Explanatory Gaps against Physicalism and Panpsychism

A Ph.D. thesis by
Damian Aleksiev

Submitted to
Central European University
Department of Philosophy

In partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Supervisor:
Tim Crane

Associate supervisor:
Philip Goff

Budapest, Hungary
2021

COPYRIGHT NOTICE

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

Damian Aleksiev

Budapest, April 20th, 2021

ABSTRACT

How does consciousness fit within the physical world? I examine—and ultimately reject—two metaphysical theories with different answers to this question. The first theory is physicalism. It posits that the world is fundamentally only physical. The second theory is panpsychism. It posits that the physical world is fundamentally only conscious. I argue that both physicalism and panpsychism are false. Physicalism faces an explanatory gap between physical reality and human consciousness. Panpsychism faces an under-discussed explanatory gap between consciousness and the structure of physical reality. I conclude that physicalism cannot metaphysically explain human consciousness, while panpsychism cannot metaphysically explain the structure of physical reality.

The conclusion of my thesis is not entirely negative. I argue that although neither physicalism nor panpsychism can solve the problem of consciousness, the Russellian framework used by many panpsychists should not be rejected. I suggest that the Russellian framework, in conjunction with a non-panpsychist theory of fundamental reality, could be the way forward in solving the problem of consciousness.

In Chapter 1, I introduce the problem of consciousness and the metaphysical terminology I use throughout the thesis. In Chapter 2, I argue against physicalists who employ the 'phenomenal concepts strategy.' I show that the phenomenal concepts strategy rests on implausible assumptions about the nature of phenomenal concepts. In Chapter 3, I argue against Jonathan Schaffer's ground functionalism. I show that human consciousness remains especially puzzling even if exploratory gaps are abundant. In Chapter 4, I argue against panpsychism. I show that the physical structure of spacetime would lack a metaphysical explanation if panpsychism were true. Finally, in Chapter 5, I evaluate several other theories of consciousness and speculate which one of them is the most promising.

ACKNOWLEDGEMENTS

I am most grateful to Philip Goff for his enduring belief in me, his academic guidance, and his friendship throughout the writing of both my MA thesis and this Ph.D. dissertation. I first met Philip in 2015. He has inspired me to be a better philosopher ever since. I started this project under his guidance in Budapest, and he remained actively involved with it as an associate supervisor after moving to Durham. This dissertation would not have been possible without his many comments and support.

I am grateful to Tim Crane, who kindly took over the supervision of this project at CEU after Philip went to Durham. Tim influenced me to write clearly and concisely, and his feedback significantly improved the quality of the whole dissertation. Without Tim's guidance, this dissertation would have been far less refined.

For comments and useful discussions that helped me improve specific chapters, I thank Miri Albahari, Hanoch Ben-Yami, Jamie Elliott, Zhiwei Gu, Howard Robinson, Simon Rippon, and Jonathan Schaffer.

I thank James Cartlidge, Jamie Elliott, Zhiwei Gu, Garrett Mindt, Greta Rauleac, and Marta Santuccio for their friendship. I feel blessed to have met you all, and I will forever remember our many well-spent (and misspent) times in Budapest. Many other people have marked my six years in Budapest with kindness and beauty of character. I will forever carry these memories within me, and I am forever grateful to you all.

I am grateful to Robin Hendry and Daniela Slanickova for providing a warm and welcoming home for me in Durham and to Philip Goff for inviting me to be a visiting student. I thank Giacomo Giannini, Kirsty Jones, Jamie Taylor, and Andy Tomas for making me fall in love with Durham.

Finally, I thank Bori Kara and my parents Emil Aleksiev and Lenče Jovanova. This dissertation would not have been possible without your support and love.

TABLE OF CONTENTS

Chapter I:
The Problem of Consciousness

"There is nothing that we know more intimately than conscious experience, but there is nothing that is harder to explain." (Chalmers, 1995: 200)

"Every era has its Weltanschauung and in much contemporary philosophy the doctrine of 'physicalism' plays this role." (Gillett & Loewer, 2001: ix)

"The test of an ontological picture is its power: how much it enables us to explain and tie together." (Martin & Heil, 1999: 49)

1.1. Consciousness and the Physical World

1.1.1. Strange Companions

Mind and matter appear to be entirely different. Introspection reveals the mind to be the realm of sensation, thought, and feeling. From simple color and sound sensations to complex emotions such as love and despair: experiences define what it means to be human. Yet, experiences do not exhaust the whole of reality. Observation moreover reveals a material world 'outside' of the mind. Matter obeys the laws of physics and constitutes our bodies. Matter seems to exist independently of anyone's experience.

There are well-known *correlations* between the mind and the brain. Yet, their metaphysical relation is a mystery and the topic of an age-old philosophical debate. This is largely because there seems to be no intelligible connection between mind and matter. Experiences, as they appear in introspection, have a specific qualitative richness; they feel a certain way from the subject's point of view. The brain, on the other hand, as it appears to observation, is a material organ, an incredibly complex lump of tissue. Neither introspection reveals the existence of matter nor observation reveals the existence of experiences. Although they are partners involved in an intricate dance of correlations, mind and matter are eerily silent about one another.

The metaphysics of consciousness aspires to *metaphysically explain* the correlations between mind and matter. What are mind and matter, and how do they fit together? These are some of the key questions asked by metaphysicians of consciousness and are the questions that I attempt to answer in this dissertation. In this chapter, I start with the basics: I define the meaning of the terms 'mind' and 'matter'—or more precisely, of 'consciousness' and 'the physical world'—and I formulate the puzzle of their relation.

1.1.2. Consciousness

I understand the term 'consciousness' to mean *experience*. On this usage, something is a conscious state iff it essentially *feels* a specific way.[1, 2] Sensations and emotions, pleasures and pains, wills and desires, the feeling of reading these words on a page or a screen, are just a few among the countless examples of human experiences.

I take experience to be an *essential mark of the mind*.[3] Although a full catalog of human experiences might be impossible, and although there might be mental states beyond experiences, it is clear the experience is *indispensable* for the existence of consciousness. My project is about experiences and about how experiences fit within the physical world. Thus, in what follows, when I speak of the mind or consciousness, I speak of them in terms of experience.

The ordinary language usage of 'consciousness' is different from the philosophical usage that I have adopted. In its ordinary language usage, 'consciousness' is often understood more narrowly to refer to the human capacity for self-reflection or self-awareness. In the ordinary language usage, someone might say that human beings are conscious while rabbits are not conscious, although both humans and rabbits undergo experiences. In contrast, in my usage, both humans and animals are conscious, and likewise, any entity that has *any* kind of experience is conscious.

