property they are not cognitively independent, she (*ivi:* 65) argues that to require that any time two concepts reveal the nature of the same property they must be cognitively inter–dependent, as the third premise of the argument from revelation does, would beg the question against type–B physicalism.

Damnjanovic (2012: 77–8) offers the <CASSIUS CLAY> – <MOHAMED ALI> example as a case for dual carving. He says that one may fully master the concept <CASSIUS CLAY>, know anything there is to know about the nature of Cassius Clay, and then coming to form the concept <MOHAMED ALI> in a way that does know allow them to infer *a priori* that the two concepts are necessarily coextensive. Recall, however, that a concept is said to be transparent iff «it is *a priori* (for someone possessing the concept and *in virtue of possessing the concept*) what it is for that entity to be part of reality» (Goff, 2017: 15, emphasis mine). And now I seriously doubt that *just in virtue of possessing the concept <CASSIUS CLAY>* one would get to know anything about the nature of Cassius Clay himself. In general, again, whether proper names have any meaning at all, rather than merely designating their referents without the mediation of any meaningful mode of presentation, has been largely put into question (see Cumming, 2019).

As for Damnjanovic's (*ivi:* 78) <WATER>–<RETAW> case, it goes as follows.

[7]It is not entirely clear whether the concept <THAT SHAPE> really is such that it does not provide *any meaningful mode of presentation at all* of its referent. For after all, such concept characterizes its referent *as a shape*. Hence, one could say, it *does* present its referent under a certain aspect, although very minimal.

He says that a child might start to call water 'retaw' as a game. One, then, could come to learn the child's concept <RETAW> without knowing that it is necessarily coextensive with <WATER>. As far as I can see there are two ways in which this could turn out. If one got to know that <RETAW> is just a child's way to name water, then clearly by accepting this as the relevant background knowledge K as true in our world they would be in a position to rationally judge that <WATER> and <RETAW> are necessarily coextensive, in accordance with the principle whereby, at least in most paradigm cases, if two concepts are both rigid designators and coextensive, they are not cognitively independent, at least relatively to the relevant amount of background knowledge K. But even if one did not know this, what <RETAW> reveals of its referent, at best, is that it is a colorless, liquid stuff which a child loves, *et cetera*, none of which is an essential property of the concept's referent itself. Thus, the principle of cognitive inter–dependence of *transparent* co–referential concepts is arguably still safe.

Anyway, Damjanovic (*ivi:* 79) merely dismisses these worries by stating that they beg the question against type–B physicalism, just as Levin does.

I thereby conclude that none of the examples I have mentioned thus far uncontroversially confutes the third premise of the argument from revelation, namely the principle of cognitive inter–dependence of transparent co–referential concepts: one could accept most of what Díaz–León and Elpidorou say and still accept the argument, thereby rejecting physicalism.

As I have already stressed, I acknowledge that this does adjudicate the match *per se*. To show that none of the counterexamples that have been offered against a thesis uncontroversially works is not to show that the thesis is true. Yet I think it is fair to say that there are good *prima facie* reasons to believe the principle of cognitive inter–dependence of transparent co–referential concepts or something close to it holds, and that the burden of offering an uncontroversial counterexample to it is on physicalists.

As for Levin and Damjanovic, absent independent reasons besides one's need to

save physicalism to believe that phenomenal concepts have such a unique nature, working differently from any other kind of concepts, their defence of dual carving strategy smells *ad hoc*, at least to me. As Tomasetta (2015: 107) writes,

> That physicalism is indeed more a worldview than a well–grounded philosophical thesis is further buttressed by the almost religious fervency with which materialist views are often held (Bonjour 2010, p. 4). A fervour that is evident, for example, in Dennett's (1989, p. 37) declaration that "dualism is to be avoided at all costs", a position which is certainly not well suited to a rational inquirer.

That being the case, once again I think it is fair to say that the burden of offering an uncontroversial counterexample to the third premise of the argument from revelation which does not involve phenomenal concepts themselves is on physicalists. Yet I want to concede that there *might* possibly be counterexamples to the principle of cognitive inter–dependence of transparent co–referential concepts, though not uncontroversial, which do not involve phenomenal concepts.

However, the argument I have drawn to the case that both phenomenal transparency and phenomenal translucency are inconsistent with physicalism, both under the assumption that physical and phenomenal concepts are cognitively independent and under the assumption that they are not are neutral with respect to the principle of cognitive inter–dependence of transparent co–referential concepts. Hence, as far as I can see, appealing to dual carving won't save physicalism anyway.

Below – 3.6 and 3.7 – I am going to offer some remarks on another few of the most renowned physicalist responses to the knowledge argument that have been offered in the recent (and not so recent) literature in the attempt to show that it is at least necessary and possibly necessary and sufficient, in order for each of them to go through, that both phenomenal transparency and phenomenal translucency are denied. Before we go ahead, there is one last case for dual carving I whish to examine.

3.5 The Powerful Qualities view: Mørch (2018) on why dispositionalism may entail panpsychism.

The last case for dual carving I whish to examine is the one put forth in Taylor (2013). Taylor appeals to the so called powerful qualities view to defend the possibility of dual carving. Here I will elaborate a bit on some arguments that have recently been developed by Hedda Hasel Mørch (2018) to the case that the only essentially dispositional qualities that we know may be phenomenal properties themselves. *Prima facie*, one might think that the arguments of Mørch may be deployed to cast doubt on whether the powerful qualities view really counts as a genuine counterexample to the principle of cognitive inter–dependence of transparent co–referential concepts. However, I will suggest that under certain assumptions the powerful qualities view *might*, indeed, be deemed as a genuine case of dual carving. Still, when combined with either full–revelation/phenomenal–transparency or partial–revelation/phenomenal–translucency the powerful qualities view still leads to a panpsychist view. Hence, defending physicalism via the powerful qualities view still commits one to deny both full–revelation/phenomenal–transparency and partial–revelation/phenomenal–translucency.

The powerful qualities view is mostly associated with the work of Charles Burton Martin and John Heil (see Martin, 2007; Martin and Heil 1998; 1999; Heil, 2003, ch. 11. See also Strawson 2008_{b}). As we saw in 1.3, we may understand dispositional properties as those properties that inherently or essentially ground certain dispositional powers of their bearers. That is to say, the dispositionality of a property is the influence the property has on the behaviors exhibited by an object that possesses it, based on specific stimuli (see Taylor, 2013).

As for qualitative properties, Taylor (*Ibid.*) defines them as those properties we need not identify via the powers they confer to their bearers. Notice, crucially, the the latter definition has merely to do with our way of referring to some given properties, of thinking about them or characterizing them. That is to say: from

the fact that we need not identify some given properties via the powers they confer on their bearers does not follow that the nature of those properties is such that they inherently ground some patterns of behaviour of their bearers based on certain stimuli. In fact, this is what Martin and Heil (1999: 46) say:

> [...] every property is at once dispositional and categorical—or, as we prefer, dispositional and qualitative. Dispositionality and qualitativity are built into each property; indeed, they are the property. [...] A property [...] is a two–faced dispositional–qualitative coin. Properties are not compounds, however. [...] A property just is a certain dispositionality that just is a certain qualitativity. What we propose boils down to a surprising identity: the dispositional and the qualitative are identical with one another and with the unitary intrinsic property itself. The suggested identity is surprising only because we have grown used to distinguishing the dispositional and the qualitative. Once it is recognized that these are simply different ways of representing the selfsame property, the identity can be seen to resemble countless others. For any intrinsic, irreducible property, then, what is dispositional and what is qualitative are one and the same property differently considered: considered as what the property exhibits of its nature, and considered as what the property is directive and selective for as its manifestations.

Thus, under the powerful qualities view qualitativity and dispositionality are understood as conceptually distinct ways of considering what it is for one and the very same property to be instantiated. Notice that to bring his point home Taylor (*Ibid.*) does not need to prove the powerful qualities view is true: it is enough for him to show it is coherent and intelligible, for provided it is possible to make sense of the idea whereby one may understand what it is for a given entity to be part of reality in two or more conceptually distinct ways, the physicalist "dual carver" is in a position to argue that physical and phenomenal concepts function just as concepts of qualities and concepts of disposition under the powerful qualities view.

What is more, *prima facie* the powerful qualities view may also be seen as allowing physicalists to avoid the threat posed by two arguments for panpsychist Russelian monism – in the following just panpsychism for brevity. Let me explain[8]. Assume first, a dispositionalist conception of theoretical physics. Combined with a physicalist worldview whereby only fundamental physical properties carve nature at its joints, such a dispositionalist conception implies that all fundamental entities are essentially dispositional – unless physical concepts are taken to be either mildly opaque or radically opaque, which, as we saw in 1.8.4, can hardly be made compatible with physicalism. Panpsychists, then, urge that dispositions require qualitative grounds as their realizers. Moreover, they argue that phenomenal properties are or at least may be the only qualities we know, and suggest that phenomenal properties may be the qualitative grounds of physical dispositions. This idea is then deployed to draw two distinct but structurally similar arguments to the case that panpsychism should be preferred over possible alternative views (see Mørch, 2018: 1–2).

The first argument (see, e.g., Strawson, 2008_a; Alter and Nagasawa, 2012; Chalmers, 2013; Mørch, *Ibid.*) claims, roughly, that panpsychism has all the advantages of physicalism while at the same time avoiding its disadvantages and avoiding the problems of dualism as well. It keeps the advantages of physicalism and avoids the issues of dualism in that, as a form of monism, it is compatible with the causal closure of the physical. However, and unlike standard physicalism, by positing phenomenal properties as fundamental it promises to close the physico–phenomenal explanatory gap.

The second argument for panpsychism claims that, since phenomenal properties are (arguably) the only qualities we know, positing them as fundamental allows to avoid a (Kantian) form of nominalism whereby the qualitative/categorical grounds of the physical world are unknowable or even inconceivable (Mørch, *Ibid.*; Strawson, 2006; Seager, 2006; Alter and Nagasawa, 2012).

[8]In the following, I will mostly follow Mørch's (2018) exposition which, however, is far more rich and articulated than my brief sketch could be. Hence, I advice interested readers to go and have a look at Mørch's paper.

The fact now is that physicalists may appeal to the powerful qualities view to offer a response to these two arguments. In fact, they may accept a dispositionalist conception of theoretical physics but deny that dispositions require qualitative grounds over and above themselves, claiming that dispositions *just are* qualities, and hence are to be taken as fundamental and irreducible properties in themselves. Of course, in order to work such a strategy needs to assume that no fundamental physical quality/disposition is essentially phenomenal. The view whereby disposition do not require qualitative grounds over and above themselves and are to be taken as fundamental and irreducible in themselves is called by Morch (2018) *dispositionalism*[9].

One may say that dispositions are purely structural properties, and then appeal to arguments against so called ontic structural realism (Ladyman and Ross, 2007) to undermine dispositionalism (see Mørch, 2018: 4). Roughly, ontic structural realism claims that physical properties are purely structural or relational, but denies that purely relational properties require essentially non relational *relata*. Thus, on this view, physical properties are treated like points in a graph – that are individuated only in virtue of their position with respects to other points – or as numbers – that are individuated only in virtue of their relations to other numbers. According to many (e.g., Van Frassen, 2006, see also Mørch, 2018: 4) such a view leads to a *reductio ad absurdum*, for it implies that the (instantiated) physical is virtually indistinguishable from the (non–instantiated) mathematical.

The fact is, however, that dispositionalists and/or sponsors of the powerful qualities view need not say that dispositions are purely structural properties. As Mørch (2018: 5) writes:

> [...] powers have a relational aspect, in the sense that they are essentially directed or aimed toward producing their manifestations. [...] But powers can also be thought of as characterized by a kind of primitive force or energy (or "oomph", as some have called it) that we may seem to have an intuitive, positive grasp of. This aspect is not purely structural

[9]Dispositionalism is defended – though not always via the powerful qualities view – by Shoemaker (1980), Ellis (2002), Molnar (2003), Mumford (2004) and Bird (2007), among others.

> or relational, at least not in the sense of being capturable in purely logico–mathematical terms. The property of instantiating causal power has also traditionally been regarded as a paradigmatic criterion of real, concrete existence (for example by Plato's Eleatic Stranger).

Another option opponents of dispositionalism may resort to is to appeal to an objection raised by Armstrong (1997) called the "always packing never traveling" objection (see also Goff, 2017: 140; Mørch, 2018: 5). According to such an objection, a world in which only powers or dispositions are instantiated , as per a dispositionalist view, would be never actually realized, but it would exist only in potency. However, dispositionalists may retort that the always packing never traveling objection conflates potencies, or powers, with potentials (see Mørch, 2018: 5; Mumford, 2004: 174). In fact, while the latter are not actual by definition, the former may very well be actual: a crystal glass is actually fragile even no sharp object is hitting it and it is still intact.

Finally, anti–dispositionalists may contend that dispositions require intrinsic grounds, where, recall, in 1.3 intrinsicality has been defined, in line with a very standard approach, as such that if a an intrinsic property P is instantiated, than the fact that X is P does not hold partly in virtue of any other entity besides X itself and/or some of its parts. Russell (1927: 325) argued that intrinsic properties are necessary to avoid a vicious infinite regress. Seager (2006: 142) has argued that concreteness requires intrinsicality.

However, once again dispositionalists may reply that dispositions understood as powers, that is, as potencies rather than as potentials, may very well be taken as intrinsic: the fragility of a crystal glass does not seem to hold in virtue of any other entity besides the glass itself; it would hold even if no sharp object with the power of breaking the glass existed in the universe. Likewise, a given charged particle may instantiate the power of attracting or repelling other possible charged particles even if no other charged particle existed in the universe, and hence such power may never be manifest (Mørch, 2018: 8).

I will now briefly expose some arguments that are developed by Mørch (2018) to the case that the at least some phenomenal properties may be essentially dispositional and that, possibly, the only essentially dispositional properties we know are phenomenal properties themselves.

Mørch (2018: 9) starts by examining what the necessary and sufficient conditions may be for some given property to count as a disposition/power as understood under dispositionalism. As we saw, powers should be intrinsic and not purely relational. Moreover, they should metaphysically necessitate the states of affairs they bring about. However, they should also be intrinsically defeasible: that is, they should metaphysically necessitate their outcomes *ceteris absentibus*, namely in the absence of other powers interfering with them (Mørch, *Ibid.*).

Mørch (2018: 10–12), then, suggests that it is inconceivable that motivational experiences, most notably pain and pleasures, may not elicit certain volitions in the subjects who have them. In particular, it seems inconceivable that the experience of pain may not elicit in those who have it the volition to avoid it, meaning that it seems inconceivable that people may feel pain without trying to avoid it, or at the very least without *desiring* or *whishing* to try to avoid it. Likewise, it seems inconceivable that someone may feel pleasure without *eo ipso* whishing to try to pursue it. Notice first, that Mørch (*Ibid.*) is not claiming that it is inconceivable that pain does not cause avoidance behaviour, or that pleasure does not cause pursuit behaviour. Rather, she's claiming that it is inconceivable that pain does not elicit *the volition* to initiate avoidance behaviour. That is to say, although it is very much conceivable that someone may want to avoid pain without actually avoiding it or even without trying to avoid it, it seems inconceivable that someone be in pain without *wanting* or at least *desiring* to try to avoid it: while volitions may be severed from their effects, namely actions, it seems that phenomenally conscious motivational states may not be severed from the volitions they elicit. Notice, plus, that the principle whereby powers metaphysically necessitate their outcomes only in the absence of other powers interfering with them is apparently still verified. For instance, it is very

much conceivable that one might not want to avoid pain, or even want to pursue it, because she believes that it will elicit more pleasure and/or will produce some other desirable outcome(s) in an immediate future – as when enduring the sting of a needle to do blood tests, cleaning a wound, or in masochism (Mørch, 2018: 10).

Moreover, and crucially, it seems inconceivable that pain may not ground *ceteris absentibus* the volition to initiate certain patterns avoidance–behavior and that pleasure may not ground *ceteris absentibus* some patterns of pursuit–behavior *in virtue and just in virtue of their phenomenal character*, that is in virtue of how the way it is like for someone to have them. On this regard, Mørch (2018: 11) writes:

> To support this, try to imagine a painful experience, such as stepping on a sharp nail, as vividly as possible in terms of how it feels. The phenomenal character of such an experience can only be described as intrinsically disagreeable and repulsive. How could such an experience make someone try do anything else than avoid it solely in virtue of feeling like that? Correspondingly, the phenomenal character of pleasure can only be described as intrinsically agreeable and attractive. How could it make subjects try to do anything else than try to pursue it solely in virtue of feeling like that? It seems someone who conceives of pain and pleasure as having different effects in virtue of how they feel must either not know (or fail to vividly imagine or remember) how they feel, or implicitly assume that their effects do not derive from how they feel.

Now, one could argue that such a conceivability criterion for some given property to count as a genuine disposition may be too strong. However, it is hard to deny that it is sufficient, provided by inconceivability one means ideal inconceivability, namely, to use once again the words of Mørch (2018: 9) «inconceivability by an ideal rational subject who is fully acquainted with the nature of the items she is conceiving of». This brings us to a first point. Supposing Mørch is on the right track in saying that it is inconceivable, and perhaps even ideally inconceivable, that certain motivational experiences may not ground certain volitions in those who have them

ceteris absentibus, in virtue of a line of reasoning on which I've already elaborated extensively throughout the essay, *prima facie* one might think that the arguments of Mørch may be deployed to cast doubt on whether the powerful qualities view can really be used to defend the possibility of dual carving. For take the concept <PHENOMENAL PAIN> and the concept <POWER TO ELICIT THE VOLITION TO INITIATE PAIN–AVOIDANCE BEHAVIOUR IN THE ABSENCE OF INTERFERING POWERS> and suppose for the sake of argument they were cognitive independent, meaning that a subject mastering both the concepts and accepting the relevant amount of background knowledge K as true would not be put on that basis alone in a position to rationally judge that the two are necessarily coextensive, no matter the amount of background knowledge she accepts as true. As one could say, from this would follow that our subject could conceive at least one situation in which the referent of <PHENOMENAL PAIN> may be instantiated without the referent of <POWER TO ELICIT THE VOLITION TO INITIATE PAIN–AVOIDANCE BEHAVIOUR IN THE ABSENCE OF OTHER INTERFERING POWERS> being instantiated, namely a situation in which someone is in pain without wishing or desiring to try to avoid it in the absence or other interfering powers, for if she could not conceive such a situation, then she would be in a position to rationally judge that the two concepts are necessarily coextensive. But that it is inconceivable that pain does not elicit the volition to initiate some pain–avoidance behaviour *ceteris absentibus* is precisely what Mørch states. Given this, one may cast doubt on whether the powerful qualities view can really be deployed to argue for the possibility of dual carving which, recall, states that two concepts may be cognitive independent and yet necessarily coextensive and both transparent, *contra* the third premise of the argument from revelation, and hence assumes that phenomenal and physical concepts are cognitively independent.

What is more, dual carving claims that the nature of one and the very same entity may be revealed in two or more conceptually independent ways. Hence, physicalist dual carvers do accept that phenomenal concept are either transparent or translucent.

I now whish to argue that phenomenal transparency, for rather straightforward reasons, enhances Mørch's point that, provided it is inconceivable that at least some phenomenal properties are essentially dispositional, then those experiences really are essentially dispositional.

As Mørch (2018: 11–12) argues, assuming that it is inconceivable that the experience of pain did not go along with the volition to initiate some pain–avoidance behaviour *ceteris absentibus*, or even that it is metaphysically necessary that whenever there is pain *eo ipso* there's the the volition to initiate some pain avoidance behaviour, is not sufficient, *per se*, to prove that the experiential property of being in pain necessarily elicit its outcomes in virtue and only in virtue of their qualitativity, namely its painfulness. In fact, one may accept that pain necessitates some given outcomes but buy into a version of the so called *governing laws view*, whereby it does so not in virtue of its qualitative nature, but rather in virtue of some irreducible laws that necessarily connects it to its outcomes but are extrinsic to its qualitative nature (see Armstrong, 1978; Morch, 2018: 9). Alternatively, one may adopt some kind of Humean *regularity view*, or appeal to epiphenomenalism, whereby phenomenally conscious mental states have no effect whatsoever[10].

What Mørch is trying to establish, however, is that dispositionalism entails or at the very least might entail panpsychism, meaning that *if* dispositionalism is true, then panpsychism most likely follows. Thus, on the assumption that dispositionalism is true – and hence that dispositions are to be taken as fundamental and irreducible properties – if it is true that pain cannot be conceived of if not as necessarily eliciting the volition to avoid it, the most plain thing to say is that the latter fact holds in virtue and only in virtue of the qualitativity of pain, that is, in virtue of its painfulness.

At this point, revelation enters the scene. First of all, as we just saw Mørch

[10]This allows me to mention another point. As Mørch herself (2018: fn. 18) notes, opponents may argue that 'pain necessitates the volition to avoid it *ceteris absentibus*' may be read as 'pain is necessarily followed by the volition to avoid it, excepts when it is not', which is a truism. However, on the assumption that dispositionalism is true – and hence it is true that powers are irreducible properties – 'pain necessitates the volition to avoid it *ceteris absentibus*' reduces to 'pain necessitate the volition to avoid it in the absence of other powers'. The latter is not a truism.

(2018:9, emphasis mine) defines ideal inconceivability as «inconceivability by an ideal rational subject who is fully acquainted with the nature of the items she is conceiving of». But phenomenal transparency states precisely that deploying a phenomenal concept *does get us* fully acquainted with the nature of phenomenal properties themselves. Hence, assumed phenomenal transparency, from the fact that one cannot think about about pain phenomenally and just phenomenally without conceiving it as necessitating the volition to initiate some avoidance behaviour seems to follow that pain really is essentially dispositional, meaning that it really is part of the nature of pain that it necessitates the volition to initiate some avoidance behaviour just in virtue of the way it feels.

Moreover, I've already elaborated extensively on the reason why phenomenal transparency leads, when combined with the positive *desideratum* of physicalism – namely with the *desideratum* whereby the phenomenal metaphysically supervene upon the physical – to at least some fundamental physical entities being essentially entirely phenomenal, *contra* the negative *desideratum* of physicalism, which states that there cannot be fundamental phenomenal entities. At any rate, just to recap, assume a dispositionalist conception of theoretical physics, according to which what physical concepts reveal of their referents are dispositional/functional roles. Notice that a dispositional conception of theoretical physics does not imply, *per se*, neither the powerful qualities view nor dispositionalism – *i.e.*, it does not imply that all there is to the nature of physical properties is their being dispositions/qualities, nor that fundamental physical dispositions do not require categorical/qualitative grounds and are irreducible in themselves. To establish, via physicalism, dispositionalism and a version of the powerful qualities view a further assumption is needed, namely physical transparency. In fact, if physical concepts characterize their referents in terms of their causal/dispositional roles, and if they are transparent, then all there is to the nature of physical properties is their having certain dispositional roles and so, provided, via physicalism, that every fundamental entity is a physical entity, it follows that every fundamental entity is dispositional, and hence at least some

dispositions are fundamental and irreducible, as per dispositionalism.

Therefore, assume, besides a dispositionalist conception of theoretical physics, that both physical and phenomenal concepts are transparent, and that, as per a physicalist overall worldview, only fundamental physical properties carve nature at its joints, meaning that every phenomenal concept is ultimately necessarily coextensive with a fundamental physical one, and hence every phenomenal property is ultimately necessarily nothing over and above a fundamental physical property, regardless of whether subjects are in a position to rationally judge the latter to be the case only on the basis of mastering the relevant concepts or not. Provided it is inconceivable that pain conceived phenomenally and only phenomenally may not elicit the volition to try to avoid it, in virtue of phenomenal transparency we have that certain phenomenal properties are essentially dispositional. Moreover, given a dispositionalist conception of theoretical physics and the transitivity of the relation of necessary co–referentiality between concepts, we have that at least some fundamental physical dispositional are both essentially dispositional and essentially phenomenal.

From this follows that at least some fundamental physical entities are both essentially dispositional and essentially phenomenal. This is already sufficient *per se* to rebut the appeal to the powerful qualities view to defend physicalism, which, as we saw, needs to assume that no fundamental physical disposition is essentially phenomenal and, more in general, is sufficient to confute physicalism, which states that no experience–involving entity can be fundamental. Hence, those who wanted to appeal to the powerful qualities view to defend dispositionalism against the two arguments for panpsychism I've mentioned at the beginning of this section are committed to deny phenomenal transparency, *contra* what Taylor (2013) suggests. Let me now briefly elaborate on what further remarks Mørch (2018) adds to the case that dispositionalism might entail panpsychism.

I've just discussed Mørch's claim that it seems inconceivable, and perhaps even ideally inconceivable, that certain motivational states, most notably pains and pleasures, do not elicit certain volitions in those who have it. To make her case

that dispositionalism entails panpsychism Mørch needs a further premise, namely that the only essentially dispositional qualities that we know of or can conceive are phenomenal ones. In fact, she (2018: 12–3) argues that when it comes to non–phenomenal properties derived from physics it seems that we have no problem in conceiving, say, that two solid objects may pass through each other, or that two negatively charged particles may attract each other instead of repelling each other. Notice that in order to establish a genuine parallel with the pain example, physical properties must not be conceived in terms of the dispositions physics ascribes to them. For it is clear that if physical properties are conceived in terms of their dispositions, it then becomes inconceivable that they come apart from their effects. But this reflects only a logical relation between concepts, not a connection between distinct properties (Mørch, *Ibid.*). And now, again, it seems conceivable that two solid objects may pass through each other when solidity is conceived as a quality, that is in terms of how it looks or feels, rather than as a disposition. When it comes to electric charge, we may not have any non–dispositional concept available, for we may have no means of understanding how charged particles look or feel, their qualitative character. But again, this does not seem to undermine the case that, unlike in the case of pain and other phenomenally conscious motivational states, when physical properties are conceived as qualities rather than as dispositions it becomes conceivable that they may come apart from their effects[11].

One may contend that there may other weaker, less demanding criteria than the conceivability/inconceivability one for some given property to count as a genuine quality/disposition but argue that such criteria are still sufficient, so that after all the conceivability/inconceivability criterion might not be necessary (see Mørch, 2018: 9). On this regards, Mørch (*Ibid.*) writes the following:

> [...] supposing some phenomenal properties actually did satisfy the conceivability criterion, but no physical properties satisfied it, then even if physical properties still were to satisfy some weaker criterion, it

[11] Morch (2018: 13) argues against the possibility of conceiving neutral – *i.e.*, neither physical nor phenomenal – qualities – *i.e.*, intrinsic, non–purely–structural properties.

> would seem reasonable to conclude that our experiences or conceptions of physical powers are not as revelatory as our experiences or concepts of phenomenal powers. That is, if phenomenal properties satisfy the conceivability criterion but physical properties do not, then the physical concepts/experiences should at the very least not be taken to afford us complete insight into the nature of dispositionality, given that the kind of insight that enables us to recognize necessity by means of inconceivability is in fact available in the phenomenal case. And if our conception of physical dispositionality is incomplete, then it must have some hidden aspect, which panpsychists may argue is phenomenal.

To put the matter in the terminology I've been using thus far in the essay, on this picture we would have that phenomenal concepts are transparent and physical concepts are translucent. That is to say, although it is part of the nature of the fundamental entities that they have certain dispositions, and subjects are put in a position to know this *a priori* just on the basis of mastering the corresponding physical concepts, what they are *not* put in a position to know *a priori*, only on the basis of mastering the relevant physical concepts, is that physical properties play the dispositional roles they play, meaning that they necessarily elicit certain outcomes *ceteris absentibus*, in virtue and only in virtue of their phenomenal/qualitative nature.

This brings me to a potential ambiguity of Mørch. In fact above I've argued that, assumed a dispositionalist conception of theoretical physics, to infer dispositionalism and hence a version of the powerful qualities view the physicalist is committed to assume physical transparency, from which follows that all there is to the nature of fundamental physical properties is their being dispositions. As we just saw, Mørch seems to suggest, instead, that physical concepts are translucent. In 2.13 I've argued that physical translucency leads at least to a form of property dualism, whereby there are two ontologically distinct aspects or components to the nature of fundamental physical entities, and possibly either to a form of mysterianism or to a form of pluralism. On the former view, no combination of concepts, no matter

how extended and no matter the concepts it deploys, could ever get us to know the entire nature of fundamental physical entities; on the latter, there would be a plurality of ontologically distinct essential aspects of fundamental physical entities. That panpsychist Russelian monism may entail property dualism is something that also Chalmers (2019: 13–4) suggests:

> The key respect in which micro–idealism goes beyond constitutive Russellian monism is its purity: it holds that all (and not merely some) fundamental properties of micro–entities are mental. [...] In holding that all fundamental properties are of the same kind, micro–idealism offers a simple and unified monistic view of nature. By contrast, impure versions of panpsychism have both mental and nonmental properties at the fundamental level, yielding a sort of property dualism.

Let me go a bit more slowly. The point I'm trying to make is that since on the powerful qualities view *all there is* to the nature of qualities is their being dispositions and *vice versa*, concepts that characterize certain (fundamental) properties as dispositions, as physical concepts do, would have to be transparent, and not translucent.

In other words, it is not entirely clear whether, on Mørch's view, (fundamental) phenomenal properties *ground* certain dispositions or whether, instead, they *are* dispositions. On the former approach, even assuming that, as Mørch suggests, it is inconceivable that certain phenomenal properties did not had the power to elicit certain effect *ceteris absentibus*, it may still be conceivable that the the power were grounded on something else and hence occurred in the absence of the phenomenal properties. Again, this is something that Chalmers (2019: 16, emphasis mine) has noted in discussing Mørch's view:

> Mørch (forthcoming) has argued for a phenomenal powers views on which phenomenal states are or metaphysically ground certain causal powers or dispositions. For example, the experience of pain might ground a disposition to avoid certain situations, while the experience of love might

> ground a disposition to associate with certain people. On one version of this view, the phenomenal state without the power is inconceivable and metaphysically impossible *(even if the power without the phenomenal state is conceivable and possible).*

Above I've suggested that, supposing Mørch is on the right track in saying that it is inconceivable that certain phenomenal states may not elicit the volition to initiate certain patterns of behaviours *ceteris absentibus*, one might want to cast doubt on whether the powerful qualities view may really be deployed to argue for the possibility of dual carving, for it is not clear whether concepts of qualities and concepts of disposition really are cognitive independent. Now: if the relation between qualities and dispositions is taken to be one of grounding, then it might conceivable that certain powers did not go along with certain (phenomenal) qualities, which implies that concepts of powers and concepts of dispositions may turn out, after all, to be cognitively independent. For recall that two concepts are cognitively independent iff a subject mastering both of them and accepting a relevant amount of background knowledge as true would be in a position to rationally judge that they are necessarily coextensive. But the relation of (necessary) co–referentiality is symmetrical by definition, meaning that if the concept A is co–referential with B, B must be co–referential with A. So, not only would it have to be inconceivable that certain (phenomenal) qualities did not go along with certain powers: it would also have to be inconceivable that the powers did not go along with the (phenomenal) qualities. If, on the other hand, the relation between qualities and dispositions is taken to be one of identity, than if sponsors of the powerful qualities view could somehow manage to show that it is conceivable that certain powers did not go along with certain qualities even though there is a relation of identity between the former and the latter, then the powerful qualities view may count as a form of dual carving. I'm skeptical as to whether the latter may be done. At any rate, I don't need to decide on this matter here, for as I argued, when combined with either full–revelation/phenomenal–transparency or partial/revelation/phenomenal–translucency the powerful qualities view still leads

at least to a micropsychist theory whereby at least some fundamental entities are essentially experience–involving.

To sum up, Mørch's point is that dispositionalism may entail panpsychism. I agree with this, provided revelation is assumed. However, dispositionalism claims that dispositions do not require qualitative grounds. The latter point may be defended via the powerful qualities view (though it does not need to), which states that, since dispositions just are qualities, they don't need qualitative grounds over and above themselves. It is not entirely clear whether on Mørch's view the relation between fundamental (phenomenal) qualities and powers is one of identity or one of grounding. Provided sponsors of the powerful qualities view could somehow manage to show that it may be conceivable that certain powers did not go along with the corresponding qualities and/or *vice versa* even though there is a relation of identity between the two, the powerful qualities view *may* count as a form of dual carving. I'm skeptical that the latter may be done. Be as it may, physicalists sponsors of the powerful qualities view are still committed to deny revelation. Moreover, since, assumed that only fundamental physical properties carve nature at its joints, partial–revelation/phenomenal–translucency also leads to some fundamental physical entities being at least partially essentially experience–involving, physicalist sponsors of the powerful qualities view need to deny partial–revelation/phenomenal–translucency as well.

3.6 The ability hypothesis

According to the so called ability hypothesis, of which Lewis (1983_c; 1988), Nemirow (1980; 1990) and Teller (1992) are the main sponsors, what Mary acquires upon her release is not a new amount of propositional knowledge regarding color vision, but rather a bunch of abilities, such as the ability to imagine, recognize, remember, *et cetera* certain experiences. In other words, according to the ability hypothesis the new knowledge of Mary is not a form of know–that, or propositional knowledge, but rather a form of know–how, or practical knowledge. Crucially, as Tye (2000) argues,

the ability hypothesis does *not* imply that there cannot be, not even in principle, any amount whatsoever of conceptual/propositional–knowledge which Mary will acquire or will be put in a position to acquire upon her release. Indeed, it would be absurd to say so. The ability hypothesis only claims that there is *one* kind of knowledge which Mary could *only* acquire via having the experience of seeing something red, and *this* kind of knowledge is a form of know–how (Nida–Rümelin and O'Conaill, 2023: sec. 4.3). The latter, however, is not incompatible with the possibility of there being some amount of propositional knowledge which Mary could acquire both via having the experience of seeing something red and in some other ways.

What is more, as Loar (1990/1997) and Goff (2017: 70–71) (but see also Chalmers, 1996: 129), among many others, have argued, to my eyes, rather cogently, overall the ability hypothesis seems deeply implausible, for there is a very strong *prima facie* case that Mary will indeed acquire some propositional knowledge upon her release. For instance, in wondering wether other people have the same kind of experience she is having, she might consider the force of an argument such as this (see Loar, *Ibid.*; Goff, *Ibid.*):

[1]It is like <THIS> for me to see red.

[2]If it is like <THIS> for me to see red, then it must be like <THIS> for every other human being to see red.

[3]Jacopo is a human being.

Conclusion: It is like <THIS> for Jacopo to see red.

Provided upon her release Mary may consider the force of an argument such as the one just presented, the only way to make sense of this is to say that the phrase 'it is like <THIS> for me to be red' is truth–evaluable, which requires that Mary has some new conceptual knowledge (Goff, *Ibid.*). Moreover, it seems something reasonable to say that Mary could not evaluate the strength of such an argument if not upon having seen something red at least once.

At any rate, as long as we agree that Mary may deploy some phenomenal concepts upon he release, since both full–revelation/phenomenal–transparency and partial–revelation/phenomenal–translucency are inconsistent with physicalism, it will still be at least necessary for sponsors of the ability hypothesis to deny both of them, arguing that phenomenal concepts are either mildly opaque or radically opaque, though it may not be sufficient, for physicalist sponsors of the ability hypothesis need also argue that the kind of kind of knowledge Mary could only acquire upon her release is not conceptual/propositional.

3.7 The acquaintance hypothesis

Another version of the no–propositional–knowledge–strategy claims that what Mary acquires is a form of knowledge by acquaintance. Earl Conee (1994) is among the main sponsors of such an approach.