All experiences have an *essential phenomenal character*. The phenomenal character of an experience is the specific way that experience feels like; it makes each experience distinct from all other experiences. Since phenomenal characters are felt, they are particularly difficult to define or describe. Often, such attempts hinge on analogies or metaphors referring to other experiences. This makes experiences apt to be the perfect topic for poetry, literature, and art in general, and not so much the topic of the natural sciences. For example, when the French 15th century poet François Villon asked:

[1] This usage of 'consciousness' has become customary in the contemporary metaphysics of consciousness following Thomas Nagel (1974).
[2] For readers that might not be academic philosophers, the expression 'iff' is a philosophical term of art. It is the short-hand form of the logical connective *'if and only if.'*
[3] Beyond experience, there might be other essential marks of mind. *Intentionality* is another much-discussed candidate for a mark of the mental in the philosophy of mind literature.

"Where are the snows of yesteryear?" he did not intend to get an explanation from the natural sciences but to evoke a feeling of deep nostalgia for the times and people long gone.

Phenomenal characters are *qualitative*. They appear to be something that can only be fully felt but cannot be fully described. This is not to say that experiences are without structure. At least some pluralities of experiences clearly instantiate structures and can be described quantitatively. Colors can be described in terms of hue and saturation, sounds in terms of frequency, while pleasures and pains in terms of intensity, and so on. However, although interesting in their own right, such descriptions seem inapt to capture how experiences feel like. This indicates that experiences might be best described negatively, as *beyond structural*.

The data of experience is *subjective* and is the result of *introspection*. Each experience is the experience of a subject; it comes with its own first-person point of view. Moreover, experiences appear to be epistemically *private* to individual subjects. Each conscious subject appears to be the only witness of their own experiences. My experiences are my own, and the same seems to hold true for all other subjects and their experiences. Privacy should not be misunderstood to mean that each experiencer has completely unique experiences. Quite plausibly, many subjects might be feeling the same pains, or seeing the same colors, and so on. Nevertheless, each subject experiences them from *their own* point of view, and moreover, it appears no other subject has *that* same point of view.

Epistemically, the mind resembles a black box. My mind is epistemically closed off to other subjects, and the minds of other subjects are epistemically closed off to me. At best, we seem capable of indirectly inferring or intuiting the feelings of others. Plausibly, other humans have consciousness, while inanimate objects such as tables, chairs, or fruits lack consciousness. Nevertheless, we can never feel what another person feels, know if they experience the same feelings as we do, or even know if they feel at all. To give a tangible example, no amount of external observation will ever make someone understand what joy feels like unless she has experienced joy.

In summary, experiences are essential to consciousness. Experiences are known via introspection. They have essential phenomenal characters that are qualitative and beyond structural, are subjective, and appear to be epistemically private.

1.1.3. The Physical World

I use the term 'physical' to refer to the property of having a structure described by our *best current physics*. On this usage, an entity E is a *physical entity* iff it essentially has a structure described by our best current physics. Particles with mass such as quarks and leptons, the four fundamental physical forces, and spacetime are all good candidates for physical entities according to current physics.

The notion of '*matter*' is closely related, although not synonymous to that of the physical. In the contemporary scientific literature, 'matter' is typically used in a narrow way to refer to physical entities with *mass*, such as quarks and leptons. This usage entails that all material entities are physical while not all physical entities are material. For example, entities like fields and spacetime have no mass, yet, if they exist, they must be physical entities.[4] With this in mind, henceforth, I typically speak of 'physical entities and not of matter unless this is on purpose.

All physical entities have an essential *mathematical structure*. The vocabulary of contemporary physics is the quantitative vocabulary of mathematics. Mathematics is taken apt to describe the essential structures of the physical entities *fully*. Galileo Galilei, who is widely considered to be one of the founding figures of modern science, clearly expresses this attitude in *The Assayer* (1623/2008) when describing how to read the "book of the universe:"[5]

[4] In contrast to the contemporary usage, throughout history 'matter' seems to have been understood more broadly, to refer to mind-independent stuff in general. Moreover, 'matter' in its historical usage is often understood as endowed with properties that might not count as physical given current physics. For example in Aristotle's philosophy, 'matter' is understood as having value, purpose, and potentiality.

[5] Galilei was not the first to connect mathematics and physics. The mathematization of physics can be traced back to antiquity. Ancient Greek scientists such as Ptolemy and Archimedes relied heavily on mathematics, and so did medieval (Christian and Islamic) scientists. Moreover, the mathematical approach to physics was widespread in Galilei's own era and was shared by his contemporaries such as Copernicus, Kepler, and Gilbert, to name just a few.

> [The book of the universe] is written in mathematical language, and its characters are triangles, circles, and other geometrical figures; without these it is humanly impossible to understand a word of it, and one wanders around pointlessly in a dark labyrinth. (1623/2008: 183)

In addition to the use of a mathematical vocabulary, physics relies on *empirical verification*.[6] The mathematical models of physics need to be backed by reproducible experiments. The equations have to fit the current data and make verifiable predictions about any future data. Arguably, the great technological successes of modern civilization are precisely due to the use of these experimental methods. Discoveries in physics have allowed us to predict and control nature and are behind most of the technology at use in our daily lives.

The data of physics is *objective* and is the result of *observation*. The abstract nature of mathematics is understood as something beyond individual human differences and attitudes towards the world. Any being capable of rational thought should be able to grasp the language of mathematics. It does not matter how I or any other subject feels qua subject; the data of physics is the same for everyone. Thus, the data of physics can be accurately described as third-person and *public*.

It is not clear how physics translates into *metaphysics*. The great predictive power of physics is undeniable; it indicates that the descriptions of physics must match at least some of reality's objective structure. However, it is unclear what exactly physics reveals about objective reality. Perhaps, physical descriptions should be taken literally, and objective reality is fundamentally nothing but an abstract structure. Or, perhaps, there is more to objective reality than an abstract structure, and physical descriptions do not capture the full essences of physical entities.

Moreover, current physics is *not* a complete theory of physics. Instead, current physics is the conjunction of two *incompatible* theories: the theory of relativity and quantum

[6] Like the connection between physics and mathematics, the connection between physics and experimentation is not new and has a long historical foundation. Influential advocates of experimental verification include theologians such as Robert Grosseteste (12th-13th centuries) and Roger Bacon (13th century), as well as philosophers such as Francis Bacon (16th-17th centuries).

theory. The theory of relativity and quantum theory cannot both describe the objective structure of reality equally well.

The above uncertainties about defining 'physical' relate to a famous dilemma devised by Carl G. Hempel that subsequently became known as *Hempel's Dilemma*.[7] The dilemma considers two contentious proposals (horns) that seem to exhaust the options of defining 'physical':

Horn-1: define 'physical' in terms of current physics.

Horn-2: define 'physical' in terms of future physics.

Horn-1 states that the physical entities are the entities described by current physics. The main problem with Horn-1 is that *current physics is wrong*. There is a univocal consensus between philosophers and physicists that current physics is not complete. Horn-2 states that the physical entities are the entities that a future (complete) physics might describe. The main problem with Horn-2 is that *we do not know what future physics might describe*. Assuming a complete physics is possible, perhaps it would be nothing like current physics. It might even include God, souls, irreducible minds, or similar entities that are clearly not physical based on current standards.[8]

To avoid Hempel's dilemma, I adopt a middle path between the horns. My preferred definition of 'physical' has current physics as a *base* and *restricts* what can count as future physics in light of this base. I agree that current physics is incomplete and that future physics might posit very strange entities. Nevertheless, as long as the entities of future physics resemble the entities of current physics sufficiently well, I see no reason not to count them as physical. Current physics is mathematical; moreover, current physics has no reference to supernatural entities. I take supernatural entities to be hypothetical entities whose essential structures and relations are not fully describable and explainable by physics. The added clause that physical entities are not supernatural seems to distinguish them from other reputed objective entities. Thus, I

[7] See Hempel (1969).
[8] For more on Hempel's Dilemma see Ney (2008).

take it is reasonable to think future physics should likewise be mathematical and non-supernatural.

In summary, despite the difficulties, I understand the physical entities as having essential mathematical structures described by our best current physics. They are not supernatural, and their structure is objective and publicly knowable via observation.

1.2. Physicalism

1.2.1. What is Physicalism?

Physicalism can be roughly defined as the metaphysical thesis that *everything fundamental is physical*. Thus, physicalism allows for only one type of fundamental entity, namely physical entities. According to physicalists, physics describes the fundamental building blocks of the universe. Everything else that might exist—including minds—exists in virtue of these physical building blocks.

Physicalism is the orthodox view in the current philosophy of mind debates. Because of this, it is the starting point of my thesis. A testament to the popularity of physicalism comes from a recent survey among academic philosophers conducted by the editors of the PhilPapers website.[9] In the study, 56.5% of the 3000-plus interviewed academic philosophers declared they "accept or lean toward physicalism." In opposition, only 27.1% of the participants said they prefer some version of non-physicalism.[10]

The dominance of physicalism began in the 1950s with the *type identity theory* of Ullin Place (1956), Herbert Feigl (1958), and J. J. C. Smart (1959). Then, in the 1960s and early 1970s, the *analytic functionalism* of David Lewis (1966, 1972, 1980) and David Armstrong (1968), as well as the *role functionalism* of Hilary Putnam (1967) and Jerry Fodor (1974) solidified physicalism as the paradigmatic doctrine of our era in academic philosophy.

Given the above definition of physicalism, a clear understanding of physicalism requires a clear understanding of 'fundamental' and 'physical.' I have already elaborated on the notion of 'physical.' In the next section, I will characterize the notion of 'fundamental.' Along the way, I will also introduce the metaphysical framework I intend to use throughout the thesis.

[9] http://philpapers.org/surveys/results.pl

[10] They chose the 'accept or lean toward non-physicalism' option. The rest 16.4% chose the 'other' option.

1.2.2. Fundamentality, Grounding, and Reality

The *fundamental* entities are the entities that metaphysically determine and explain the existence of all other entities. In contrast, nothing determines nor explains their existence. The *dependent* entities are all the entities that are not fundamental. All entities are either fundamental or dependent. Thus, the sum of the fundamental entities is the ontological base on which everything rests.

Grounding is a relation of *metaphysical determination* and *explanation* between the more and the less fundamental entities or facts. All grounding relations obtain between *grounds*–corresponding to the fundamental entities, and *groundees*–corresponding to the less fundamental entities. The fundamental entities are the ultimate grounds; they ground everything else yet are themselves ungrounded. The dependent entities are always grounded (either in other dependent entities or ultimately in the fundamental entities).

Although the idea of grounding is ancient,[11] Kit Fine (2001), Jonathan Schaffer (2009a), and Gideon Rosen (2010) brought grounding to the forefront of contemporary philosophy. Despite the prima facie intuitive nature of grounding, there are ongoing debates in the literature concerning its exact properties.[12] What matters for my purposes here are only a few relatively uncontroversial properties of grounding. As I already mentioned, grounding is a relation of metaphysical determination and explanation. Moreover, grounding is *asymmetric*; the determination and explanation go one way, from the ground up, but not vice-versa. Third, grounding is *synchronic*. The determination and explanation in grounding cases involve entities that exist at the same time (synchronically). Finally, grounding is *transitive;* all groundees have an ultimate fundamental ground.

Grounding orders the metaphysical hierarchy. It clearly outlines the ontological structure of the world. However, grounding alone does not reveal anything about the

[11] Although it was not explicitly and rigorously defined, grounding appears to be implicitly present throughout the history of philosophy. Typically, grounding claims have been expressed as existence 'in virtue of' assertions.

[12] For an overview of the many disagreements in the grounding literature see Raven (2015).

natures of the individual entities that play the roles of grounds and groundees in the hierarchy. It is the job of both natural science and philosophy to posit what entities exist, what their natures are, and where do they belong in the metaphysical hierarchy.

All physicalists agree that all fundamental entities are physical. Moreover, physicalists agree that experiences are dependent. Nevertheless, physicalists often disagree about the exact way in which experiences and dependent entities in general exist. The orthodox position amongst physicalists appears to be that the fundamental and the dependent entities *do not* exist in the same sense of 'exist.' According to these physicalists, the dependent entities are somehow ontologically inferior or reducible to the fundamental entities. In contrast, other physicalists deny this. Regardless of which side is right, a clear understanding of the spectrum of physicalist positions is impossible without an analysis of the *two senses* of 'exist' that seem to be at play in the literature.

The fundamental entities are most naturally understood in an *ontologically privileged sense* of 'exist.' This is the sense of 'exist' deployed in serious discussions of ontology. It is what we typically mean by 'exist' when we ask whether God, Platonic universals, or quantum particles exist. I refer to the entities that might exist in this sense as 'Real' entities.[13]

> *Real entities*: an entity E is Real iff E exists in the ontological privileged sense of 'exist.'

The Real entities are *metaphysically irreducible,* and everything that reduces, reduces to the Real entities. To use a rough metaphor, the Real entities are the *building blocks of the world*; they are the entities that God must create and structure so that the rest of the world follows without further action. Epistemically, in Ted Sider's (2009, 2011)

[13] My usage of 'Real' is inspired by Kit Fine's (2001) terminology (although I do not strictly follow Fine in his usage of the term). The literature is abundant with examples of entities (or properties) that could be classified as existing in in the ontologically privileged sense of 'exist.' Platonic forms, Aristotelian and Cartesian substances, Armstrong's (1979) 'sparse' universals, Lewis' (1983, 1986) 'perfectly natural' properties, and Barnes' (2012) 'building blocks,' are just a few examples of entities/ properties that can be classified as Real.

terminology, truths about the Real entities 'carve nature at the joints.' I use the term 'Reality' to refer to the world as characterized purely in terms of the Real entities.