Notoriously, Russell (1910; 1912) drew a distinction between two kinds of knowledge a person might come to have, namely *knowledge of truths – i.e.*, propositional knowledge – and *knowledge of things – i.e.*, objectual knowledge. Russel held that knowledge of things is split into two sub–kinds, namely *knowledge by description* and *knowledge by acquaintance*, and maintained that while knowledge by description is grounded on a subject knowing at least some true propositions regarding a given particular, knowledge by acquaintance cannot be reduced to knowledge–of–things/propositional–knowledge, in that it is both simpler than and logically independent from the latter. More specifically, Russell (1912: 73) described knowledge by acquaintance as a form of direct knowledge: we know by acquaintance those particulars that are immediately and directly presented to us (on this see Gertler, 2012 and Pallagrosi and Cortesi, forth. among many others). He (1910: 108) also described knowledge by acquaintance as presentational, rather than *re*presentational. In particular, as Gertler (*ivi:* 95) underlines, the relation of acquaintance has been typically understood as both epistemically and metaphysically direct, meaning that it does not depend upon any process of inference – nor upon the awareness of anything

else besides the item one is aware of or presented with – nor is it mediated by any other relation besides the one of direct–awareness/presentation.

Now, according to Conee (1994: 144) knowledge by acquaintance «requires the person to be familiar with the known entity in the most direct way that it is possible for a person to be aware of that thing». Such a statement clearly suggests that before being released Mary could not get to know the way it is like for someone to see something red by acquaintance, for it seems clear that having an experience firsthand provides a subject with a sort of direct access to the way it is like to have it which is prevented to those who haven't had the experience themselves.

In light of this – as Goff (2017: fn. 6) argues – physicalist fans of the acquaintance–hypothesis would better assume a conception of acquaintance such that being acquainted with a given entity/property/... does not convey any information about the latter, for otherwise the acquaintance hypothesis would strengthen the argument rather than offering a response to it (Goff, *Ibid.*). Indeed, Conee (*ivi:* 139) explicitly contests that there is some information Mary lacks before being released. In fact, if the relation of acquaintance were construed as conveying a certain amount of (non–propositional) information about its targets, then a sub–version of a weaker, epistemic version of the knowledge argument may be drawn. This epistemic version of the argument would go as follows:

[1]By assumption, Mary has complete physical information about color vision already before her release.

[2]There is some information about color vision which Mary acquires by acquaintance upon her release. Moreover, Mary could not have acquired this information prior to being released. Hence, this information is new for her.

Conclusion: not all information about color vision is physical information.

What I now whish to do, is to offer an "enhanced" version of an argument Anna Giustina (2021) calls the *argument from phenomenal–concept–acquisition*. As far as I can see, as it is formulated by Giustina the argument already allows to

establish a conclusion which, as I just argued, may be deployed to draw a weaker, epistemic version of the knowledge argument – assumed that what Mary gains is a form of knowledge by acquaintance – namely that being acquainted with the phenomenology of one's own experiences *per se* conveys (non–propositional) information about the phenomenology itself. The "enhanced" version of the argument I will offer contains Giustina's original argument within it, but combines with a further assumption, namely full–revelation/phenomenal–transparency, even though as far as I can see very similar conclusions may be reached even if instead of full–revelation/phenomenal–transparency one assumed partial–revelation/phenomenal–translucency. By combining the original formulation of the argument with phenomenal transparency I hope to reach a further conclusion, namely that it is necessary and possibly necessary and sufficient for physicalist sponsors of the knowledge–by–acquaintance–hypothesis to deny the principle Liu (forth.) calls PHENOMENAL NARD2, on which I've already elaborated in 2.9, in order to rebut the argument. But since, provided certain minimal assumptions are in place, full–revelation/phenomenal–transparency is arguably logically equivalent with PHENOMENAL NARD2, if the latter is denied, full–revelation/phenomenal–transparency is denied as well.

Giustina (*ivi:* 6–7) starts by assuming a non–atomistic view of concepts whereby given concepts may be "built up", so to say, by composition from other ones: for instance, the concept <BACHELOR> may acquired by combining <UNMARRIED> and <MALE>. Acquiring a concept means coming to have it or possess it, when one possesses a concept when she is able to deploy it in thought, cognition, action–guidance, *et cetera* (*Ibid.*). Those concepts that are possessed but not acquired are called innate (*Ibid.*) Then, Giustina (*Ibid.*) proceeds to draw a distinction between *basic* and *non–basic* phenomenal concepts. Non–basic phenomenal concepts are those that are acquired by composition from other phenomenal concepts. Among this family of concepts we may include concepts like <EXCRUCIATING PAIN> <BITTER–SWEET SENSATION> and so on (*Ibid.*). It is clear that not all phe-

nomenal concepts – indeed, not all concepts in general – could be non–basic, on the pain of an infinite regress.

Basic phenomenal concepts, then, are those that are not acquired by composition from other phenomenal concepts. These may be concepts like <PAIN> <VISUAL PERCEPTION> <SENSATION> and the like. Truth be told, it may not be so straightforward to establish whether, say, <PAIN> is basic or whether, instead, it is obtained by composition from, say, <CORPOREAL SENSATION> and <UNPLEASANT STATE>. Likewise, <COLOR PERCEPTION> may have <VISUAL PERCEPTION> as one of its components. At any rate, strictly speaking it is not necessary that we agree on what concepts count as basic and what other count as non–basic, as long as we agree that, assumed a non–atomistic view of concept[12], if some concepts are acquired by composition from other ones then at least some concepts need to be basic.

That being said, the argument from phenomenal–concepts–acquisition goes as follows (Giustina, *ivi:* 8):

(P1) (Almost) all basic phenomenal concepts are acquired.

(P2) For most basic phenomenal concepts, if they are acquired, they are acquired by introspection.

(P3) If all introspective states are conceptual, then it is not the case that most basic phenomenal concepts are acquired by introspection.

(C) Not all introspective states are conceptual.

Here is how according to Giustina each premise of the argument may be justified. Denying the first premise would imply a rather implausible form of nativism whereby most of our basic phenomenal concepts or all of them were innate, *i.e.*, possessed but not acquired. Although there may be some innate basic phenomenal concepts, which is what justifies the '(almost)' bit in the premise, it seems very implausible to say

[12]But as Giustina (*ivi:* fn. 12) argues, on an atomistic conception *all* concepts are basic. Hence, as far as I can see the argument may still be drawn, for it targets basic phenomenal concepts.

a huge number of them, including concepts like <OLFACTORY EXPERIENCE>, <SWEET>, <BITTER>, <COLOR PERCEPTION> and the like, are innately possessed by us.

As for the second premise, it states that, provided most of our basic phenomenal concepts are acquired, they must be acquired via introspection, or at least it is most likely that the latter is the case. On some transparency[13] theories, some phenomenal concepts such as color–phenomenal–concepts like <PHENOMENAL RED> are not acquired via introspection, but rather via the perception of some mind–independent entities such as colored surfaces. The problem with this approach, is that while it may be viable for concepts referring to visual experiences and other perceptual–experiences and maybe for some concepts referring to some bodily and/or somatosensory experiences, when it comes to mood–concepts like <NOSTALGIA> and other concepts of this sort it is not at all clear whether it can be profitably pursued, for it is very controversial whether, e.g., moods and other emotive states really get us in contact with some mind–independent objects. Moreover, sponsors of the transparency–approach would have to say that introspection does not get us in contact, in turn, with the phenomenology of the emotive condition in question, but just with the object it supposedly is about[14].

What is more, provided most of our basic phenomenal concepts are acquired via introspection, it is most likely that they are acquired via the introspection *of the experience they are about*[15]. As we shall see, this point will be of some importance when it will come to showing that the combination of the acquain-

[13]Basically, transparency theorists claim that introspection does not give access to the way it is like to undergo a given experience, but only on the mind–independent objects the experience is about. The literature on transparency is vast and articulated: exploring it in depth transcends the scopes of this dissertation. Interested readers may have a look at Harman (1990), Speaks (2009). Tye (1992, 1995, 2000), and Thau (2002), as well as Crane and French (2021) and for critical discussions, Martin (2002a), Smith (2008), Stoljar (2004) and Soteriou (2013).

[14]Here I am merely trying to summarize Giustina's (2021) defense of this and the other premises of the argument, which is much more rich and articulated than the few remarks I'm drawing here. Hence, I advice readers to go and have a look at her paper.

[15]As Giustina (2021: 9–10) notes, it may be the case that some phenomenal concepts may possibly be acquired via the introspection of some other experiences. For instance, the concept <PHENOMENAL RED2> may be acquired by extrapolation upon the introspection of "reddish$_1$" experiences and "reddish$_3$" ones. However, this is far from being uncontroversial. Moreover, it clearly cannot be true of all or even most of our basic phenomenal concepts.

tance hypothesis with full–revelation/phenomenal–transparency (as well as partial–revelation/phenomenal–transparency) threatens physicalism.

As for the third and last premise of the argument, here is how it is defended by Giustina (*ivi:* 15–17). First of all, an introspective state is said to be conceptual iff it involves the deployment of some concept(s) already possessed by the introspector. Now suppose that, as the second premise of the argument states, it is true that most of our basic phenomenal concepts are acquired via the introspection of the experiences they are about, and assume for the sake of argument that all introspective states are conceptual in the sense just defined. By the law of the excluded middle, either (a) introspecting a given experience E involves the deployment of a phenomenal concept of E or (b) introspecting an experience E requires the deployment of some phenomenal concept(s) that is/are about some other experience(s) than E itself. Now (a) must immediately ruled out, for it implies the paradoxical conclusion whereby while, on the one hand, acquiring a phenomenal concept of an experience requires introspecting that experience, the introspection of the experience requires the possession of a phenomenal concept of it, hence, on this option, no phenomenal concept could ever be acquired. Concerning option (b), Giustina (*Ibid.*) argues that such an option doesn't work for *basic* phenomenal concepts, for if acquiring a given basic phenomenal concept C required an introspective state, and the latter, in turn, required the possession of some other phenomenal concept(s) than C itself, then basic phenomenal concepts could not constitute the "foundational layer" upon which all other non–basic phenomenal concepts may be formed. From this follows that, as this premise of the argument states, if most basic phenomenal concepts are acquired, then not all introspective states can be conceptual. The conclusion of the argument follows accordingly[16].

[16]As Giustina herself (2021: 4) acknowledges, there may be *pure phenomenal demonstratives* a la Gertler (2001: e.g., 323) like <THIS> and <THAT> (see also Pallagrosi and Cortesi, forth: fn. 11). Another possible exception to the conclusion of the argument from phenomenal–concepts–acquisition may be constituted by what Chalmers (2003_a) calls *direct phenomenal concepts*. These are concepts that are formed upon attending to the phenomenology of a given experience and whose content is at least partially constituted by the phenomenology itself. In her (2021), Giustina seems at least open to the possibility that, provided pure demonstratives and/or direct phenomenal concepts exist, this very peculiar kind(s) of concepts may be involved in what she calls primitive introspection,

My "enhanced" version of the argument from phenomenal–concepts–acquisition goes as follows:

[1]Full–revelation/phenomenal–transparency: phenomenal concepts are transparent. [assumption]

[2]If phenomenal concepts are transparent, phenomenal concepts exist. [from 1]

[3]If phenomenal concepts exist, (almost) all of them must be acquired.

[4]Non–basic phenomenal concepts are acquired by composition from basic ones.

[5](Almost) all basic phenomenal concepts must be acquired.

[6]For most basic phenomenal concepts, if they are acquired, they are acquired by introspecting they experience they are about.

[7]If all introspective states were conceptual, it could not be the case that most basic phenomenal concepts were acquired by introspecting the experience they are about.

[8]Some introspective states are not conceptual [from 3–7: the argument from phenomenal–concepts–acquisition].

[9]The essence of experiences is fully revealed to subjects pre–conceptually. [from 1 and 8]

namely, the supposedly non–conceptual sort of introspection which grounds the formation of any other phenomenal concept, as per the argument from phenomenal–concepts–acquisition. Indeed, in Pallagrosi and Cortesi (forth.) we argue that a stalemate between Giustina's (2021; 2022) account of introspective knowledge by acquaintance, on which primitive introspection is taken to be utterly non–conceptual, and Chalmers' (2003_{a}) one, on which introspection involves the deployment of at least a direct phenomenal concept, may be looming in the contemporary debate. However, it seems to me that none of this affects my claim that in order to avoid the threat posed by the thought–experiment beneath the knowledge–argument it is necessary that physicalists of any variety, including physicalists fan of the knowledge–by–acquaintance–hypothesis, deny both full–revelation/phenomenal–transparency and partial–revelation/phenomenal–translucency. For as they are worded, neither phenomenal transparency nor phenomenal translucency are restricted to a peculiar sub–class of phenomenal concepts: rather, both the theses state that all phenomenal concepts, including pure demonstratives and/or direct phenomenal concepts, if anything like them exists, are transparent/translucent. So, even if all introspective states were taken to involve at least a direct phenomenal concept, assumed either phenomenal transparency or phenomenal translucency, such concept would be transparent/translucent, and provided both phenomenal transparency and phenomenal translucency are inconsistent with physicalism, my claim would still hold.

[10]If the essence of experiences is fully revealed to subjects pre–conceptually – then physicalism is false.

Conclusion: physicalism is false.

Now, the first premise of this "enhanced" version of Giustina's original argument is just an assumption I'm making: in fact, the point I'm trying to make is that if phenomenal transparency were true, physicalist advocates of the acquaintance hypothesis would be doomed anyway, hence it is necessary and possibly sufficient that they deny phenomenal transparency itself as well as phenomenal translucency. The second premise of the argument is a truism: it is self–evidently true that phenomenal concepts must exist in order to be transparent. The premises from [3] to [8] are just a reiteration of Giustina's argument, so I won't say anything more than what Giustina says and I have already attempted to summarize above. Of course, what is of greater relevance for our purposes here are the last two premises, from which follows the conclusion.

In 2.9 I argued that from phenomenal transparency follows the thesis of revelation as stated by Michelle Liu (forth: 3): «By having an experience–token with phenomenal property Q, S is in a position to know that 'Q is X', where the predicate 'X' captures the essence of Q» – I won't repeat the argument here. Following Liu (*Ibid.*), I've also shown, in the same section, that from the thesis of revelation as thus stated follows a version of what Liu calls the NARD principle, namely PHENOMENAL NARD2: «all there is to the essence of Q is X if and only if in having an experience–token with Q, S experiences the essence of Q as X and only as X» (Liu, *Ibid.*). Thus, from phenomenal transparency follows PHENOMENAL NARD2. More precisely, in 2.9 I argued that there are good reasons to believe that phenomenal transparency and PHENOMENAL NARD2 are logically equivalent, meaning that they are either both true or both false (of course, provided one assumes that phenomenal concepts exist in the very first place, plus a number of other minimal and rather plausible assumptions, such as that coming to have an experience may already be sufficient to form a phenomenal concept of it).

Now, as we saw the argument from phenomenal concepts acquisition allows to infer from the claim that most of our basic phenomenal concepts are acquired the conclusion that some introspective states are not–conceptual. Indeed, offering and/or defending a precise and detailed account of the nature of introspection is something that goes far beyond the scopes of this dissertation. Schwitzgebel (2019: sec 1) preliminarily characterizes introspection as «a process by means of which we learn about our own currently ongoing, or very recently past, mental states or processes», though he (*Ibid.*) acknowledges that no simple account of introspection is widely accepted. Above we saw that knowledge by acquaintance is typically defined as the kind of knowledge we have of those items we are directly presented with or, equivalently, of those items we are immediately aware of. Following Giustina (2022: 1) we may therefore define introspective knowledge by acquaintance as the knowledge we have of our own experiences by being directly and immediately aware of them in introspection. Moreover, It seems something plain to say that what we are most directly and immediately aware of in introspection is the way of our experiences feel, namely the way it is like to have them. At any rate, since Giustina's argument from phenomenal–concepts–acquisition is an argument about phenomenal concepts, namely concepts that characterize given experiences in terms of of the way it is like for someone to have them, it seems obvious to interpret its conclusion – accepting Schwitzgebel very minimal preliminary definition of introspection – as stating that we have a way of knowing the way it like to have our own experiences that precedes the formation of phenomenal concepts.

But according PHENOMENAL NARD2 the way it is like to have certain experiences coincides with (part of) their essence. So, we have that from phenomenal transparency follows PHENOMENAL NARD2 and from the argument from phenomenal–concepts–acquisition follows that we have a way of knowing what it is like to have certain experiences that precedes the formation of phenomenal concepts. Hence, by combining the argument from phenomenal–concepts–acquisition with phenomenal transparency, as I did, one is allowed to conclude that we have a way of knowing

the essence of our own experiences which is pre–conceptual in that it precedes the formation of the relevant phenomenal concepts – which is what the 9th premise of the argument states.

Now, according to a physicalist worldview experiences are nothing over and above a kind of physical states. So, assumed physicalism and assumed the picture we are considering, there would be a kind of physical states whose essence is either fully – phenomenal transparency – or partially – phenomenal translucency – revealed to subjects in introspection by acquaintance and pre–conceptually, that is prior to the formation of phenomenal concepts. This would require at the very least a radical reconception of what the physical is, whereby the essence of at least some physical states is immediately and directly revealed to subjects in introspection. I don't think many physicalists would be on board with such a radical reconception. Moreover, assumed, as per physicalism, that only fundamental physical properties carve nature at its joints, we would still have that at least some fundamental physical entities are both essentially entirely physical and essentially either entirely or partially phenomenal, *contra* the NE principle and *contra* the negative *desideratum* of physicalism.

If my analyses are correct, it is at least necessary and possibly sufficient, for physicalists sponsors of the knowledge–by–acquaintance hypothesis to deny PHENOMENAL NARD2 in order for them to avoid the threat posed by the knowledge–argument. But since – as I argued in 2.9 – there are good reasons to believe that (given certain minimal assumptions) PHENOMENAL NARD2 and phenomenal transparency are logically equivalent, if PHENOMENAL NARD2 is denied then phenomenal transparency is also denied and *vice versa*.