The consensus view among contemporary metaphysicians is that all fundamental entities are Real entities.[14] This is a highly plausible thesis that I likewise accept, and more importantly, is a thesis that contemporary physicalists, by and large, appear to accept.[15] Thus, I further qualify physicalism as follows:

> *Physicalism*: the metaphysical thesis that all fundamental entities are physical and Real.

The fundamental entities might be naturally describable as Real, but what about the dependent entities? Ordinary language is full of existence assertions that do not involve a commitment to Real entities. We talk about macroscopic objects such as tables or walls, fictional characters like Harry Potter, social institutions such as courts or ministries, and so on. We say that these entities exist, and there appear to be facts about them. Yet, it seems that ordinarily, we do not intend to say that these entities are Real. If so, there must be another sense of 'existence' at play here that is different from the above-mentioned ontologically privileged sense. The ordinary language usage of 'exists' seems to involve a *nominal sense* of 'exist.' Henceforth, I refer to the entities that exist in this sense as 'Nominal' entities.

> *Nominal entities*: an entity E is Nominal iff E exists in the nominal sense of 'exist.'

When in the role of groundees, Nominal entities are *metaphysically reducible* to the Real entities that ground them. The existence of Nominal groundees is *derived* from

[14] This appears to be the position of Fine (2001), Dorr (2005), Schaffer (2009a), and Sider (2009, 2011), amongst others. For a defense of this position, see Sider (2009).

[15] It is worth noting that there are respectable anti-realist views in metaphysics. These views reject the Reality of the fundamental entities. It is questionable whether an anti-realist theory could provide a genuine metaphysical explanation of experiences. Thus, since the topic of this dissertation are *metaphysical* explanations of experiences, my focus is solely on Realist theories of the fundamental entities.

the existence of the Real entities. They are nothing 'over and above' their Real grounds; or, in David Armstrong's (1997: 12) words, they are a metaphysical 'free lunch.' Epistemically, if there are truths about the Nominal groundees, these are made true in virtue of their Real grounds. Nominal groundees correspond to how humans (and rational subjects in general) describe and think about the Real entities. Their existence is a useful convention; however, it does not carve nature at its joints.

It is an open question whether dependent entities are Nominal or Real. Historically and in the current literature, many philosophers and scientists have thought that at least some dependent entities are Real. Chemical compounds, living organisms, and minds have been posited as plausible examples of Real dependent entities.[16]

In light of the above, I distinguish two instances of grounding based on whether or not the dependent entities are Nominal:[17]

> *Reductive grounding*: a grounding relation is reductive iff it involves a Real fundamental ground and a Nominal groundee.
>
> *Non-reductive grounding*: a grounding relation is non-reductive iff it involves a Real fundamental ground and a Real groundee.

Physicalists agree that experiences *exist*.[18] However, they disagree whether experiences are Nominal or Real. This disagreement results in two broad versions of physicalism that I call: reductive and non-reductive physicalism.

Reductive physicalism, as I define it, includes all versions of physicalism that posit that experiences are *reductively grounded* in the physical entities. According to reductive physicalists, experiences (and, plausibly, all dependent entities) are

[16] Typically, the resulting position has been labelled '*emergentism*.' I discuss the relation between emergentism and physicalism in more detail in Chapter 5.
[17] Fine (2001: 27) takes reductive grounding to be the norm, while non-reductive grounding to be the odd exception. On the other hand, Schaffer's (2009a, 2017, forthcoming) notion of grounding is exclusively non-reductive.
[18] This claim is true for all versions of physicalism except *eliminative physicalism*. Eliminative physicalism is a radical view that is beyond the scope of this dissertation.

Nominal. Reductive physicalists include type-identity physicalists like Papineau (2002), analytic functionalists like Lewis (1966, 1972) and Armstrong (1968, 1997), and grounding physicalists like Derk Pereboom (2011, 2013), among others.

Traditionally, the label reductive physicalism has been reserved for the identity theory. In contrast to tradition, I use reductive grounding as a broad label for many reductive relations with the same ontological commitments to Reality.[19]

The reader might object to my conflation of type-identity and reductive grounding. After all, the reader might say: "is not type-identity symmetric while reductive grounding is asymmetric?" In response, I argue that type-identity claims contain is a hidden asymmetry that is indispensable to their ontological implications.

Type-identity claims are easily translatable to grounding claims involving Nominal groundees. For example, take the famous (although empirically dubious) claim that 'pain is C-fibers firing' (CFF). 'Pain is CFF' does not state that both pain and CFF are ontologically equal. Instead, 'pain is CFF' is a reductive claim. It states that pain is metaphysically reducible to CFF. Thus, CFF must be ontologically *superior* to pain. A natural way to read this reductive claim is in terms of CFF being Real and pain being Nominal. In this reading, pain is CFF as described from the subjective perspective of humans. Thus it seems, type-identity claims have a *hidden asymmetry*. Without this asymmetry, pain and CFF would be ontologically equal, and 'pain is CFF' would *not* entail physicalism. Grounding makes this hidden asymmetry *visible*. Thus, it seems, type-identity and reductive grounding have the same ontological commitments to Reality.

Grounding is more informative than type-identity. Grounding clearly reveals the *directionality* of type-identities and other reductions. It makes clear what reduces to what and what is more fundamental vs. what is less fundamental. Thus, I typically speak of reductive grounding instead of type-identity or other finer-grained reductive relations.

[19] 'Small-g' relations, in Wilson's (2014) terminology.

Non-reductive physicalism, on the other hand, as I define it, includes all versions of physicalism that posit that experiences are *non-reductively grounded* in the physical entities. According to non-reductive physicalists, experiences (and, plausibly, other dependent entities) are Real. Non-reductive physicalists include role functionalists like Putnam and Fodor, and Johnathan Schaffer (2017, forthcoming) who defends ground functionalism.

It is important to note that both reductive and non-reductive physicalists accept there are facts about the dependent entities, including experiences. As Montero and Papineau point out: "Physicalists do not deny that there are mental and biological facts. Of course there are. They don't want to eliminate these facts, but illuminate their nature." (2016: 184) Reductive physicalists deny that these facts are made true in virtue of Real entities 'over and above' the fundamental physical entities. Non-reductive physicalists like Schaffer, on the other hand, see no problem in the claim that these facts are made true in virtue of dependent Real entities.

1.2.3. Types of Physicalism

All physicalists agree that fundamental Reality is physical. Yet, physicalists disagree about the metaphysics of experiences and about how to explain the existence of experiences. To shed light on these difference, I classify all versions of physicalism along *three dimensions*, based on two metaphysical and one epistemic question:

> *The 1st metaphysical question* is: are experiences Real?
>
> *The 2nd metaphysical question* is: are the fundamental physical entities purely structural?
>
> *The epistemic question* is: can experiences be intelligibly explained in virtue of fundamental Reality?