This also allows me to rebut an objection Damnjanovic (2012: 76) raises against the argument from revelation. Damjanovic asks us to imagine a person, Thomas, who's tasting peaches for the very first time in his life. As he (*Ibid.*) writes, one way to make the thesis of revelation compatible with physicalism is to claim that upon tasting peaches Thomas gets to know by acquaintance the taste of peaches itself. Once

the newly acquired knowledge of Thomas is interpreted as knowledge by acquaintance, according to Damnjanovic (*Ibid.*) «The argument from Revelation fails [...] because it incorrectly supposes that Thomas' complete knowledge of the taste of peaches implies that he knows certain truths about the nature of peaches». However, as far as I can see, even if the newly acquired knowledge of Thomas was interpreted only as knowledge by acquaintance, the physicalist would still be committed to deny PHENOMENAL NARD2, and hence both partial–revelation/phenomenal–translucency and full–revelation/phenomenal–transparency. For again, assumed PHENOMENAL NARD2 – *i.e.*, that the nature of experiences is at least partially constituted by the way it is like for subjects to undergo them – and assumed, as per physicalism, that only fundamental physical properties are spare/carve nature at its joints, it follows that at least some fundamental physical entities are both essentially entirely physical and essentially at least partially phenomenal. This holds regardless of whether (part of) the nature of phenomenal properties is known conceptually/propositionally or only by acquaintance.

If one combined the argument from phenomenal–concepts–acquisition with phenomenal translucency instead of phenomenal transparency, she would obtain the conclusion that part of the nature of experiences, though not the entirety of it, is revealed to subjects in introspection prior to the formation of phenomenal concepts. As far as I can see, such a conclusion would still most likely be at odds with a physicalist worldview.

Chapter 4

Looking back, and then forward

4.1 Part one: looking back

There is something puzzling about the way the contemporary debate on the metaphysics of conscious phenomena is perceived by some scholars. One the one hand, it is undeniable that the debate is way more lively now than it was fifty or sixty years ago. The 50^{s} and the 60^{s} of the last Century were characterized by the sheer predominance of physicalist materialism[1]. Back at the time, a picture of the people working in the field would have looked like one of those austere family portraits from the late 1800^{s} or the early 1900^{s}. Compared to then, a family portrait of the people working in the field now would look like the cover of *Sgt. Pepper's Lonely Hearts Club Band* by *The Beatles*. Although most – though not all – philosophers nowadays still take for granted that some version of physicalism ought to be true, after the publication of the works of authors like Nagel (1974), Kripke (1980), Jackson (1982), Levine (1983), Chalmers (1996) and Nida–Rümelin (1997; 2007), among others, forms of dualism have been rehabilitated and are now deemed at least as worthy of being considered alongside physicalism. In the 2000^{s}, there has been a revive of Russelian monism and panpsychism – of which Strawson (1994; 2008$_{a}$; 2009) has

[1] I am thinking about authors like Armstrong (1968); Feigl (1958), Lewis (1966) Place (1956), Putnam (1960), and Smart (1959) among others.

been among the pioneers, though not the sole one[2]. Recently, there have been some "stirrings" – as Chalmers (2019) calls them – of idealism – of which Foster (1982) and Sprigge (1983) have been among the pioneers[3]. Even illusionism – of which Dennet (1978; 1981; 1988; 1989) is among the pioneers[4] – is nowadays taken seriously (rather surprisingly, I must say). Indeed, one can agree that there is something to Siewert's (2011: 242) point that a non–negligible part of the contemporary research in the philosophy of mind has been and still is dominated by «the tyrannizing anxieties and ambitions of mind–body metaphysics».

Still, talking with a number of scholars, both early–career and ahead in their path, more than once I have sensed a growing feeling of dissatisfaction with the theme, as well as a certain skepticism with respect to the possibility of making any real progress in the near future. For sure, the point I am making is largely anecdotal, but the impression some people have is that many philosophers writing on the metaphysics of consciousness are somehow biased, meaning that few of them have changed their mind in the last few decades, and a number of them spend most of their time clarifying their own positions – sometimes even reiterating old points – rather than considering the "options on the table", so to say, namely the possible stances one could take with respect to certain issues, and subject them to slow and careful rational scrutiny. As a result, a number of scholars is starting to feel as if the debate is not going anywhere (again: the point I'm making is largely anecdotal)[5].

[2]See also Bruntrup and Jaskolla (2017) Chalmers (2013), Mathews (2003), Rosenberg, (2004), Skrbina (Ed.) (2009), Goff (2017), and Coleman (2014; 2017) among others.

[3]But see also Adams (2007), Kastrup (2017; 2018), Albahari (2019), Chalmers (2019), Builes (forth.), Bolender (2001), Goldschmidt and Pearce (forth.), Meixner (2017), Pelczar (2015), Yetter–Chappell (forth.), among others

[4]See also Frankish, (2012; 2016) Kammerer (2021) Shabasson (2022) Pereboom (2011, 2017) Graziano (2013, 2016) Humphrey (2016).

[5]Whether this impression these scholars have of the debate not going anywhere corresponds to reality is something that may be disputed. For a direction of the current mind–body dialectic may, perhaps, be individuated. In what he refers to as the "Hegelian synthesis argument," Chalmers (2016) has recently examined the development of the debate drawing inspiration from Hegel's dialectical approach involving thesis, antithesis, and synthesis. Chalmers traces this evolution through various forms of materialism, dualism, and panpsychism. The trajectory of the dialectic suggests a movement that brings consciousness progressively closer to the foundational essence of all existence. The latest stance is a type of panpsychism termed *cosmopsychism* which posits that the entire externally defined cosmos functions as an internally conscious subject (See also Albahari, 2019: 1). The trajectory just sketched is also vividly portrayed by Chalmers in this passage from his (2019: 1): «When I was in graduate school, I recall hearing "One starts as a materialist, then one

This overall feeling of dissatisfaction has been vividly expressed, I think, By Galen Strawson (2016: 10–11):

> I find the current debate deeply depressing. It takes place between those like myself who know that consciousness exists, know that there is a fundamental respect in which we know exactly what it is, and know that nothing in life is more certain, and those who seek to doubt or deny some or all of these things (even as—sometimes—they deny that they deny them). As far as I know almost no one who is engaged in this debate has ever changed sides, and I gave up trying to persuade anyone long ago. The most one can do from my side, without expecting anyone who disagrees to listen, is to try to make three things clear, as vividly as possible.

It is not up to me to say whether in this work I have managed to make things clear and vivid. Yet I do hope to have reached at least one goal: to have offered a radical simplification of a debate which is oftentimes perceived as in some way "clogged". This seems to me to be a pretty substantive result. As we saw in 1.2, even timidly overlooking the debate on how physicalism as a doctrine ought to be construed in the very first place, one almost immediately finds herself entangled in a jungle of extremely subtle distinctions between kinds of metaphysical relations, kinds of properties, *et cetera*. The same goes for the debate on what phenomenal concepts are, if they exist at all. Perhaps surprisingly, even fixing reference on experiential phenomena in the very first place proves to be an harder task than one might initially have thought, if it is true that even their existence has been put into question. As I see it, the debate on the metaphysics of conscious states rotates around whether three fundamental ideas are true or not, or at the very least it has rotated around

becomes a dualist, then a panpsychist, and one ends up as an idealist" [...] Some recent strands in philosophical discussion of the mind–body problem have recapitulated this progression: the rise of materialism in the 1950s and 1960s, the dualist response in the 1980s and 1990s, the festival of panpsychism in the 2000s, and some recent stirrings of idealism». For what it's worth, I myself happen to think that a form of cosmopsychist idealism, or something close to it, may possibly be the best metaphysical framework we have available (see the very last section of the essay).

these three fundamental ideas in the last fifty years or so: (1) conscious phenomena are such that their nature is revealed to subjects in introspection; (2) conscious phenomena are physical phenomena; (3) The NE principle (Strawson, 2008_a: 60): «physical stuff is, in itself, in its fundamental nature, something wholly and utterly non–experiential». Rather self–explanatorily, physicalism is committed to the truth of (2). Moreover, I take it that the vast majority of people who declare themselves as physicalists without adding any further specification would commit themselves to (3). Plus, most physicalists would probably add a further assumption, namely that (4) only physical properties carve nature at its joints. In fact, in principle one might accept (3) and perhaps also (2) but be a neutral monist and claim that physical properties are not fundamental, in that they are grounded on some more fundamental neutral properties, which are neither physical nor conscious. This, I take it, would be at odds with a physicalist worldview. In this work, I have argued that no version of physicalism is compatible with the truth of (1), namely revelation. Substance dualism may be obtained by denying (2) and keeping (3) plus possibly (1), even though I guess there may be forms of substance dualism which deny both (1) and (2). By denying (3) and keeping (2) and possibly (1) one obtains at least a panpsychist version of Russelian monism. Indeed, I would say that (1) is one of the main motivations behind panpsychist Russelian monism, for this view is based on the idea that conscious phenomena are or at least may be the only ones whose intrinsic nature we get to know. Depending on how the details of one's view are further spelled out, By denying (3) and keeping (1) and (2) one may also obtain a form of idealism. In particular, assumed a dispositionalist conception of theoretical physics one may either say that physical concepts are translucent or that they are transparent. Combined with (1) and (2) physical transparency leads to a version of the powerful qualities view which, it seems to me, amounts to a form of idealism whereby physical properties just are conscious properties, and *vice versa*. Physical translucency combined with (1) and (2) leads to a form of property dualism – see, e.g., 2.13.

Crucially, none of this has to do, I think, with the purported cognitive independence or inter–dependence of physical and phenomenal concepts. Recall that the cognitive independence between two concepts C_1 and C_2 relative to a certain amount of background knowledge K has been defined as such that a subject mastering both C_1 and C_2 and accepting K as true would not on that basis alone be in a position to rationally judge that C_1 and C_2 are necessarily coextensive. But physicalism, idealism, dualism, panpsychism *et cetera*, are metaphysical theses, not epistemological ones: they concern the nature of certain properties, not the way in which we may get to know that those properties have the nature they have. Likewise, I take it that the reason why revelation threatens physicalism has to do with metaphysics, not with epistemology. In fact, I've argued that revelation would threaten physicalism both under the assumption that physical and phenomenal concepts are cognitively independent and under the assumption that they are not. Let me go more slowly. As I see it, revelation does not threaten physicalism insofar as it claims that we get to know the nature of phenomenal properties *a priori* by deploying phenomenal concepts, or in a way that it is cognitively independent from the deployment of given physical concepts: rather, it threatens physicalism in virtue of what it implies, namely insofar as it implies that phenomenal properties are essentially (partially or entirely) phenomenal, and hence essentially experience–involving. Provided phenomenal properties are taken to be nothing over and above a kind of physical ones, the latter violates the NE principle and, in virtue of the thesis whereby only fundamental physical properties carve nature at its joints, threatens to violate the negative *desideratum* of physicalism, whereby there cannot be fundamental experience–involving entities.

So, as far as I can see given certain very minimal and rather innocent assumptions the debate of the last decades on whether physicalism ought to be our default metaphysical stance in consciousness studies or not rotates around wether revelation is true or false as one of the key–issues if not *the* key–issue. I've started this work by trying to make those minimal background assumptions as clear as I could manage. I've begun by offering an ostensive definition of conscious or phenomenal properties.

Once again, allow me to stress that the fact that I've deployed Nagel's iconic expression to fix reference on phenomenal properties just because I could not come up with any better expression does not imply that I've assumed any theory–laden account of those properties from the start. As long as it is clear that by terms like consciousness, phenomenality, experience, phenomenal properties *et cetera* I *just* mean the most folk–psychologically obvious feature or combination of features which the states that figure in the list of positive examples – see 1.1 – uncontroversially share and those that figure in the list of negative examples most likely lack, I'm fine: people who happen to have problem with Nagel's expression – say, because it is theory–laden, because it conflicts with whatever prior understanding the might have, or for whatever other reason – may call that feature or combination of features however it pleases them.

Offering an innocent characterization of physicalism is less straightforward, for as an old joke goes, if you ask two physicalists what physicalism is you will get three different answers. Still, in virtue of the brief overview of the relevant recent literature offered in 1.2 I believe the vast majority of contemporary physicalists would at least buy into Coleman's (2009) conventional physicalism where, recall, conventional physicalism consists in the combination of two *desiderata*: (i) phenomenal properties metaphysically supervene upon fundamental physical ones and (ii) There are no fundamental experience–involving entities. The case I've been trying to make throughout the work that revelation is incompatible with physicalism so understood may be summarized as follows: assumed (i), revelation would violate (ii). Hence, if both in order for the combination of (i) and (ii) – in which conventional physicalism consists – to be true, revelation must be false.

Regarding physical entities and/or properties I have defined them as those entities/properties that (I) are treated «approximately accurately, by current or future (in the limit of inquiry, ideal) versions of fundamental physics» (Wilson, 2006: 72), (II) are not essentially experience–involving. Moreover, I've been assuming that in order for physicalism to be true at least some of the properties/entities/... physics

deals with should be fundamental. Indeed, it seems reasonable to suppose that on a physicalist worldview only fundamental physical properties carve nature at its joints, where fundamental properties have been defined, rather straightforwardly, as those properties such that, necessarily, they are not instantiated in virtue of anything else, and fundamental entities have been defined as those entities that do not depend on anything else to exist. Again, I think most philosophers that I know of who declare themselves as physicalists without any further specification would at least buy into these minimal *desiderata* for some given entity or property to count as physical.

I've assumed a Finean/non–modal account of essentiality and of essential properties. While it seems clear that if a given property is essential to a certain entity in the Finean sense it is also necessarily possessed by that entity and hence essential to it in the modal sense, it is not uncontroversial whether the opposite is the case. So, I would say that everything I've been arguing for throughout the work is compatible with a Kripkean/modal account of essentiality as well.

The definition of physical concepts I've adopted is very minimal: these are the concepts that are deployed in our best physical theories to characterize physical entities, properties, states and processes. Even though part of this work has been devoted to an investigation of the consequences a dispositionalist conception of theoretical physics might have upon the debate on the metaphysics of consciousness, nothing in my arguments for the incompatibility between revelation and physicalism hinges upon such a conception, nor upon the so called powerful qualities view, whereby qualities and dispositions are not distinct kind of properties, but rather distinct ways of considering what it is for one and the very same kind of properties to be instantiated.

Offering a characterization of phenomenal concepts that could be accepted by the broadest possible audience has been trickier, for the debate on what phenomenal concepts are is extremely tangled. Yet, again, the working definition of phenomenal concepts I've adopted is based on some assumptions which I believe may be accepted by a vast number of readers: (1) people have experiences; (2) people might have

all sorts of thoughts and (true or false) beliefs about their own and other subjects' experiences; (3) whenever a subject has an experience, there is something it is like for that subject to live through time t, the notion of what–it's–like–ness and similar ones having been pointed to ostensively and minimally characterized in 1.1; (4) there is a way of thinking about experiences in terms of the way it is like for someone to have them, which is (arguably) different from other possible ways of thinking about those experiences which could also yeld true or false beliefs about them (e.g., community–relational concepts, neuro–physiological concepts, and so on).

In the second chapter, which is centered on Chalmers' conceivability argument, I've laid down my case that both full–revelation/phenomenal–transparency and partial–revelation/phenomenal–translucency are incompatible with physicalism. After having described the two–dimensional semantic apparatus Chalmers deploys to draw the conceivability argument, I've shown that the the ideal primary negative conceivability of phenomenal zombies is logically equivalent with the cognitive independence of physical and phenomenal concepts, meaning that the two must be either both true or both false. Then, I've made a case that both phenomenal transparency and phenomenal translucency would confute physicalism both under the assumption that physical and phenomenal concepts are cognitively independent, hence phenomenal zombies are conceivable, and under the assumption that they are not, hence phenomenal zombies are inconceivable. So, I hope to have established that it is necessary and most likely necessary and sufficient that revelation is denied in order for physicalists to avoid the threat posed by Chalmers' two–dimensional semantic apparatus. For, as I argued in 2.10, 2.13 and 2.14 and then recapped in 3.1, assumed that every metaphysically possible world which is a minimal exact physical duplicate of the actual world must also be a duplicate of the actual world with respect to consciousness, full–revelation/phenomenal–transparency leads into one of the following views: micro–psychism, panpsychist Russelian monism or a form of idealism. Whereas partial–revelation/phenomenal–transparency, when combined with the claim that every metaphysically possible world which is a minimal exact physical

duplicate of the actual world is also a duplicate of the actual world with respect to consciousness leads to one of the following views: property dualism, pluralism or mysterianism plus, possibly, a panpsychist version of Russelian monism or a form of micro–psychism. Of course, one could reject Chalmers' semantic apparatus in the very first place. But this would be like winning by flipping the table. I would say that in order to rebut an argument one must at the very least assume the language in which the premises of the argument are formulated, for otherwise the argument could not even be formulated, but then there would be nothing to confute. Besides, the only result one would obtain by rejecting Chalmers' version of two–dimensional semantics is to allow for the possibility of there being radically opaque concepts. As I argued in 1.8.4 among other places, physicalism is incompatible with physical concepts being radically opaque, whereas by taking phenomenal concepts to be radically opaque revelation is *eo ipso* denied. More in general, although the thesis of revelation may be expressed in two–dimensional terms, strictly speaking it does not need two–dimensional semantics in order to be formulated. Provided it is accepted that experiences and phenomenal concepts – ostensively pointed to and minimally characterized as in 1.1 and 1.7 – do exist, to get revelation one only needs the notion of essence – understood either in a primitivist/Finean sense or in a Kripkean/modal one – and possibly a roughly post–Fregean account of concepts and the notion of *a priori* knowledge. So, even if Chalmers' two–dimensional semantic apparatus happened to be rejected *in toto*, revelation may still be formulated, and it would still threaten physicalism.

There are striking similarities between Chalmers' conceivability argument and Jackson's knowledge argument, which has been the topic of the third chapter, as well as between Kripke's modal argument, which has also been treated in the third chapter, and the argument from revelation. As I've suggested, there are basically four possible reactions to the thought–experiment on which the knowledge–argument is based, namely (a) there is something Mary learns upon her release, but all the facts about color vision are physical facts; (b) there is something Mary learns

upon her release, and not all the facts about color vision are physical facts; (c) there is nothing Mary learns upon her release, and all the facts about color vision are physical facts; (d) there is nothing Mary learns upon her release, but not all the facts about color vision are physical facts. (b) and (d) are incompatible with physicalism, for according to physicalists all the facts (about color vision) are physical facts. Besides, given the way in which the thought–experiment is formulated (d) seems to me to be self–refuting. (a) and (c) map into the following two possible options: (A) physical and phenomenal concepts are cognitively independent and every metaphysically possible world which is a minimal exact physical duplicate of the actual world is also a duplicate of the actual world with respect to consciousness; (B) phenomenal and physical concepts are cognitively inter–dependent and every metaphysically possible world which is a minimal exact physical duplicate of the actual world is also a duplicate of the actual world with respect to consciousness. Given this, all one has to do is to reiterate the arguments I have offered in the second chapter to the case that both full–revelation/phenomenal–transparency and partial–revelation/phenomenal–translucency confute physicalism both under the assumption that physical and phenomenal concepts are cognitively independent and under the assumption that they are not.