In response to the *1st metaphysical question*, as I already argued, physicalists can embrace either reductive or non-reductive physicalism.

Are experiences Real?	
No: *Reductive Physicalism*	Yes: *Non-reductive Physicalism*

Table 1: Reductive and non-reductive physicalism.

Second, in response to *the 2nd metaphysical question*, physicalists can either embrace pure physicalism or extended physicalism.

Pure physicalism, as I define it, is the view that fundamental Reality *is* a pure causal structure. This view amounts to taking the mathematical models of physics literally. It entails that there is nothing more to the nature of the fundamental Reality than the structure and dynamics described by physics. The label 'pure physicalism' comes from Philip Goff (2017). However, in the literature, the same view can be found under many names: 'theory-based physicalism' (Stoljar, 2001), 'physicSalism' (Strawson, 2006b), 'narrow physicalism' (Chalmers, 2015), 'structuralism,' and 'causal-structuralism.'

Extended physicalism, as I define it, is the view that fundamental physical Reality *is not* only a causal structure. Proponents of this view are happy to accept that the mathematical models of physics accurately describe the structure and dynamics of fundamental Reality. However, instead of taking the models literally, they think that the models are about objects (or an object) with essential properties that are not fully describable by quantitative models. If so, there is more to the fundamental Reality's nature than the structure and dynamics that physics describes. In the literature, this view can be described as 'impure physicalism' (Goff, 2017), 'object-based physicalism' (Stoljar, 2001), 'broad physicalism' (Chalmers, 2015), and 'Russellian physicalism' (Montero, 2015).

Are the fundamental physical entities purely structural?	
Yes: *Pure physicalism*	No: *Extended physicalism*

Table 2: Pure and extended physicalism

Pure physicalism appears to be the most popular physicalist doctrine and is often assumed (either explicitly or implicitly) by both philosophers and physicists. Thus, throughout this thesis, I typically assume pure physicalism. I deal with extended physicalism only in the final chapter of the thesis (Chapter 5).

Finally, in response to *the epistemic question*, physicalists can be either epistemically reductive or epistemically non-reductive. Traditionally, this difference has been framed in terms of *theory reduction*. Theory reduction requires the derivation of the *laws* of a higher-order scientific theory from the laws of a more fundamental scientific theory. However, in more recent debates, this demand has been largely abandoned and replaced with the *weaker* demand for an *intelligible connection* between grounds and groundees. Following Joseph Levine (1983), this weaker notion of epistemic reduction is typically expressed in terms of *explanatory gaps*.

Chalmers (2003) gives an influential characterization of physicalist views based on whether or not they accept the presence of an explanatory gap between the physical facts (as grounds) and the experiential facts (as groundees).

Type-A physicalists, in Chalmers' terminology, think that there *is no* explanatory gap between the physical and experiential facts. Type-A physicalists are *epistemically reductive* physicalists. The analytic functionalism of Lewis and Armstrong is a paradigmatic example of type-A physicalism.

Type-B physicalists, in Chalmers' terminology, think that there *is* an explanatory gap between the physical and experiential facts. Type-B physicalists are *epistemically non-reductive* physicalists. Contemporary physicalists are, by and large, type-B physicalists. Type-B physicalists include *phenomenal concepts strategists* like Brian Loar (1990/1997), Christopher Hill (1997), David Papineau (1998, 2002), Katalin Balog (1999, 2012b), and Esa Diaz-Leon (2014, 2016),[20] as well as non-reductive physicalists like Schaffer (2017, forthcoming).

[20] The name 'Phenomenal Concept Strategy' is due to Stoljar (2005).

Can experiences be intelligibly explained in virtue of fundamental Reality?	
Yes: *Type-A physicalism*	No: *Type-B physicalism*

Table 3: Type-A and type-B physicalism.

The notion of 'explanatory gap' and the divide between type-A and type-B physicalists is crucial for this dissertation. I return to and elucidate the notion of 'explanatory gap' towards the end of this Chapter. Moreover, Chapter 2 and Chapter 3 deal with different versions of type-B physicalism.

I end this section with one final note on terminology: the metaphysical and the epistemic senses of 'reduction' are different and must not be conflated.[21] My distinction between reductive and non-reductive physicalism involves *metaphysical* reduction. In contrast, closing explanatory gaps involves a weak kind of *epistemic* reduction. For clarity of presentation, I use the unqualified term 'reduction' to refer strictly to metaphysical reduction. In contrast, when speaking of epistemic reduction, given that my focus is on explanatory gaps, I typically use Chalmers' terminology of type-A and type-B views or the qualified expression 'epistemic reduction.'

1.2.4. The Causal Argument for Physicalism

Why think physicalism is true? So far, I have defined physicalism and its many versions but have not given any reason to believe in it. As David Papineau (2001) points out, historically, physicalism was not the dominant view. Instead, quite the opposite, the history of philosophy is largely the history of various versions of non-physicalism. According to Papineau, the main reason for the recent overwhelming popularity of physicalism is *empirical*. Specifically, Papineau suggests that the rise of physicalism is due to the strong empirical evidence for a thesis known as the *'completeness of physics*.' This is the thesis that "all physical effects are due to physical causes" (2001: 7).

[21] See Crane (2001: 54–5, 2010) for a detailed analysis of the difference between metaphysical and epistemic reductions and the implications of these notions for physicalism.

The Causal Argument is the most influential argument for reductive physicalism in the current literature.[22] Versions of the argument can be found across the literature, going back to Lewis (1966) and Armstrong (1968). Here, I focus on Papineau's (2001: 9) version:

P-1. *The Completeness of Physics:* "All physical effects are fully determined by law by prior physical occurrences."
P-2. *Mental Causal Influence:* "All mental occurrences have physical effects."
P-3. *No Universal Overdetermination:* "The physical effects of mental causes are not all overdetermined."
C. "Mental occurrences must be identical with physical occurrences."

P-1, The Completeness of Physics, according to Papineau, is the crucial premise in the Causal Argument. P-1 is an empirical premise that, as Papineau argues, is based on two important empirical discoveries. First, current research in physics has shown that all causation is the manifestation of a small number of energy-conserving fundamental forces: gravity, electromagnetism, the weak nuclear force, and the strong nuclear force. Second, current research in physiology has found no evidence for any mental forces. Thus, based on current empirical knowledge, we can inductively conclude that all physical causal effects are caused by physical causes. Moreover, all currently known physical effects are caused in a lawlike manner. Thus, the physical world seems to form a causally closed whole in which all causation is lawlike and physical. Rejecting P-1 seems to entail rejecting our best empirical science.

P-2, Mental Causal Influence, expresses the common assumption that mental states have physical effects. It seems evident that our sensations, feelings, beliefs, and desires causally influence our actions. For example, if I feel hungry, I might eat, if I believe it is raining, I might take an umbrella with me, and so on. By itself, this premise does not make any further assumptions about the metaphysics of experiences. Mental states might be Real or not. In either case, the premise simply asserts that they have physical effects.