At any rate, throughout the third chapter I've also elaborated on a number of possible sub–versions of the two physicalism–friendly overall possible reactions that there might be had to the thought–experiment of Mary. In particular, I've given a closer look at the following possible strategies to cope with the knowledge argument: The old–facts–new–ways strategy – 3.2 – the phenomenal–concepts–as–demonstratives strategy – 3.3 – the dual carving strategy – 3.4 and 3.5 – the ability hypothesis – 3.6 – and the acquaintance hypothesis – 3.7. Discussing the phenomenal–concepts–as–demonstratives strategy has also allowed me to argue that Kripke's modal argument may also be reduced to the argument from revelation, and in particular to the thesis of revelation itself. As far as I can see, none of the approaches just mentioned is consistent with revelation. Moreover, it seems to me that it is

necessary *and sufficient* that revelation is denied in order for most of these strategies to constitute a genuinely physicalist response to the argument. The only exception may be constituted by the ability hypothesis. In fact, although it is still necessary that physicalist sponsors of the ability hypothesis deny revelation to be safe, it might not be necessary and sufficient, for they are also committed to deny that the kind of knowledge Mary acquires and could only acquire upon undergoing a visual experience of something red is conceptual/propositional[6]. However, I follow Loar (1990/1997), Goff (2017: 70–71) and Chalmers (1996: 129), among several others, in thinking that overall the ability hypothesis is deeply implausible, for there is a strong *prima facie* case that Mary will, indeed, acquire a new piece of conceptual/propositional knowledge, or at least that she will be put in a position to acquire it, and that she could not be put in such a position if not upon undergoing the visual experience of something red. As for dual carving, to my knowledge no uncontroversial case for it being a real possibility has been offered in the relevant literature. At any rate, my arguments for the incompatibility between revelation and physicalism are neutral both with respect to the second premise of the knowledge argument, which states that physical and phenomenal concepts are cognitively independent, and with the third one, which is denied by dual carvers and states that no two concepts can be transparent with respect to the same entity and cognitive independent. So, even if dual carving happened to be shown to be a real possibility, revelation would threaten physicalism anyway. Yet once more, as far as I can see this is because physicalism is a metaphysical thesis: it concerns a metaphysical relation purportedly holding between kinds of properties, not a relation between kinds of concepts. Provided one can manage to show that it may be conceivable that certain dispositions were

[6]One may say that physicalist sponsors of the acquaintance hypothesis need to deny, besides revelation, that coming to be acquainted with a certain entity/property conveys information about that entity/property. For otherwise a version of the knowledge argument may be drawn whose conclusion is that not all information about color vision is physical information. So, after all it might be necessary but not necessary and sufficient for them to deny revelation. However, I guess type–B physicalists may say that there might be non–physical information concerning physical facts. At any rate, I might be contempt with the most modest result of showing, as I attempted to do, that in the case of the acquaintance hypothesis it is at least necessary, though possibly not necessary and sufficient (but I'm not sure about this), that revelation is denied in order for it to constitute a genuinely physicalist response to the argument.

instantiated without certain qualities being instantiated even though there is a relation of identity between the two, which is something I'm quite skeptical about, the powerful qualities view may be deployed to argue for the possibility of dual carving. However, on the assumption that phenomenal properties are at least among the qualities whose essence we get to know via phenomenal concepts, the powerful qualities view leads to full–revelation/phenomenal–transparency which, assumed a dispositionalist conception of theoretical physics, leads to panpsychism and possibly to a form of idealism. So, physicalist sponsors of the powerful qualities view still need to deny revelation.

The acquaintance hypothesis is the last possible strategy to cope with the knowledge argument I have examined. I have offered an "enhanced" version of the argument Anna Giustina (2021) calls the argument from phenomenal–concepts–acquisition. As it is formulated by Giustina, the argument already allows to conclude that being acquainted with a certain phenomenal property *per se* conveys some amount of non–conceptual information about that property. In virtue of how the thought–experiment of Mary is formulated this implies, in turn, that not all information about human color vision is physical information, for by assumption Mary had all the relevant physical information already prior to her release. Combined with either phenomenal translucency or phenomenal transparency as a further premise, the argument of Giustina arguably leads to the conclusion that the (full) nature of phenomenal properties is revealed to subjects pre–conceptually, that is just in virtue of those properties being instantiated. The latter view is most likely inconsistent with physicalism.

4.2 Part two: looking forward

Thus far, I haven't offered any positive argument for revelation. Rather, I've only been reasoning hypothetically: *if* either full–revelation/phenomenal–transparency or partial–revelation/phenomenal–translucency were true, physicalism would be doomed regardless of whether certain other assumptions – most notably, the cognitive

independence of physical and phenomenal concepts (premise 2 of the argument from revelation) and the principle of cognitive inter–dependence of transparent co–referential concepts (premise 3 of the argument from revelation) – are in place or not.

On the assumption that my arguments are correct, the fact that up to now I've only been reasoning hypothetically leaves an obvious strategy available to physicalists: they might simply deny both phenomenal transparency and phenomenal translucency.

The time has now come to confess myself: I believe revelation is true. Hence, I believe conventional physicalism is almost certainly false. The aim of this last section of my work is thus twofold: to suggest that we have a very strong *prima facie* reason to believe revelation is true and to offer *few* remarks on what I suppose is or may be the best metaphysical framework we may have available to accommodate the obvious datum of phenomenal consciousness plus revelation. Let us begin with the first aim.

4.2.1 Revelation: an argument from obviousness?

Along with Goff (2017: 108) I believe the persuasive power of slow and careful reflection on a certain claim or combination of claims is sometimes not sufficiently appreciated, Now, denying revelation amounts to saying that it is not essential to a certain experience – the instantiation of a given phenomenal property – or, equivalently, that it is only contingent to it, that whenever it occurs there's a subject who feels *in precisely that way*. This means that experiences may not be experienced in the way the are experienced and still be the entities they are. Please stop for a moment to slowly and carefully think about what is being said here:

> *It is not essential to the experience of pain that it is painful–for–someone, that it hurts. There might be pains that are not painful. Or, conversely, you might have the worst hurtful sensation right now, and THAT might not be pain.*

Honestly, I cannot think of any other, better way to defend revelation besides saying that for me it is simply obvious: it seems close to be self–evident to me that it

is essential to the experience of pain that it is painful for those who have it, that it is essential to the experience of pleasure that people are pleased when they experience it, that it is essential to the experience of melancholy that people feel melancholic when they have it, and so on. To obtain revelation, one has only to add that people are put in a position to think about the way in which their experiences feel by having them. But even if one could somehow manage to show that we have *no way at all* of thinking about the way in which our own experience feel, and I really cannot see how the latter could be done, as far as I can see the fact that those experiences are essentially phenomenal would still be sufficient *per se* to disprove physicalism. I am sure many readers won't be convinced. Some may even feel disappointed and even annoyed by the fact that I haven't offered any subtle and articulated line of defense. I'm afraid these readers may keep on being disappointed, for as I said I have no better defense of revelation to offer. Sadly, I think there is something to Strawson's (2016: 86) point that hoping to convince anyone who's not already convinced that experiences exist and that it is essential to them that they are precisely as they are experienced may be vain.

It is important to be clear on something, though. The thesis of revelation does *not* imply, *per se*, neither the phenomenon known as *self intimation*, nor the phenomenon known as *infallibility*. Self–intimation claims that whenever a subject instantiate some given phenomenal property or properties, she could not but form the belief that she's instantiating those phenomenal properties, or that she's having the relevant experience(s). Infallibility claims that it is impossible for the subject of an experience to form false beliefs about the phenomenal property or properties she's instantiating, or about how her experience feels. Revelation is perfectly compatible with the negation of both. It is compatible with the fact that some experiences may simply "flow away" without the subjects who have them forming any belief about them, including the belief that they are having them. Likewise, it is perfectly compatible with the possibility of subjects making all sorts of introspective errors. Revelation only claims that having an experience *puts one in a position* to get to

know its (full) nature. Once again, in turn this rests on three fundamental ideas: (1) there is something it is like for subjects to have experiences (2) it is essential to an experience that they are experienced in the way the are experienced – *i.e.*, it is essential to a phenomenal property P that whenever it is instantiated a subject feels thus–and–so; (3) subjects may have beliefs about the way it is like for them to have certain experiences, *of which some may be true and some may be false*. The beliefs about the way it is like for someone to have a certain experience concern the (full) nature of that of the phenomenal property whose instantiation essentially constitutes the experience itself. Among these beliefs, *the true ones and only the true ones* get one to know the (full) nature of the phenomenal property instantiated in – not by – a given experience.

The latter point concludes my discussion of what reasons there might be to believe revelation is true.

4.2.2 What now?

This section will be by far the most speculative one of the entire essay: it raises a question and points to some problems left open by my research, rather than offering any solution to them. That question that will be raised is the following. On the assumption that the arguments I've been drawing throughout the work are correct, the truth of revelation confutes conventional physicalism. As I've just suggested, we may also have a very strong *prima facie* reason to believe revelation is true. What is the best metaphysical framework we should buy into to accommodate the obvious datum of phenomenal consciousness plus revelation, then? My intuitions tell me that a form of idealism may possibly be the best metaphysical framework we have available. According to idealism, every fundamental entity is essentially experience–involving. However, I haven't had the time to slowly and carefully think about most of the points I will touch upon. Besides, an in–depth philosophical analysis of any of those points would require *way more* than a single, concluding section of a single dissertation. It is not my ambition to offer any proper line of

defense for any of the claims I will mention: rather, and way more modestly, I would like to try to suggest that they *may*, *possibly*, be defended, even though, needless to say, for any argument that has ever been offered in philosophy, there has been a counterargument.

There are two claims which, when combined together, as far as I can see would lead to idealism or something close to it: (1) If revelation is true, then some phenomenal properties are fundamental; (2) we'd better believe that the space of fundamental properties is unitary, namely, that there is only one kind of fundamental properties. I strongly believe (1) is true, and a large part of this work has been devoted to argue for it being true. I am less sure about (2). However, I think the probabilities of (2) being true are higher than those of it being false. Hence, my sympathies for idealism.

The reason why I think the probabilities of (2) being true are higher than those of it being false is that its denial would most likely lead to there being brute, inexplicable metaphysical necessities. This point has been made vivid, I think, by David Builes (forth: 5–12), even though several philosophers have independently argued against the existence of brute necessities[7]. Suppose for the sake of argument the space of fundamental qualities, the latter being defined as in 1.3, consisted in a number of "disjunct" kinds of properties, which were not connected in any way. There does not seem to be an intelligible explanation for why *exactly those and only those kinds of fundamental properties* should be the only possible ones. Suppose the space of fundamental qualities was taken to consist of 3 distinct kinds – say, physical properties, phenomenal ones and a further, fundamentally distinct, kind of properties, call them "shmlisical" ones. Why exactly 3 and not 2, or 4, or 57, or infinitely many? No matter what the number of kinds of qualities is posited to be, there does not seem a satisfactory non–arbitrary answer to such a question: it will always seem as if the fact that there is that number and precisely that number of fundamentally distinct kinds of properties is ultimately a brute fact. But as Builes (*ivi:* 6) notes, the way in which metaphysical necessity is typically understood appears to preclude

[7]Besides Builes (forth.), see Chalmers (2002; 2003_{a}) Goff (2019), Van Cleve, (2018) and Strawson (2009), among many others.

its being brute in this manner. Metaphysical necessity is commonly characterized as the most comprehensive form of objective necessity. However, a necessity subjected to arbitrary and *ad hoc* limitations of this sort would not qualify as the broadest objective necessity. One could say that the space of fundamental properties is generated by some kind of underlying mechanism, but, such an explanation does not seem to apply if the disjunct quality–spaces are taken to be fundamentally distinct in kind (Builes, *ivi:* 9). Which leads us to a crucial point. One might want to argue that the assumption that there is only one kind of fundamental properties is ultimately as arbitrary as the assumption that there are 2, or 57, or infinitely many. The fact is, however, that positing an underlying common kind to the space of fundamental qualities seems to be, at least *prima facie*, the only way to make intelligible that those fundamental qualities might be generated by the same sort of underlying mechanism. In other words, in order for the space of fundamental qualities to be intelligibly generated by the same sort of underlying mechanism, and so in order to avoid brute necessities, there has to be a form unity to the space of qualities itself, and a *consequence*, rather than a presupposition, of the latter fact is that there are no fundamentally distinct kinds of fundamental qualities (see Builes, forth: fn. 27).

Provided it is preferable to avoid brute metaphysical necessities, it seems to me that the fact that the truth of revelation leads to some properties being fundamental may lead rather easily, in turn, to a form of idealism.

Now, there are a number of ways in which this, in turn, may be further spelled out. One could buy into a version of the powerful qualities view whereby physical dispositions are identical to (fundamental) phenomenal qualities. Provided one could somehow manage to show that it may be conceivable for physical dispositions to occur in the absence of phenomenal qualities and/or *vice versa* even though there is a relation of identity between the two, this would count as a form of dual carving. On this approach, phenomenal zombies would be conceivable even though not metaphysically possible and there would be something Mary learns upon her

release even though all the facts would be "physico–phenomenal" facts. One that was skeptical as I am about dual carving being a real possibility but still wanted to buy into a version of the powerful qualities view would have to say that phenomenal concepts of qualities and physical concepts of disposition are not cognitively independent, hence phenomenal zombies are neither conceivable nor possible, there is nothing Mary learns upon her release, and all the facts are physico–phenomenal ones. I doubt dual carving could be true. Still, I'm inclined to believe physical and phenomenal concepts are cognitively independent. There's another route which might be taken, though. One could say that physical dispositions are grounded on phenomenal qualities, rather than being identical with the latter. Notice that this approach would still count as idealistic, for since physical dispositions would be grounded on something else, namely phenomenal qualities, by definition they would not be fundamental: only phenomenal qualities would be. At any rate, the cognitive independence of physical and phenomenal concepts would arguably be preserved.

I think there is something to Mørch's (e.g., 2018; 2019) point that it may be inconceivable, and perhaps even ideally inconceivable, that certain phenomenally conscious mental states did not ground certain volitions in those who have them *ceteris absentibus* in virtue and only in virtue of their phenomenal character. Hence my intuition tell me that it might indeed be the case that it is part of the nature of phenomenal qualities that they ground certain dispositions. Still, I'm yet undecided between a version of the powerful qualities view, with or without dual carving, and an approach in which dispositions are taken to be grounded on the fundamental phenomenal qualities rather than being identical with the latter. I think each option has its own issues. I skeptical about dual carving. Denying dual carving but still buying into the powerful qualities view commits one to say that physical and phenomenal concepts are not cognitively independent, and I'm not sure whether I would be willing to agree to the latter. In fact, I've got strong intuitions to the opposite. Taking physical dispositions to be grounded on rather than identical with fundamental phenomenal qualities saves the cognitive independence of physical and

phenomenal concepts but, assumed the form of idealism we're discussing, commits one to say that even though it may be conceivable that physical dispositions might be grounded on something else than phenomenal properties and hence occur in the absence of them, the latter would not be metaphysically possible, for it is assumed that fundamental phenomenal qualities are the only possible kind of fundamental properties. Again, this is something I'm not sure I would be willing to say.

Of course another huge issue – perhaps *the* issue – of idealism and, more in general, of any panpsychist view, has to do with subjects. Provided phenomenal properties are such that necessarily whenever one of them is instantiated a subject S is involved for whom it is like W to live through time t – as per the minimal working definition offered in 1.1 – taking them to be the only possible kind of fundamental properties, and hence taking consciousness to be fundamental and ubiquitous in the material realm, would imply that even the states that figure in the list of negative examples provided in 1.1 and, more in general, any object we observe out there, necessarily involve the instantiation of some fundamental phenomenal property/ies and hence of «subjects, or aggregates of subjects, or part of a wider subject» (Albahari, 2019: 21). Admittedly, the latter conclusion is *very* counterintuitive. Besides, the usual (de–)combination problems that afflict any panpsychist theory arise: how is it the case that subjects like you, me, a cat, a horse either emerge from arrays of micro–subjects or de–merge from a cosmic macro–subject?

I don't have any solution to such a deep problem. But I somehow intuit that a possible way to address it might be to deploy an theory of subjects as "mental lives" or as "sums" or "collections" of experiences. On such a view, subjects do not *have* experiences: they *are* experiences, or "experiential streams". Assumed an account of experiences, in turn, as instantiations of phenomenal properties over spans of time, this view would collapse on a version of the so–called bundle theory, whereby subjects would be nothing over and above bundles of phenomenal properties, possibly held together by a relation of "co–consciousness". So, on this view strictly speaking subjects would not exist: only phenomenal properties would. Hence, there would be nothing

to compose or decompose. Or better, subjects would be nothing over and above the experiences they "have", but they would not constitute a substratum occupying a distinct ontological category with respect to the experience/phenomenal–properties themselves.

The question then arises of what the bearer(s) of phenomenal properties are. Provided experiences are instantiations of phenomenal properties, rather than being themselves bearers of phenomenal properties, and supposed subjects are nothing over and above the experience they "have", what is it that instantiates the phenomenal properties whose instantiation constitutes the experiences in which subjects, in turn, essentially consist in? A possible, tentative reply might be: the universe itself. According to this possible, tentative reply, the universe as a whole is the only fundamental entity (see Schaffer, e.g., $2009_{a,b}$; $2010_{a,b,}$), which at each point in time instantiates an incredibly vast number of phenomenal properties, which in turn ground physical dispositions and by doing so constitute everything that exists and takes place over that moment, including the experiences in which subjects essentially consists in, as well as everything else.