[22] In Chapter 3 I present another argument for physicalism, that is Schaffer's (2017, forthcoming) argument for his version of non-reductive physicalism.

Rejecting P-2 entails embracing *epiphenomenalism* about the mind. Epiphenomenal entities are causally infective (at least in relation to the physical world).[23] Epiphenomenalism is standardly considered to be a highly implausible view. For one, it is not clear if epiphenomenalism is coherent.[24] Moreover, even if it is coherent, epiphenomenal entities seem to be metaphysical oddities; they appear to be unlike any other entities we know to exist. All examples of putative epiphenomenal entities are controversial. Finally, embracing epiphenomenalism seems to entail a deep divide between the physical and the mental as two independent causal realms. Such a result is neither empirically verifiable nor metaphysically parsimonious. Thus, based on all of these reasons, it seems that epiphenomenalism, even if coherent, is an option of last resort and not a good reason to reject P-2.

P-3, No Universal Overdetermination, rejects the idea of universal overdetermination. Overdetermination is the thesis that some effects have multiple distinct yet simultaneous causes. Each of these causes is sufficient for the effect. Thus, as long as one cause obtains, the effect will be produced.

In the case of experiences, overdetermination entails that the physical effects of mental causes have two sufficient causes: (a) a mental cause and (b) a physical cause. Take the voluntary movement of an arm as an example. In this example, the arm's movement seems to be a physical effect of the will to move the arm as a *mental cause*. However, the arm's movement also has another distinct and *purely physical cause*—the firing of neurons. Thus, it seems either the mental cause alone or the physical cause alone could have produced the same physical effect. If so, either the mental cause or the physical cause would have been sufficient to move the arm. Thus, it seems the arm would have moved even if there was *no* mental cause. We already have strong empirical reasons to believe in the existence of the physical cause. Thus, the mental cause must either be redundant or overdetermining the physical effect (namely, the arm's movement).

[23] Epiphenomenalism does not necessarily entail complete causal inefficacy. Epiphenomenal experiences might be causally effective with other experiences.
[24] The principle that causal efficiency is a prerequisite for existence is known as the 'Eleatic Principle,' the 'Causal Criterion,' or 'Alexander's dictum' (after the British emergentist Samuel Alexander).

Although overdetermination is coherent, at best, it seems to be a highly improbable and rare occurrence. There are no clear examples of systematic overdetermination in nature. Thus, believing without further argument that overdetermination happens systematically only in the mental case but nowhere else appears unreasonable. Thus, it seems we have strong reasons to reject overdetermination and instead accept P-3.

The above leads us to *the conclusion* of the argument. In light of the completeness of physics, experiences appear to be either epiphenomenal or causally redundant (or overdetermining physical effects). Thus, since all of these options are undesirable, it is most reasonable to conclude that experiences *are* physical, or in my terminology, that experiences *are reductively grounded* in the physical. If so, physicalism is true.

The causal argument is a good argument. The argument moves from the empirical findings of natural science (in P-1) to the rejection of two implausible theses (epiphenomenalism in P-2 and overdetermination in P-3) to the acceptance of physicalism (in the conclusion). Rejecting it entails either rejecting causal closure or embracing epiphenomenalism or overdetermination. All of these positions, although plausibly coherent, seem less desirable than accepting physicalism.

1.3. The Problem of Consciousness

1.3.1. Why Is Consciousness a Problem?

The problem of consciousness is roughly the problem of metaphysically explaining how experiences fit within the physical world. Physicalists say experiences fit because experiences *are* physical (or grounded in the physical). However, as I will soon show, there are strong reasons to doubt the physicalist project. I start this chapter by giving a somewhat rough and intuitive exposition of the problem and illustrating it with examples from history. I give a more rigorous analysis of the problem in the subsequent section.

A good metaphysical theory must be *explanatory*. In ground-theoretic terms, the grounds must *metaphysically explain* their groundees. Although there are many ways to understand 'metaphysical explanation,' there is broad agreement in the literature that, at the very least, there should be an *intelligible connection* between grounds and groundees. The lack of an intelligible connection, as I already mentioned, is standardly expressed in the current literature as an '*explanatory gap*.'

Physicalism appears to be explanatory when it comes to physical groundees. Many macroscopic physical phenomena appear to have explanations in terms of more fundamental physical phenomena. Empirical research reveals that water is H_2O, that lighting is electrical discharge, that heat is molecular motion, and so on. In all of these cases, the explanans is a macroscopic, physical kind (water, lighting, heat, etc.), while the explanandum is a more fundamental physical kind (H_2O, electrical discharge, molecular motion). In all of these cases, once a physical ground has been discovered, there appears to be no residual mystery left about the obtaining of the explanans. Thus, in all of these cases, the physical grounds appear to metaphysically explain their physical groundees. Thus, there appeared to be no explanatory gaps between physical grounds and their physical groundees.

In contrast, there seems to be *no* intelligible connection between the physical facts (as grounds) and the experiential facts (as groundees). We know that the mind and the

brain are correlated. Nevertheless, despite the amazing advances of the natural sciences, the data of the natural sciences appears incomplete when it comes to experiences. The phenomenal characters of experience—namely, how experiences feel like from the subjective point of view—seem to remain outside of the grasp of the natural sciences. The postulation of a physical ground does not appear to shed light on the nature of experiences.

If the above is right, physicalism is *not explanatory* when it comes to experiences. If so, consciousness is a problem for the physicalist because there is (or at least appears to be) an explanatory gap between the physical and the experiential facts.

Explanatory gap worries about experiences are not new. They have been raised in different guises by many philosophers throughout history. Leibniz, in the 18th century, famously illustrated his own worries on the same topic with an analogy:

> Imagine there were a machine which by its structure produced thought, feeling, and perception; we can imagine it as being enlarged while maintaining the same relative proportions, to the point where we could go inside it, as we would go into a mill. But if that were so, when we went in we would find nothing but pieces which push one against another, and never anything to account for a perception. (1714/1998: 270)

The same problem was discussed throughout the 19th century by many scientists and philosophers. Emil du Bois-Reymond, Thomas H. Huxley, and John Tyndall argued that the natural sciences are inapt to fully explain experiences.[25] Huxley expressed his concerns particularly vividly when he wrote:

> How it is that anything so remarkable as a state of consciousness comes about as a result of irritating nervous tissue, is just as unaccountable as the appearance of Djin when Aladdin rubbed his lamp. (1866)

[25] For a detailed overview of the history of the explanatory gap in the 19th century, see Tennant (2007).