I suppose the old de–composition problem in a new garb might possibly still arise, namely the challenge of explaining how is it that the bundles of suitably arranged phenomenal properties de–merge from a single all–encompassing cosmic subject/object. How is it that the world in which we all live in – a world made of subjects and objects, and perspectives on those objects, and perceptions of them, and thoughts, and feelings, and memories and desires ... – de–merges from a supposedly a–perspectival, non–dual cosmic subject/object?

I have some *very* vague intuitions about this, but no answer to offer at all. All this would be a matter for another work, and most likely for *a number* of other works. I will stop here, for the moment.

Bibliography

[1] R. Adams. Idealism vindicated. In P. Van Inwagen and D. Zimmerman, editors, *Persons: Human and Divine*, pages 35–54. Oxford University Press, Oxford, 2007.

[2] M. Albahari. Perennial idealism: A mystical solution to the mind-body problem. *Philosophers' Imprint*, 19, 2019.

[3] T. Alter and Y. Nagasawa. What is russellian monism? *Journal of Consciousness Studies*, 19(9-10):67–95, 2012.

[4] T. Alter and D. Pereboom. Russellian monism. In E. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2023 Edition.

[5] D. M. Armstrong. *Perception And The Physical World*. Humanities Press, New York, 1961.

[6] D. M. Armstrong. *A Materialist Theory of Mind*. Humanities Press, New York, 1968.

[7] D. M. Armstrong. *A Theory of Universals: Universals and Scientific Realism, Vol II*. Cambridge University Press, New York, 1978.

[8] D. M. Armstrong. *A World of States of Affairs*. Cambridge University Press, Cambridge, 1997.

[9] A. Arnauld and P. Nicole. *Logic or the Art of Thinking*. Cambridge University Press, Cambridge, 1662[1996].

[10] P. Audi. Grounding: Toward a theory of the in-virtue-of relation. *The Journal of Philosophy*, 109(12):685–711, 2012.

[11] K. Balog. Conceivability, possibility, and the mind-body problem. *Philosophical Review*, 108(4):497–528, 1999.

[12] K. Balog. In defense of the phenomenal concept strategy. *Philosophy and Phenomenological Research*, 84(1):1–23, 2011.

[13] K. Balog. Acquaintance and the mind-body problem. In S. Gozzano and C. Hill, editors, *New Perspectives on Type Identity: The Mental and the Physical*, pages 16–43. Cambridge University Press, Cambridge, 2012.

[14] R. Barcan Marcus. Essentialism in modal logic. *Noûs*, 1(1):91–96, 1967.

[15] T. Bayne and M. Montague, editors. *Cognitive Phenomenology*. Oxford University Press, Oxford, 2011.

[16] A. Beghetto. *The Radical Problem of Consciousness: A Critical Clarification of the Illusionist-Realist Debate*. PhD thesis, 2023.

[17] A. Bird. *Nature's Metaphysics: Laws and Properties*. Clarendon Press, Oxford, 2007.

[18] N. Block. On a confusion about a function of consciousness. In *Consciousness, Function, and Representation*. MIT, Cambridge, MA, 1995/2007.

[19] N. Block and R. Stalnaker. Conceptual analysis, dualism, and the explanatory gap. *Philosophical Review*, 108:1–46, 1999.

[20] J. Bolender. An argument for idealism. *Journal of Consciousness Studies*, 8(4):37–61, 2001.

[21] D. Braddon-Mitchell. Qualia and analytical conditionals. *Journal of Philosophy*, 100:111–135, 2003.

[22] D. Braddon-Mitchell and F. Jackson. *Philosophy of Mind and Cognition (2nd edn)*. Blackwell, Oxford, 2007.

[23] F. Brentano. *Psychology from an Empirical Standpoint*. Routledge, 1874[2009].

[24] C.D. Broad. *The Mind and its Place in Nature*. Routledge, London, 1925.

[25] B. Brogaard and J. Salerno. Why counterpossibles are non-trivial. *The Reasoner*, 1.1:5–6, 2007a.

[26] B. Brogaard and J. Salerno. A counterfactual account of essence. *The Reasoner*, 1.4:4–5, 2007b.

[27] B. Brogaard and J. Salerno. Remarks on counterpossibles. *Synthese*, 190:639–60, 2013.

[28] C. D. Brown. A properly physical russellian physicalism. *Journal of Consciousness Studies*, 24(11-12):31–50, 2017.

[29] R. Brown. Deprioritizing the a priori arguments against physicalism. *Journal of Consciousness Studies*, 17(3-4):47–69, 2010.

[30] R. G. Brown and J. Ladyman. *Materialism: A Historical and Philosophical Inquiry*. Routledge, New York, 2019.

[31] D. Builes. Modal idealism. Forthcoming in Oxford Studies in Philosophy of Mind.

[32] D. Builes. The world just is the way it is. *The Monist*, 104(1):1–27, 2021.

[33] A. Byrne. Perception and conceptual content. In E. Sosa and M. Steup, editors, *Contemporary Debates in Epistemology*, page 231–50. Blackwell, Oxford, 2005.

[34] S. Campbell. The conception of a person as a series of mental events. *Philosophy and Phenomenological Research*, 73(2):339–358, 2006.

[35] R. Carnap. *Meaning and Necessity*. University of Chicago Press, 1947.

[36] P. Carruthers. Conscious experience versus conscious thought. In *Consciousness: Essays from a Higher-Order Perspective*. Oxford University Press, Oxford.

[37] P. Carruthers. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge University Press, Cambridge, 2000.

[38] P. Carruthers. Phenomenal concepts and higher-order experiences. *Philosophy and Phenomenological Research*, 68(2):316–36, 2004.

[39] P. Carruthers and B. Veillet. The phenomenal concept strategy. *Journal of Consciousness Studies*, 14(9):212–36, 2007.

[40] D. Chalmers. Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3):200–19, 1995.

[41] D. Chalmers. *The Conscious Mind*. Oxford University Press, Oxford, 1996.

[42] D. Chalmers. Materialism and the metaphysics of modality. *Philosophy and Phenomenological Research*, 59:473–96, 1999.

[43] D. Chalmers. Does conceivability entail possibility? In T. Gendler and J. Hawthorne, editors, *Conceivability and Possibility*. Oxford University Press, Oxford, 2002.

[44] D. Chalmers. The content and epistemology of phenomenal belief. In Q. Smith and A. Jokic, editors, *Consciousness: New Philosophical Perspectives*, pages 1–54. Clarendon Press, Oxford, UK, 2003a.

[45] D. Chalmers. Consciousness and its place in nature. In S. P. Stich and T. A. Warfield, editors, *Blackwell Guide to the Philosophy of Mind*, pages 102–142. Blackwell, 2003b.

[46] D. Chalmers. Epistemic two-dimensional semantics. *Philosophical Studies*, 118(1/2):153–226, 2004.

[47] D. Chalmers. The foundations of two-dimensional semantics. In M.G. Garcia-Carpintero and J. Macià, editors, *Two-Dimensional Semantics*. Oxford University Press, Oxford, 2006a.

[48] D. Chalmers. Strong and weak emergence. In P. Clayton and P. Davies, editors, *The Re-emergence of Emergence*. Oxford University Press, Oxford, 2006b.

[49] D. Chalmers. The two-dimensional argument against materialism. In B. Beckermann, A. Mclaughlin and S. Walter, editors, *The Oxford Handbook of Philosophy of Mind*. Oxford University Press, Oxford, 2009.

[50] D. Chalmers. The two-dimensional argument against materialism. In *The Character of Consciousness*, pages 141–206. Oxford University Press, Oxford, 2010a.

[51] D. Chalmers. Two-dimensional semantics. In *The Character of Consciousness*, pages 541–68. Oxford University Press, Oxford, 2010b.

[52] D. Chalmers. Panpsychism and panprotopsychism. *Amherst Lecture in Philosophy*, (8), 2013.

[53] D. Chalmers. The combination problem for panpsychism. In G. Brüntrup and J. Jaskolla, editors, *Panpsychism: Contemporary Perspectives*, pages 179–214. Oxford University Press, Oxford, 2016.

[54] D. Chalmers. The meta-problem of consciousness. *Journal of Conscious Studies*, 25(9-10):6–61, 2018.

[55] D. Chalmers. Idealism and the mind-body problem. In W. Seager, editor, *The Routledge Handbook of Panpsychism*, pages 353–73. Routledge, New York, 2019.

[56] P. M. Churchland. Eliminative materialism and the propositional attitudes. *Journal of Philosophy*, 78:67–90, 1981.

[57] P.M. Churchland. Reduction, qualia and the direct introspection of brain states. *Journal of Philosophy*, 82:8–28, 1985.

[58] P.M. Churchland. *Matter and Consciousness*. MIT Press, Cambridge, MA, revised edition, 1988.

[59] P.S. Churchland. *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. MIT Press, Cambridge, MA, 1986.

[60] S. Coleman. Mind under matter. In D. Skrbina, editor, *Mind that Abides*. Benjamins Publishing, Amsterdam, 2009.

[61] S. Coleman. The real combination problem: Panpsychism, micro-subjects, and emergence. *Erkenntnis*, 79(1):19–44, 2014.

[62] S. Coleman. Panpsychism and neutral monism: How to make up one's mind. In G. Brüntrup and L. Jaskolla, editors, *Panpsychism: Contemporary Perspectives*, page 518–596. Oxford University Press, New York, 2017.

[63] S. Coleman. Natural acquaintance. In J. Raleigh, T.and Knowles, editor, *Acquaintance: New Essays*, pages 49–74. Oxford University Press, Oxford, 2019.

[64] R. G. Collingwood. On the so-called idea of causation. *Proceedings of the Aristotelian Society*, 38(1):85–112, 1938.

[65] E. Conee. Phenomenal knowledge. *Australasian Journal of Philosophy*, 72:136–50, 1994.

[66] F. Correia. (finean) essence and (priorean) modality. *Dialectica*, 61:63–84, 2007.

[67] B. Cortesi. The thesis of revelation in the philosophy of mind: A guide for the perplexed. *Argumenta*, 2023.

[68] T. Crane. Cosmic hermeneutic vs. emergence. In C. MacDonald and G. MacDonald, editors, *Emergence in Mind*. Oxford University Press, Oxford, 2010.

[69] T. Crane and C. French. The problem of perception. In E. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2021 Edition.

[70] C. Daly. Scepticism about grounding. In F. Correia and B. Schnieder, editors, *Metaphysical Grounding: Understanding the Structure of Reality*, pages 81–100. Cambridge University Press, Cambridge, 2012.

[71] N. Damnjanovic. Revelation and physicalism. *Dialectica*, 66(1):69–91, 2012.

[72] S. Dasgupta. The possibility of physicalism. *The Journal of Philosophy*, 111(9/10):557–92, 2014.

[73] D. Dennett. Why you can't make a computer that feels pain. In *Brainstorms*, page 190–229. MIT Press, Cambridge, MA, 1978.

[74] D. Dennett. *The Intentional Stance*. MIT Press, 1981.

[75] D. Dennett. Quining qualia. In A. Marcel and E. Bisiach, editors, *Consciousness in Contemporary Science*, page 42–77. Oxford University Press, New York, 1988.

[76] D. Dennett. The origins of selves. *Cogito*, 3(3):163–173, 1989.

[77] D. Dennett. *Consciousness Explained*. Little, Brown, New York, 1991.

[78] R. Descartes. *Meditations on First Philosophy*. Cambridge University Press, Cambridge, 1985.

[79] J. L. Dowell. The physical: Empirical, not metaphysical. *Philosophical Studies*, 131(1):25–60, 2006.

[80] F. Dretske. *Naturalizing the Mind*. MIT Press, 1995.

[81] J. W. Dunne. *An Experiment With Time*. A & C Black, London, second edition, 1929.

[82] E. Díaz–León. Do a posteriori physicalists get our phenomenal concepts wrong? *Ratio*, 27(1):1–16, 2013.

[83] E. Díaz–León. Phenomenal concepts: Neither circular nor opaque. *Philosophical Psychology*, 29(8):1186–99, 2016.

[84] A. Eddington. *The Nature of the Physical World*. Cambridge University Press, Cambridge, 1928.

[85] B. Ellis. *The Philosophy of Nature: A Guide to the New Essentialism*. McGill Queens University Press, Montreal, 2002.

[86] A. Elpidorou. A posteriori physicalism and introspection. *Pacific Philosophical Quarterly*, 98(1):474–500, 2015.

[87] A. Elpidorou and G. Dove. *Consciousness and Physicalism: A Defense of a Research Program*. Routledge, New York, 2018.

[88] B. A. Farrell. Experience. *Mind*, 59:170–198, 1950.

[89] H. Feigl. The 'mental' and the 'physical'. *Minnesota Studies in the Philosophy of Science*, 2(2):370–497, 1958.

[90] K. Fine. Essence and modality: The second philosophical perspectives lecture. *Philosophical Perspectives*, 8:1–16, 1994.

[91] K. Fine. Ontological dependence. *Proceedings of the Aristotelian Society*, 95(1):269–90, 1995a.

[92] K. Fine. Senses of essence. In D. Sinnott-Armstrong, W. Raffman and N. Asher, editors, *Modality, Morality and Belief: Essays in Honor of Ruth Barcan Marcus*, pages 53–73. Cambridge University Press, 1995b.

[93] K. Fine. The question of realism. *Philosophers' Imprint*, 1(2):1–30, 2001.

[94] J. Fodor. Special sciences (or: The disunity of science as a working hypothesis). *Synthese*, 28(2):97–115, 1974.

[95] J. Fodor. *The Language of Thought*. Cromwell, New York, 1975.

[96] J. Fodor. *Concepts: Where Cognitive Science Went Wrong*. Oxford University Press, Oxford, 1998.

[97] J. Foster. *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*. Routledge, 1981.

[98] J. Foster. *The Case for Idealism*. Routledge, 1982.

[99] J. Foster. *A World for Us: The Case for Phenomenalistic Idealism*. Oxford University Press, 2008.

[100] K. Frankish. The anti-zombie argument. *Philosophical Quarterly*, 57(229):650–666, 2007.

[101] K. Frankish. Quining diet qualia. *Consciousness and Cognition*, 21(2):667–676, 2012.

[102] K. Frankish. Illusionism as a theory of consciousness. In K. Frankish, editor, *Illusionism as a Theory of Consciousness*, page 13–48. Imprint Academic, Exeter, 2017.

[103] G. Frege. Über sinn und bedeutung. *Translations from the Philosophical Writings of Gottlob Frege*, 1952.

[104] B. Gertler. Introspecting phenomenal states. *Philosophy and Phenomenological Research*, 63(2):305–28, 2001.

[105] B. Gertler. *Self-Knowledge*. Routledge, New York, NY, 2011.

[106] B. Gertler. Renewed acquaintance. In D. Smithies and D. Stoljar, editors, *Introspection and Consciousness*, page 89–123. Oxford University Press, Oxford, 2012.

[107] C. Gillett and B. Loewer, editors. *Physicalism and its Discontents*. Cambridge University Press, Cambridge, 2001.

[108] A. Giustina. Introspection without judgment. *Erkenntnis*, 86:407–427, 2021.

[109] A. Giustina. Introspective knowledge by acquaintance. *Synthese*, 200:128, 2022.

[110] P. Goff. Ghosts and sparse properties: Why physicalists have more to fear from ghosts than zombies. *Philosophy and Phenomenological Research*, 81:119–37, 2010.

[111] P. Goff. A posteriori physicalists get our phenomenal concepts wrong. *Australasian Journal of Philosophy*, 89(2):191–209, 2011.

[112] P. Goff. Real acquaintance and physicalism. In P. Coates and S. Coleman, editors, *Phenomenal Qualities: Sense, Perception and Consciousness*. Oxford University Press, Oxford, 2015.

[113] P. Goff. *Consciousness and Fundamental Reality*. Oxford University Press, New York, 2017.

[114] P. Goff. *Galileo's Error: Foundations for a New Science of Consciousness*. Pantheon Books, New York, 2019.

[115] P. Goff and S. Coleman. Russellian monism. In Uriah Kriegel, editor, *The Oxford Handbook of the Philosophy of Consciousness*. 2020.

[116] T. Goldschmidt and K. Pearce. *Idealism: New Essays in Metaphysics*. Oxford University Press, Oxford, forthcoming.

[117] M. Graziano. *Consciousness and the Social Brain*. Oxford University Press, Oxford, 2013.

[118] M. Graziano. Consciousness engineered. *Journal of Consciousness Studies*, 23(11-12):98–115, 2016.

[119] A. Gupta and S. Mackereth. Definitions. In E. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2023 Edition.

[120] P. Guyer and R-P. Horstmann. Idealism. In E. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2023 Edirion.

[121] B. Hale. *Necessary Beings*. Oxford University Press, Oxford, 2013.

[122] G. Harman. The intrinsic quality of experience. *Philosophical Perspectives*, 4:31–52, 1990.

[123] C. Hartshorne. Causal necessities: An alternative to hume. *Philosophical Review*, 63(4):479–499, 1954.

[124] J. P. Hawthorne. Advice to physicalists. *Philosophical Studies*, 108:17–52, 2002.

[125] J. Heil. *From an Ontological Point of View*. Oxford University Press, New York, 2003. Reprinted 2009.

[126] J. Heil. Powerful qualities. In A. Marmodoro, editor, *The Metaphysics of Powers: Their Grounding and Manifestations*, page 58–72. Routledge, New York, 2010.

[127] T. Hildebrand. Two types of quidditism. *Australasian Journal of Philosophy*, 94(3):516–532, 2016.

[128] C. S. Hill. Imaginability, conceivability, possibility and the mind-body problem. *Philosophical Studies*, 87:61–85, 1997.

[129] C. S. Hill and B. P. McLaughlin. There are fewer things in reality than are dreamt of in chalmers's philosophy. *Philosophy and Phenomenological Research*, 59:446–454, 1999.

[130] T. Hofweber. Ambitious, yet modest, metaphysics. In D. Chalmers, D. Manley, and R. Wasserman, editors, *Metametaphysics: New Essays on the Foundations of Ontology*, page 260–289. Oxford University Press, Oxford, 2009.

[131] E. L. Holman. Phenomenal concepts as bare recognitional concepts: Harder to debunk than you thought, ... but still possible. *Philosophical Studies*, 164(3):807–827, 2013.