The problem of consciousness was well apprehended by G. M. Moore (1903), C. D. Broad (1925), and Bertrand Russell (1927) in the early 20th century. While discussing the experience of the color yellow, Moore writes:

> We may try to define it by describing its physical equivalent; we may state what kind of light-vibrations must stimulate the normal eye, in order that we may perceive it. But a moment's reflection is sufficient to [show] that those light-vibrations are not themselves what we mean by yellow. (1903: 62)

Discussions of the problem of consciousness reached their peak in the last decades of the 20th century. Robert Kirk (1974b, 1974a), Thomas Nagel (1974), Saul Kripke (1972/1980), Frank Jackson (1982), Howard Robinson (1982), Joseph Levine (1983) (who coined the phrase 'explanatory gap'), Colin McGinn (1989), and David Chalmers (1995, 1996), among others, have questioned, in different ways, whether there is an intelligible connection between the physical and the experiential facts. This intelligible connection is what Colin McGinn had in mind when he asked: "How is it possible for conscious states to depend on brain states? How can technicolour phenomenology arise from soggy grey matter?" (1989: 349)

The lack of an intelligible connection between the physical and the experiential facts is at the essence of David Chalmers' (1995, 1996, 2003) famous distinction between the 'easy problems' and the 'hard problem' of consciousness.

The easy problems, in Chalmers' terminology, are all those problems of consciousness that are "straightforwardly vulnerable to explanation in terms of computational or neural mechanisms" (1995: 201). These include explaining cognitive functions such as "discrimination, integration, access, report, control" (2003: 104) and similar. The 'easy' problems are not easy because they are easy to solve. Quite the opposite, some of the 'easy' problems might be extremely hard to solve. What makes them 'easy' is that these problems are *in principle* solvable by the natural sciences.

The hard problem of consciousness, in Chalmers' terminology, is the problem of experience. Chalmers argues that even if we have solved all the essay problems and

thus have a full account of the structure, dynamics, and functions of the brain, it remains a further question: "why is the performance of these functions accompanied by experience?" (2003: 104). If the hard problem is sound, then there is at least one undeniable datum that seems to escape a full explanation via the methods of the natural sciences, and this is the datum of experience.

1.3.2. The Consciousness Explanatory Gap

An intuitive way to grasp the notion of an 'intelligible connection' is in terms of what could an ideal intellect infer a priori from the fundamental facts. Imagine that an ideal intellect—for example, Laplace's hypothetical demon[26]—knows all the fundamental facts. If this were the case, would Laplace's demon be able to a priori infer any other facts? For all the facts where the answer is *'yes,'* there is an intelligible connection between grounds and groundees. For all the facts where the answer is *'no,'* there is no intelligible connection between grounds and groundees.

Chalmers illustrates this idea by using the biological facts as an example:

> We can imagine that a hypothetical superbeing—Laplace's demon, say, who knows the location of every particle in the universe—would be able to straightforwardly 'read off' all the biological facts, once given all the microphysical facts. The microphysical facts are enough for such a being to construct a model of the microscopic structure and dynamics of the world throughout space and time, from which it can straightforwardly deduce the macroscopic structure and dynamics. (1996: 35)

Laplace's demon might be a good heuristic device for finding intelligible connections; however, it is not a rigorous definition of the notion. A more rigorous way to cash out 'intelligible connection' is in terms of *a priori entailment*.

[26] Named in honor of Pierre-Simon Laplace (1814/1951) who devised a powerful though experiment intended to showcase determinism about microphysics. It is worth noting that Laplace never explicitly mentions a 'demon' in his writing, but only speaks of a vast intellect. The label 'demon' is the addition of later commentators.

Chalmers and Jackson (2001), and Chalmers (2012), give a thorough elucidation of 'a priori entailment.' In their usage, P a priori entails Q iff the material conditional 'if P then Q' is a priori knowable. A priori entailment, so construed, is about *facts* and not about laws. This makes it a "weak sort of reduction [...] compatible with various sorts of irreducibility that are manifest in science." (Chalmers, 2012: 309) Moreover, as Chalmers argues, a priori entailment is apt to provide "transparent bottom-up explanation," meaning that "once one has spelled out the lower-level facts [...] there is no residual mystery about what the high-level facts are or about how the low-level facts give rise to them." (2012: 305)

We are now finally in a position to give a rigorous definition of explanatory gaps. Given that 'intelligible connection' stands for 'a priori entailment' and since all the relevant cases I discuss involve grounding, I define 'explanatory gap' as follows:

> *Explanatory Gaps*: an explanatory gap obtains iff there is *no* a priori entailment between a ground and a groundee in a putative case of grounding.

Furthermore, given the above, I can now give a rigorous definition of the reputed explanatory gap between the physical and experiential facts that seems to be at the essence of the problem of consciousness:

> *Consciousness Gap*: The physical facts (as grounds) do not a priori entail the experiential facts (as groundees).

Explanatory gaps matter because there are good reasons to think that, as a general rule, there should be no explanatory gaps in true cases of grounding. As I already mentioned, there seem to be many grounding cases involving physical grounds and physical groundees, where there are no explanatory gaps. Water (as a groundee) and H_2O (as a ground), lighting (as a groundee) and electrical discharge (as a ground), heat (as a groundee) and the motion of molecules (as a ground), are all plausible examples of grounding without explanatory gaps. In these cases, when both the groundee and

the ground are conceived in purely physical terms, it seems that the ground a priori entails the groundee.

Explanatory gaps are counter-examples to specific *scrutability theses*. A 'scrutability thesis,' in Chalmers' (2012) terminology, is any thesis that posits a compact class of truths—called *a scrutability base*—that a priori entails all the other truths.[27] Scrutability theses are a priori in the sense that given the truths in a scrutability base, a good reasoner needs no further information to work out the entailments between them and all other truths.

The above considerations about no explanatory gaps in true grounding cases indicate that some *fundamental* scrutability thesis is likely true. A fundamental scrutability thesis is any scrutability thesis positing that "all truths are scrutable from metaphysically fundamental truths plus indexical truths." (Chalmers, 2012: 470) In other words, if a fundamental scrutability thesis is true, there should be no explanatory gaps between the fundamental facts and all other facts. Explanatory gaps that are counter-examples to a fundamental scrutability thesis indicate that the metaphysical theory that produced that thesis might likewise be false. If so, explanatory gaps are a powerful tool for the assessment of metaphysical theories.

Physicalists posit that the fundamental facts are entirely physical. These are the so-called '*microphysical*' facts. The microphysical facts involve "truths about fundamental physical entities in the language of a completed fundamental physics." (Chalmers, 2012: 111) If the above considerations about fundamental scrutability are sounds and physicalism is true, the microphysical facts should a priori entail all the other facts. As Levine puts it:

> If nature is one large, lawful, orderly system, as the materialist (or naturalist) insists it is, then it should be possible to explain the occurrence of any part of that system in terms of the basic principles that govern nature as a whole. (2001: 69)

[27] The scrutability base is 'compact' mainly to exclude trivial scrutability bases which include all the truths (Chalmers, 2012: 20–3).