[132] T. Horgan. Jackson on physical information and qualia. *Philosophical Quarterly*, 34:147–183, 1984.

[133] T. Horgan. From supervenience to superdupervenience: Meeting the demands of a material world. *Mind*, 102(408):555–586, 1993.

[134] T. Horgan. Introspection and phenomenal consciousness: Running the gamut from infallibility to impotence. In D. Smithies and D. Stoljar, editors, *Introspection and Consciousness*, page 405–422. Oxford University Press, Oxford, 2012.

[135] T. Horgan and J. Tienson. Deconstructing new wave materialism. In C. Gillett and B. Loewer, editors, *Physicalism and Its Discontents*, page 307–318. Cambridge University Press, Cambridge, 2001.

[136] R. J. Howell. *Self-awareness and the Elusive Subject*. Oxford University Press, Oxford, England, 2023.

[137] D. Hume. *A Treatise of Human Nature*. Oxford University Press, Oxford, 1739-40[2000].

[138] N. Humphrey. Redder than red: Illusionism or phenomenal surrealism? *Journal of Consciousness Studies*, 23(11-12):116–123, 2016.

[139] E. Husserl. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. First Book: General Introduction to Pure Phenomenology*. Nijhoff, The Hague, 1913[1983].

[140] E. Husserl. *The Crisis of the European Sciences and Transcendental Phenomenology*. Northwestern University Press, Evanston, IL, 1936[1970].

[141] F. Jackson. Epiphenomenal qualia. *Philosophical Quarterly*, 32:127–136, 1982.

[142] F. Jackson. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford University Press, New York, 1998.

[143] W. James. *Some Problems of Philosophy*. Longmans, Green and Co., New York, 1911[1979].

[144] M. Johnston. How to speak of the colors. *Philosophical Studies*, 68(3):221–263, 1992.

[145] F. Kammerer. The illusion of conscious experience. *Synthese*, 198(1):845–866, 2021.

[146] I. Kant. *The Critique of Pure Reason*. Cambridge University Press, Cambridge, 1781/1787[2009].

[147] B. Kastrup. An ontological solution to the mind–body problem. *Philosophies*, 2(2), 2017.

[148] B. Kastrup. The universe in consciousness. *Journal of Consciousness Studies*, 25(5-6):125–155, 2018.

[149] J. Kim. Multiple realization and the metaphysics of reduction. *Philosophy and Phenomenological Research*, 52(1):1–26, 1992.

[150] J. Kim. Events as property exemplifications. *Supervenience and Mind*, 1993.

[151] J. Kim. *Mind in a Physical World: An Essay on the Mind-Body Problem*. MIT Press, Cambridge, MA, 1998.

[152] J. Kim. *Physicalism, or Something Near Enough*. Princeton University Press, Princeton, 2005.

[153] S. Kripke. *Naming and Necessity*. Blackwell, Oxford, 1980.

[154] T. Kroedel and M. Schulz. Grounding mental causation. *Synthese*, 193(6):1909–1923, 2016.

[155] J. Ladyman. The scientistic stance: The empirical and materialist stances reconciled. *Synthese*, 178(1):87–98, 2011.

[156] J. Ladyman and D. Ross. *Every Thing Must Go: Metaphysics Naturalized*. Clarendon Press, Oxford, 2007.

[157] G. W. Leibniz. *New Essays on Human Understanding*. Cambridge University Press, Cambridge, 1765[1981].

[158] G. W. Leibniz. *Philosophical Texts*. Oxford University Press, Oxford, 1998.

[159] J. Levin. Nagel vs. nagel on the nature of phenomenal concepts. *Ratio*, 20.3:293–307, 2007a.

[160] J. Levin. What is a phenomenal concept? In T. Alter and S. Walter, editors, *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, page 87–110. Oxford University Press, 2007b.

[161] J. Levin. Once more unto the breach: Type b physicalism, phenomenal concepts, and the epistemic gap. *Australasian Journal of Philosophy*, 97(1):57–71, 2019.

[162] J. Levine. Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, 64:354–361, 1983.

[163] J. Levine. *Purple Haze: The Puzzle of Consciousness*. Oxford University Press, Oxford, 2001.

[164] J. Levine and K. Trogdon. The modal status of materialism. *Philosophical Studies*, 145(3):351–362, 2009.

[165] D. Lewis. An argument for the identity theory. *Journal of Philosophy*, 63(1):17–25, 1966.

[166] D. Lewis. Art, mind, and religion. *The Journal of Philosophy*, 66(1):22–27, 1969.

[167] D. Lewis. Counterparts of persons and their bodies. *The Journal of Philosophy*, 68(7):203–211, 1971.

[168] D. Lewis. Survival and identity. In A. O. Rorty, editor, *The Identity of Persons*, page 17–40. University of California Press, 1976.

[169] D. Lewis. Postscripts to "survival and identity". In *Philosophical Papers, Volume I*, page 73–77. Oxford University Press, 1983a.

[170] D. Lewis. New work for a theory of universals. *Australasian Journal of Philosophy*, 61(4):343–377, 1983b.

[171] D. Lewis. Postscript to 'mad pain and martian pain'. In *Philosophical Papers (Volume 1)*, page 130–132. Oxford University Press, 1983c.

[172] D. Lewis. What experience teaches. *Proceedings of the Russellian Society*, 13:29–57, 1988.

[173] D. Lewis. Reduction of mind. In S. Guttenplan, editor, *Companion to the Philosophy of Mind*, page 412–431. Blackwell, 1994.

[174] D. Lewis. Should a materialist believe in qualia? *Australasian Journal of Philosophy*, 73(1):140–144, 1995.

[175] M. Liu. Revelation and the appearance/reality distinction. Forthcoming in Oxford Studies in Philosophy of Mind.

[176] M. Liu. Explaining the intuition of revelation. *Journal of Consciousness Studies*, 27(5-6):99–107, 2020.

[177] M. Liu. Revelation and the intuition of dualism. *Synthese*, 199(3-4):11491–11515, 2021.

[178] B. Loar. Phenomenal states. In J. Tomberlin, editor, *Philosophical Perspectives 4: Action Theory and Philosophy of Mind*, page 81–108. Ridgeview, 1990. Reprinted in substantially revised form in Block, N. Flanagan, O. and Güzeldere

(G.) (Eds.) The Nature of Consciousness: Philosophical Debates. Cambridge, MA: MIT Press, 1997, 597–616. Page references to the 1997 version.

[179] B. Loar. Phenomenal states (second version). In D. Chalmers, editor, *Philosophy of Mind: Classical and Contemporary Readings*, page 295–311. Oxford University Press, 2002.

[180] Brian Loar. David chalmers's the conscious mind. *Philosophy and Phenomenological Research*, 59(2):465–472, 1999.

[181] J. Locke. *An Essay Concerning Human Understanding*. 1689[1975].

[182] Michael Lockwood. *Mind, Brain and the Quantum: The Compound "I"*. Oxford University Press, 1989.

[183] M. Loux and T. Crisp. *Metaphysics: A Contemporary Introduction*. Routledge, 4th edition, 2017.

[184] E. J. Lowe. What is the source of our knowledge of modal truth? *Mind*, CXXI(484):919–950, 2012.

[185] W. G. Lycan. *Consciousness and Experience*. MIT Press, 1996.

[186] W. G. Lycan. Giving dualism its due. *Australasian Journal of Philosophy*, 87(4):551–563, 2009.

[187] W.G. Lycan. A limited defense of phenomenal information. In T. Metzinger, editor, *Conscious Experience*. Schoningh, 1995.

[188] E. Mach. *Popular Scientific Lectures*. The Open Court Publishing Company, 1897.

[189] R. Majeed. Ramseyan humility: The response from revelation and panpsychism. *Canadian Journal of Philosophy*, 47(1):75–96, 2017.

[190] C. B. Martin. *The Mind in Nature*. Oxford University Press, 2007.

[191] C. B. Martin and J. Heil. Rules and powers. *Noûs*, 32(S12):283–12, 1998.

[192] C. B. Martin and J. Heil. The ontological turn. *Midwest Studies in Philosophy*, 23:34–60, 1999.

[193] M. G. G. Martin. The transparency of experience. *Mind and Language*, 17:376–425, 2002.

[194] F. Mathews. *For Love of Matter: A Contemporary Panpsychism*. State University of New York Press, 2003.

[195] G. Maxwell. Rigid designators and mind-brain identity. *Minnesota Studies in the Philosophy of Science*, 9:9, 1979.

[196] B. McLaughlin. The place of color in nature. page 475–502, 2003.

[197] B. McLaughlin. A priori versus a posteriori physicalism. 2005.

[198] C. McMullen. 'knowing what it's like' and the essential indexical. *Philosophical Studies*, 48:211–233, 1985.

[199] P. E. Meehl. The compleat autocerebroscopist: A thought-experiment on professor feigl's mind-body identity thesis. In P. K. Feyerabend and G. Maxwell, editors, *Mind, Matter, and Method: Essays in Philosophy and Science in Honor of Herbert Feigl*. University of Minnesota Press, 1966.

[200] U. Meixner. Idealism and panpsychism. In Godehard Brüntrup and Jörg Jaskolla, editors, *Panpsychism: Contemporary Perspectives*. Oxford University Press, Oxford, 2017.

[201] A. Melnyk. *A Physicalist Manifesto: Thoroughly Modern Materialism*. Cambridge University Press, 2003.

[202] A. Melnyk. Realization and the formulation of physicalism. *Philosophical Studies*, 131(1):127–155, 2006.

[203] A. Melnyk. In defense of a realization formulation of physicalism. *Topoi*, 37(3):483–493, 2018.

[204] F. Melvin. Intensional logic. In E. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Winter 2022 Edition.

[205] A. Mendelovici. *The Phenomenal Basis of Intentionality*. Oxford University Press, 2018.

[206] J. S. Mill. *A System of Logic: Ratiocinative and Inductive*. University of Toronto Press, 1843[1974].

[207] G. Molnar. *Powers: A Study in Metaphysics*. Oxford University Press, 2003.

[208] B. G. Montero. Post-physicalism. *Journal of Consciousness Studies*, 8(2):61–80, 2001.

[209] B. G. Montero. Must physicalism imply the supervenience of the mental on the physical? *The Journal of Philosophy*, 110(2):93–110, 2013.

[210] B. G. Montero and C. D. Brown. Making room for a this-worldly physicalism. *Topoi*, 37(3):523–532, 2018.

[211] B. G. Montero and D. Papineau. Naturalism and physicalism. In K. J. Clark, editor, *The Blackwell Companion to Naturalism*, page 182–195. Wiley, 2016.

[212] R. Moran. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton University Press, 2001.

[213] H. H. Mørch. Does dispositionalism entail panpsychism? *Topoi*, 39(5):1073–1088, 2018.

[214] K. Morris. Supervenience physicalism, emergentism, and the polluted supervenience base. *Erkenntnis*, 79(2):351–365, 2014.

[215] S. Mumford. *Laws in Nature*. Routledge, 2004.

[216] T. Nagel. What is it like to be a bat? *Philosophical Review*, 83:435–450, October 1974.

[217] N. Nelkin. Propositional attitudes and consciousness. *Philosophy and Phenomenological Research*, 49(3):413–430, 1989.

[218] L. Nemirow. Review of thomas nagel, mortal questions. *Philosophical Review*, 89:473–477, 1980.

[219] L. Nemirow. Physicalism and the cognitive role of acquaintance. In W.G. Lycan, editor, *Mind and Cognition*, page 490–499. Blackwell, 1990.

[220] A. Ney. Physicalism as an attitude. *Philosophical Studies*, 138(1):1–15, 2008.

[221] C.C.-W. Ng. Is mind–body dualism compatible with modern psychiatry? *BJPsych Advances*, 28(2):132–134, 2021.

[222] M. Nida-Rümelin. Experiencing subjects. Manuscript.

[223] M. Nida-Rümelin. What mary couldn't know: Belief about phenomenal states. In T. Metzinger, editor, *Conscious Experience*, pages 219–41. Ferdinand Schoningh, 1995.

[224] M. Nida-Rümelin. Is the naturalization of qualitative experience possible or sensible? In M. Carrier and P. K. Machamer, editors, *Mindscapes: Philosophy, Science, and the Mind*. Pittsburgh University Press, 1997.

[225] M. Nida-Rümelin. Grasping phenomenal properties. In T. Alter and S. Walter, editors, *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, page 307–338. Oxford University Press, 2007.

[226] M. Nida-Rümelin. The illusion of illusionism. *Journal of Consciousness Studies*, 23(11-12):160–171, 2016.

[227] M. Nida-Rümelin. The experience property framework: A misleading paradigm. *Synthese*, 195(8):3361–3387, 2018.

[228] M. Nida-Rümelin and D. O Conaill. Qualia: The knowledge argument. In E. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2023 Edition.

[229] P. Nordhoff. Emergent causation and property causation. In C. Macdonald and G. F. Macdonald, editors, *Emergence in Mind*, page 69–99. Oxford University Press, 2010.

[230] D. O'Conaill. The argument from revelation and the metaphysics of properties. Manuscript.

[231] D. O'Conaill. Grounding, physicalism and necessity. *Inquiry*, 61(7):713–730, 2018.

[232] J. Pallagrosi and B. Cortesi. The stalemate between causal and constitutive accounts of introspective knowledge by acquaintance. Forthcoming in Argumenta.

[233] D. Papineau. *Philosophical Naturalism*. Blackwell, 1993.

[234] D. Papineau. *Thinking about Consciousness*. Oxford University Press, 2002.

[235] D. Papineau. Phenomenal concepts and the private language argument. *American Philosophical Quarterly*, 48(2):175, 2011.

[236] M. Pelczar. *Sensorama: A Phenomenalist Analysis of Spacetime and Its Contents*. Oxford University Press, 2015.

[237] D. Pereboom. *Consciousness and the Prospects of Physicalism*. Oxford University Press, 2011.

[238] D. Pereboom. Illusionism and anti-functionalism about phenomenal consciousness. In K. Frankish, editor, *Illusionism as a Theory of Consciousness*, page 215–231. Imprint Academic, 2017.

[239] J. Perry. *Knowledge, Possibility, and Consciousness*. MIT Press, 2001.

[240] D. Pitt. The phenomenology of cognition or "what is it like to think that p?". *Philosophy and Phenomenological Research*, 69(1):1–36, 2004.

[241] U. T. Place. Is consciousness a brain process? *British Journal of Psychology*, 47(1):44–50, 1956.

[242] H. Putnam. Minds and machines. In Sidney Hook, editor, *Dimensions of Minds*, page 138–164. New York University Press, 1960.

[243] H. Putnam. Psychological predicates. In W. H. Capitan and D. D. Merrill, editors, *Art, Mind and Religion*, page 37–48. University of Pittsburgh Press, 1967.

[244] H. Putnam. The meaning of 'meaning'. In *Philosophical Papers, volume 2: Mind, Language and Reality*, page 215–74. Cambridge University Press, 1975.

[245] A. Quinton. The soul. *Journal of Philosophy*, 59(15):393–409, 1962.

[246] G. Rey. Toward a projectivist account of conscious experience. In Thomas Metzinger, editor, *Conscious Experience*, pages 123–142. Ferdinand Schoningh, 1995.

[247] T. Robertson Ishii and P. Atkins. Essential vs. accidental properties. In E. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2023 Edition.

[248] H. Robinson. *Matter and Sense: A Critique of Contemporary Materialism*. Cambridge University Press, 1982.

[249] H. Robinson. *Perception and Idealism: An Essay on How the World Manifests Itself to Us, and How It (Probably) Is in Itself*. Oxford University Press, 2022.

[250] G. Rosen. Metaphysical dependence: Grounding and reduction. In B. Hale and A. Hoffman, editors, *Modality: Metaphysics, Logic, and Epistemology*, page 109–136. Oxford University Press, 2010.

[251] G. H. Rosenberg. *A Place for Consciousness*. Oxford University Press, 2004.

[252] C. Rovane. *The Bounds of Agency: An Essay in Revisionary Metaphysics*. Princeton University Press, 1997.

[253] B. Russell. Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society*, 11:108–128, 1910.

[254] B. Russell. *The Problems of Philosophy*. Oxford University Press, 1912.

[255] B. Russell. *The Analysis of Matter*. Kegan Paul, 1927.

[256] Gilbert Ryle. *The Concept of Mind: 60th Anniversary Edition*. Hutchinson & Co., 1949.

[257] J. Schaffer. On what grounds what. In D. Chalmers, D. Manley, and R. Wasserman, editors, *Metametaphysics: New Essays on the Foundations of Ontology*, pages 347–383. Oxford University Press, 2009a.

[258] J. Schaffer. Spacetime the one substance. *Philosophical Studies*, 145(1):131–148, 2009b.

[259] J. Schaffer. The internal relatedness of all things. *Mind*, 119:341–376, 2010a.

[260] J. Schaffer. Monism: The priority of the whole. *Philosophical Review*, 119(1):31–76, 2010b.

[261] A. Schopenhauer. *The World as Will and Representation*. Dover, New York, 1966.

[262] E. Schwitzgebel. The unreliability of naïve introspection. *The Philosophical Review*, 117(2):245–273, 2008.

[263] E. Schwitzgebel. *Perplexities of Consciousness*. MIT Press, 2011.

[264] E. Schwitzgebel. Phenomenal consciousness, defined and defended as innocently as i can manage. *Journal of Consciousness Studies*, 23(11-12):224–235, 2016.

[265] E. Schwitzgebel. Introspection. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Winter edition, 2019.

[266] Eric Schwitzgebel. Do you have constant tactile experience of your feet in your shoes? or is experience limited to what's in attention? *Journal of Consciousness Studies*, 14(3):5–35, 2007.

[267] W. Seager. The 'intrinsic nature' argument for panpsychism. *Journal of Consciousness Studies*, 13(10–11):129–145, 2006.

[268] J. R. Searle. *The Rediscovery of the Mind*. MIT Press, 1992.

[269] D. Shabasson. Illusionism about phenomenal consciousness: Explaining the illusion. *Review of Philosophy and Psychology*, 13(2):427–453, 2022.

[270] S. Shoemaker. Causality and properties. In P. Van Inwagen, editor, *Time and Cause: Essays Presented to Richard Taylor*. Reidel, 1980.

[271] S. Shoemaker. Realization and mental causation. In C. Gillett and B. Loewer, editors, *Physicalism and Its Discontents*, pages 74–98. Cambridge University Press, 2001.

[272] S. Shoemaker. *Physical Realization*. Oxford University Press, 2007.

[273] T. Sider. *Writing the Book of the World*. Oxford University Press, 2011.

[274] C. Siewert. Phenomenal thought. In Tim Bayne and Michelle Montague, editors, *Cognitive Phenomenology*, pages 236–267. Oxford University Press, 2011.

[275] D. Skrbina, editor. *Mind that Abides: Panpsychism in the New Millennium*. John Benjamins, 2009.

[276] J. J. Smart. Sensations and brain processes. *The Philosophical Review*, 68(2):141–156, 1959.

[277] A. D. Smith. *The Problem of Perception*. Harvard University Press, 2002.

[278] M. Soteriou. *The Mind's Construction: The Ontology of Mind and Mental Action*. Oxford University Press, 2013.

[279] J. Speaks. Transparency, intentionalism, and the nature of perceptual content. *Philosophy and Phenomenological Research*, 79:539–573, 2009.

[280] T. L. S. Sprigge. *The Vindication Of Absolute Idealism*. Edinburgh University Press, 1983.

[281] T. L. S. Sprigge and A. Montefiore. Final causes. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 45:149–192, 1971.

[282] D. Spurrett and D. Papineau. A note on the completeness of 'physics'. *Analysis*, 59(1):25–29, 1999.

[283] R. Stalnaker. *Our Knowledge of the Internal World*. Oxford University Press, New York, 2008.

[284] D. Stoljar. The argument from diaphanousness. In M. Ezcurdia, R. Stainton, and C. Viger, editors, *New Essays in the Philosophy of Language and Mind*, Canadian Journal of Philosophy (Supplementary Volume), page 341–390. University of Calgary Press, Calgary, 2004.

[285] D. Stoljar. Physicalism and phenomenal concepts. *Mind and Language*, 20(5):469–494, 2005.

[286] D. Stoljar. *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness*. Oxford University Press, Oxford, 2006.

[287] D. Stoljar. The argument from revelation. In D. Braddon-Mitchell and R. Nola, editors, *Conceptual Analysis and Philosophical Naturalism*, pages 113–138. MIT Press, Cambridge, 2009.

[288] D. Stoljar. *Physicalism*. Routledge, New York, 2010.

[289] D. Stoljar. Physicalism. In E. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Summer 2023 Edition.

[290] D. Stoljar and Y. Nagasawa. Introduction. In P. Ludlow, Y. Nagasawa, and D. Stoljar, editors, *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*, page 1–36. MIT Press, 2004.

[291] G. Strawson. Red and 'red'. *Synthese*, 78:193–232, February 1989.

[292] G. Strawson. *Mental Reality*. MIT Press, Cambridge, Mass., 1994.

[293] G. Strawson. Real materialism. In L. M. Antony, editor, *Chomsky and His Critics*, pages 49–88. Blackwell, Oxford, UK, 2003.

[294] G. Strawson. Realistic monism: Why physicalism entails panpsychism. In *Real Materialism*, page 53–74. Oxford University Press, 2008a.

[295] G. Strawson. The identity of the qualitative and the dispositional. *Analysis*, 88(4):271–282, 2008b.

[296] G. Strawson. The self. In B. McLaughlin, A. Beckermann, and S. Walter, editors, *The Oxford Handbook of Philosophy of Mind*. Oxford University Press, 2009.

[297] G. Strawson. Consciousness isn't a mystery. it's matter. *The New York Times*, May 16 2016.

[298] G. Strawson. Identity metaphysics. *The Monist*, 104:60–90, 2021.

[299] P. Sundström. Phenomenal concepts. *Philosophy Compass*, 6(4):267–281, 2011.

[300] R. Swinburne. *Mind, Brain, and Free Will*. Oxford University Press, Oxford, 2013.

[301] T. Tahko. Fundamentality. In E. Zalta and U. Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Winter 2023 Edition.

[302] J. H. Taylor. Physicalism and phenomenal concepts: Bringing ontology and philosophy of mind together. *Philosophia*, 41(4):1283–1297, 2013.

[303] J. H. Taylor. The relation between subjects and their conscious experiences. *Philosophical Studies*, 177(11):3493–3507, 2020.

[304] P. Teller. Subjectivity and knowing what it's like. In A. Beckermann, H. Flohr, and J. Kim, editors, *Emergence or Reduction: Essays on the Prospects of Nonreductive Physicalism*, page 180–200. Berlin, 1992.

[305] M. Thau. *Consciousness and Cognition*. Oxford University Press, Oxford, 2002.

[306] A. Tomasetta. Analytic phenomenology: A guided tour. Forthcoming on Argumenta.

[307] A. Tomasetta. *Coscienza E modalità: L'argomento bi-dimensionale contro il materialismo di David Chalmers*. Aracne, Roma, 2012.

[308] A. Tomasetta. Physicalist naturalism in the philosophy of mind. *Discipline Filosofiche*, 25(1):89–111, 2015a.

[309] A. Tomasetta. *Persone Umane. Teorie Contemporanee in Metafisica Analitica*. Carocci Editore, Roma, 2015b.

[310] K. Trogdon. Revelation and physicalism. *Synthese*, 194(7):2345–66, 2017.

[311] M. Tye. The subjective qualities of experience. *Mind*, 95:1–17, 1986.

[312] M. Tye. Visual qualia and visual content. In T. Crane, editor, *The Contents of Experience*, pages 158–176. Cambridge University Press, 1992.

[313] M. Tye. *Ten Problems of Consciousness*. MIT Press, 1995.

[314] M. Tye. *Consciousness, Color, and Content*. MIT Press, 2000.

[315] M. Tye. A theory of phenomenal concepts. *Royal Institute of Philosophy Supplement*, 53:91–105, 2003.

[316] J. Van Cleve. Mind-dust or magic? panpsychism versus emergence. *Philosophical Perspectives*, 4:215–26, 1990.

[317] B. Van Fraassen. Structure: Its shadow and substance. *British Journal for the Philosophy of Science*, 57:275–307, 2006.

[318] W. R. Webster. Human zombies are metaphysically impossible. *Synthese*, 151:297–310, 2006.

[319] D.M. Wegner. *The Illusion of Conscious Will*. MIT Press, Cambridge, MA, 2018.

[320] A. N. Whitehead. *Adventures of Ideas*. The Free Press, New York, 1967.

[321] J. Wilson. Supervenience-based formulations of physicalism. *Noûs*, 39(3):426–459, 2005.

[322] J. Wilson. On characterizing the physical. *Philosophical Studies*, 131(1):61–99, 2006.

[323] J. Wilson. No work for a theory of grounding. *Inquiry*, 57(5-6):535–79, 2014.

[324] Jessica Wilson. The unity and priority arguments for grounding. In K. Aizawa and C. Gillet, editors, *Scientific Composition and Metaphysical Ground*, pages 171–204. Palgrave MacMillan, 2016.

[325] D. G. Witmer. Sufficiency claims and physicalism: A formulation. pages 57–73, 2001.

[326] S. Worley. Physicalism and the via negativa. *Philosophical Studies*, 131(1):101–126, 2006.

[327] S. Yablo. Is conceivability a guide to possibility? *Philosophy and Phenomenological Research*, 53:1–42, 1993.

[328] H. Yetter-Chappell. Idealism without god. Forthcoming in Goldschmidt and Pearce.

[329] E. Zalta. Essence and modality. *Mind*, 115:659–93, 2006.

[330] G. Zanotti. Sul perché l'argomento naturalista non può fornire ragioni a sostegno del fisicalismo. *Sistemi intelligenti*, 32(3):575–96, 2020.

[331] G. Zanotti. Physicalism and the burden of parsimony. *Synthese*, 199(3-4):11109–11132, 2021.

[332] G. Zanotti. Consciousness, neuroscience, and physicalism: Pessimism about optimistic induction. *Acta Analytica*, 2022.

[333] G. Zanotti. *Why Physicalism? An Assessment of the Main Arguments for Physicalism about the Mind*. PhD thesis, 2022.

Ringraziamenti

I lettori inglesi mi perdoneranno: desideravo in questa sezione essere il più spontaneo possibile, e ho pensato che nonostante gli anni di esercizio esprimermi in una lingua che non è la mia mi avrebbe reso inevitabilmente artificiale. I lettori italiani, invece, perdoneranno la stucchevolezza che in questi casi è sempre dietro l'angolo.

Ho trascorso 10 anni a Pavia (con alcuni intermezzi). Un profondo affetto e ricordi indelebili mi legano a quella città. Questo lavoro rappresenta il compimento di questo percorso, bellissimo e intenso. Gli ultimi 4 anni sono stati belli e intensi a loro modo. Se lo sono stati, è grazie alle persone che qui andrò elencando in ordine sparso. Non avevo bisogno di ulteriori motivi per essere malinconico all'idea che una fase si stesse per chiudere, e nuove fasi si stessero aprendo: voi quei motivi me li avete dati lo stesso.

La prima persona che desidero ringraziare è il Professor Michele Di Francesco. Il Professor Di Francesco è un filosofo acuto, il cui operato è destinato a lasciare un segno profondo nell'Università italiana: molti ricercatori in filosofia devono molto al suo esempio. Ma poi, Michele è una persona mite e di buone maniere, capace di mettere a proprio agio qualunque interlocutore e tra gli artefici, all'interno del gruppo di filosofia della mente dello IUSS (e non solo), di un clima di amicizia e collaborazione. Sono doti, queste ultime, troppo spesso sottovalutate nel mondo dell'Accademia.

Sono orgoglioso di essere stato un membro della Scuola IUSS: a chi lo ha reso possibile – la Professoressa Bambini, il Professor Sereni, e tutti gli altri docenti e collaboratori della scuola – va la mia riconoscenza.

Un ringraziamento affettuoso va poi al mio relatore, Alfredo Tomasetta. Alfredo è un ragionatore implacabile, e una persona dalla cultura sconfinata. Prima di intraprendere una conversazione con Alfredo, quale che sia l'argomento, è bene prendere un profondo respiro, perché è raro discorrere con lui e non uscirne arricchiti, o stimolati, o con in testa pensieri nuovi. Idealmente, questo lavoro vorrebbe raccogliere l'eredità del suo lavoro sull'argomento bi–dimensionale di Chalmers, che ne è stata una delle principali fonti di ispirazione. Non sta a me dire se sono riuscito in questo intento. Sia come sia, mi piacerebbe che il nostro rapporto di amicizia non si interrompesse.

Oltre a Michele ed Alfredo, l'altra artefice del clima di amicizia e collaborazione di cui sopra è Giulia Piredda: di questo la ringrazio di cuore.

Ho avuto l'opportunità di trascorrere diversi mesi presso l'università di Friburgo. Qui ho avuto l'onore di incontrare Martine Nida–Rümelin, una delle filosofe più influenti della sua generazione: la sua intelligenza non smetterà mai di sorprendermi, e non può che suscitare ammirazione reverenziale. Discutere di filosofia con Martine, sentirla presentare le proprie idee, o leggere i suoi lavori è come vedere un grande pianista esibirsi. L'argomento filosofico che è il tema centrale di questa tesi viene presentato per la prima volta in un articolo di Martine del 2007. Assieme al lavoro di Alfredo sull'argomento bi–dimensionale, le idee di Martine rappresentano quindi l'altra principale fonte di ispirazione di questo mio lavoro. E poi, Martine è una persona dalle doti umane rare: si è impegnata per accogliermi dal primo giorno che ho messo piede a Friburgo.

Un'altra persona che sono onorato di avere incontrato è il Professor Gianfranco Soldati, un uomo dall'intelletto raffinato e penetrante e dall'incredibile vastità di interessi filosofici. Una delle prime cose che si percepiscono nell'incontrarlo, è il piacere che ha nel circondarsi di noi giovani ricercatori. Confrontarmi con lui su alcuni aspetti della mia ricerca e ricevere da lui delle critiche severe ma puntuali e costruttive, mi ha aiutato a rendere più precise certe mie idee.

Giacomo Zanotti e Marco Facchin sono stati dei catalizzatori di sindrome

dell'impostore, ma proprio per questo avere a che fare con loro mi ha aiutato spesso ad uscire dalla mia letargia, a cambiare, ad impegnarmi, a migliorarmi. Condividere con loro una parte del mio percorso e avere i loro consigli ed il loro esempio, poi, mi ha permesso di fare le scelte giuste in momenti delicati.

D'altro canto, avere Arianna Beghetto come amica è stato ed è importante perché profondamente rassicurante. Anche grazie a lei ho capito di non essere solo, e che le mie ansie, le mie insicurezze, le mie preoccupazioni, i miei dubbi, erano le ansie, le insicurezze, le preoccupazioni e i dubbi di una generazione di giovani ricercatori, e forse di una generazione in generale. Questo lo intuivo già, ma avere una conferma mi è servito più di quanto credessi. Arianna è una filosofa estremamente capace: le auguro un brillante futuro, e non ho alcun dubbio che lo avrà.

L'incontro con Jacopo Pallagrosi è stato per me forse l'incontro fondamentale degli ultimi 4 anni. Non esito a dire che questo lavoro semplicemente *non esisterebbe*, se non fosse stato per Jacopo. Jacopo mi ha incontrato in un periodo di profondo scoraggiamento. Con la sua inesauribile curiosità, ha iniziato a pungolarmi. Da allora, lo fa ogni giorno. Pungolandomi costantemente, Jacopo mi spinge a credere nelle mie capacità, quali che siano (alquanto modeste, suppongo), e a pensare che, forse, di fare filosofia in fondo ne vale la pena. Con la sua ironia, Jacopo mi ricorda, come altri hanno fatto prima di lui, che nella vita non bisogna mai prendersi troppo sul serio. Oltre ad essere, assieme ad Alfredo, una delle persone più colte che io conosca, Jacopo è una persona eccezionalmente brillante, di ammirevole onestà intellettuale, e di straordinaria apertura mentale: qualunque idea abbiate, non importa quanto granitica sia la vostra convinzione, discutere con Jacopo vi farà capire che si può tornare sui propri passi, considerare il punto di vista altrui senza pregiudizi, avere dubbi. Ma soprattutto, Jacopo è un amico: una persona che è lì per te quando serve.

Se Jacopo è stato l'incontro fondamentale degli ultimi 4 anni, Simone è stato senza dubbio alcuno l'incontro fondamentale degli ultimi 10. Dal giorno in cui ho messo piede a Pavia per la prima volta, Simone è stato per me di tutto: una persona dall'infinita sensibilità, un pungolo costante, un appoggio, un esempio e una

fonte di preziosi consigli, un confidente che mi conosce come poche altre persone mi conoscono. Con lui condivido alcuni dei ricordi più divertenti che posseggo e custodisco gelosamente, ma ho condiviso anche dei momenti non facili. Per quanto io cambiassi, e per quanto le cose intorno a me cambiassero, Simone era lì, e sapere che era ed è lì, è stato continua ad essere un conforto.

Ai ragazzi dello IUSS voglio dire che mi hanno aiutato a sentirmi di nuovo a casa in un momento in cui molti dei miei amici più cari avevano preso altre strade, e Pavia mi sembrava grigia e triste: Carlotta, Veronica, Sara, Sara, Andrea, Achille, Marta, Matteo, Davide, e tutti gli altri dottorandi e borsisti.

Gli amici storici del Collegio sono il motivo per cui Pavia *è* casa. In troppi andrebbero ringraziati. Chi c'è stato sa e capirà. Qui non posso esimermi dal nominare alcuni, però, che negli ultimi 4 anni hanno contato, o meglio, che hanno continuato a contare: Giorgia, Pancu (due punti di riferimento imprescindibili), Ispra (un altro punto di riferimento, sempre nel periodo in cui Pavia mi pareva grigia e triste), Zerbone (una delle persone più generose che io conosca), Falcone (non avrei potuto desiderare coinquilino migliore), Vittorio, Alice, Prisca, Edoardo, Maria e Violante (assieme a loro e grazie a loro ho trascorso uno degli anni più belli della mia vita), Fra, Caterina, Asterix e Matilde.

I ragazzi del gruppo di Friburgo – Davide, Davide, Sharon e gli altri – hanno acceso la mia curiosità: nonostante non avessi bisogno di ulteriori conferme, mi hanno ricordato che le cose che ancora ho da imparare sono infinite. La loro compagnia è stata preziosa in un momento in cui sarebbe stato molto facile sentire di non appartenere a niente, e da loro ho imparato molto su cosa significhi stare in Accademia oggi. In Davide Dalla Rosa ho trovato un amico leale e affidabile.

Infine, desidero ringraziare la mia famiglia. A Giulia e Francesco voglio fare il seguente augurio. Qualunque siano le difficoltà che vi troverete ad affrontare nella vostra vita futura, vi auguro che possiate sempre contare l'uno sull'altra nell'affrontarle. Non ho alcun dubbio, poi, che le soddisfazioni, le gioie, i ricordi lieti ed i momenti di (non) trascurabile felicità supereranno di gran lunga le preoccupazioni,

le difficoltà ed i momenti di (non) trascurabile infelicità.

Nessun ringraziamento potrebbe mai ripagare l'amore che da una vita mio padre, Francesco Cortesi, spende per la propria famiglia e per chi gli sta intorno. Provo a ripagarlo con il mio incerto fare quotidiano. Non ho dubbi di riuscirvi solo parzialmente ed in modo maldestro. Continuerò a provare. A mio padre devo ogni atomo di ciò che sono. Da lui ho imparato il senso di responsabilità, verso me stesso e verso gli altri. Da lui ho imparato a coltivare il Bello ed il Buono. Da lui ho imparato o sto provando ad imparare un'etica del lavoro, e un'etica della vita in generale. Da lui ho imparato a non accontentarmi, a costringermi a vincere la mia pigrizia, e a sforzarmi di essere ogni giorno un pochino migliore di quanto non fossi il giorno precedente. Gli auguro che nuove primavere lo attendano dietro l'angolo.