I refer to the resulting scrutability thesis as 'Cosmic Scrutability.'[28]

> *Cosmic Scrutability*: The microphysical facts (as grounds) a priori entail all other facts (as groundees).

Anti-physicalists *reject* Cosmic Scrutability based on Consciousness Gap. They argue that Consciousness Gap is a counter-example to Cosmic Scrutability, and thus, Cosmic Scrutability must be false. If so, any version of physicalism that advocates Cosmic Scrutability must likewise be false. Consciousness Gap poses an epistemic challenge against physicalism. Since the above stipulations about fundamental scrutability appear plausible and are broadly accepted, physicalists are obliged to answer the Consciousness Gap challenge. In the light of this, I agree with Levine that Consciousness Gap is "the main obstacle to acceptance of materialism." (2001: 76)

In a nutshell, the problem of consciousness for the physicalist is this: physicalists must either *close* Consciousness Gap or find a way to demonstrate that physicalism is *compatible* with Consciousness Gap.

Type-A physicalists *accept* Cosmic Scrutability and *reject* Consciousness Gap. These physicalists stick to their guns and argue that Consciousness Gap can be closed. In their view, once all the fundamental facts are spelled out in the vocabulary of a completed physics, we should be able to a priori deduce all the experiential facts. According to type-A physicalist, "if you knew enough about the workings of my body and brain, [...] you could work out, for example, whether or not I'm currently feeling anxious." (Goff, 2017: 84–5)

Type-B physicalists, in contrast, *reject* Cosmic Scrutability and *accept* Consciousness Gap. These physicalists bite the bullet and argue that physicalism is compatible with Consciousness Gap. As I already mentioned, there are two main camps of type-B physicalists: phenomenal concepts strategy (PCS) physicalists and non-reductive physicalists.

[28] My terminology is inspired by Terry Horgan (1984) who articulated the same thesis under the name of '*Cosmic Hermeneutics*.'

Typically, type-B physicalists are PCS physicalists. PCS physicalists aspire to give a physical explanation of why Cosmic Scrutability fails, and Consciousness Gap obtains. They argue that this is so because of the *special nature* phenomenal concepts (the concepts we use to think about experiences in terms of how they feel). Thus, in their view, although Cosmic Scrutability should work in principle, the special nature of phenomenal concepts explains why it fails in the case of consciousness.

In addition to the PCS physicalists, there is also a more radical camp of type-B physicalists who reject the notion of fundamental scrutability altogether. Schaffer's (2017, forthcoming) ground functionalism is a prime example of such a view. Schaffer argues that *explanatory gaps are everywhere*, in *all* cases of grounding. Thus, according to Schaffer, no explanatory gap is special. Schaffer argues that physicalism is true and compatible with all of these explanatory gaps (including Consciousness Gap).

Contemporary physicalists are by and large type-B physicalists. One of the main reasons behind the rise of type-B physicalism is the perceived failure of type-A physicalists to defend Cosmic Scrutability and close Consciousness Gap. My focus in the next two chapters will be primarily on type-B physicalism, given its popularity and novelty. In Chapter 2, I explore how PCS physicalists respond to the Consciousness Gap challenge. In Chapter 3, I explore Schaffer's ground functionalism response to the Consciousness Gap challenge.

1.3.3. Epistemic Arguments against Physicalism

Refuting physicalism based on Consciousness Gap requires an *argument*. Consciousness Gap is an epistemic thesis. Physicalism is a metaphysical thesis. Consciousness Gap is impotent against physicalism without further premises.

Three arguments against physicalism dominate the current metaphysics of consciousness literature. These are the three *'epistemic arguments'* (in Chalmers' (2003, p. 108) terminology) against physicalism: the conceivability argument, the

knowledge argument, and the explanatory argument. Consciousness Gap features as an *epistemic premise* in all three of these arguments.

In the conceivability argument, Consciousness Gap is expressed in terms of conceivability. In the knowledge argument, Consciousness Gap is expressed in terms of knowledge. In the explanatory argument, Consciousness Gap is expressed in terms of explanation. 'Conceivability,' 'knowledge,' and 'explanation' in the three arguments are best seen as different aspects of the same underlying and more fundamental notion—a priori entailment.

The three epistemic arguments against physicalism share the *same structure* and establish the *same conclusion*. Their structure involves a move from Consciousness Gap to a metaphysical gap. All three arguments establish that physicalism is false as a conclusion. Formally stated, their structure is as follows:[29]

Step-1:	Consciousness Gap (expressed in terms of either conceivability, knowledge, or explanation) is true.
Step-2:	There is a metaphysical gap between the physical and the experiential facts.
Conclusion:	Physicalism is false.

In what follows, my main focus will be on the conceivability argument. This preference is not only aesthetic but, moreover, as I will soon show, is because the link between a priori entailment and conceivability is especially clear. Thus, I consider the conceivability argument to be the best among the epistemic arguments against physicalism.

[29] My reconstruction is based on Chalmers (2003: 108).

Chapter II:
Against Reductive Physicalism

"The conceivability of zombies is [...] the principal manifestation of the explanatory gap." (Levine, 2001: 79)

"At the root of almost all weird positions in the philosophy of mind lies this rather elementary and unremarkable conceptual fact, blown up into a metaphysical problem that appears to require an extreme solution." (Loar, 1990/1997: 609)

"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 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." (Goff, 2017: 108)

2.1. The Conceivability Argument

2.1.1. The Simple Conceivability Argument

The conceivability argument—like all the epistemic arguments against physicalism—aspires to establish a metaphysical conclusion from epistemic premises.[30] Conceivability is a property of propositions. A proposition is conceivable iff it is *rationally coherent*. The conceivability argument uses the rational coherence of specific propositions to a priori establish that physicalism is false.

The conceivability argument dates back at least to René Descartes (1645/1996).[31] The contemporary version of the argument is based mainly on the work of Robert Kirk (1974b), Saul Kripke (1972/1980), and David Chalmers (1996, 2003, 2010). Here, I focus on Chalmers' canonical formulation of the argument.

The *Simple Conceivability Argument*:

P-1. 'P & ~Q' is conceivable.
P-2. If 'P & ~Q' is conceivable, 'P & ~Q' is possible.
P-3. If 'P & ~Q' is possible, physicalism is false.
C. Physicalism is false.

'P' stands for the *total fundamental physical facts*, as given by a complete theory of physics. '*Q'* stands for *an experiential fact*, such as the fact that I feel hunger now. Assuming physicalism is true, P contains all the fundamental facts. Moreover, if physicalism is true, Q must be grounded in P. The argument posits that despite the putative grounding relation between them, P and Q are conceivable apart. In other words, that 'P & ~Q' is conceivable. Moreover, the argument posits that the conceivability of 'P & ~Q' entails a genuine possibility. Namely, that P could obtain

[30] See section 1.3.3.
[31] Descartes' original conceivability argument is about the conceivability of ghosts (disembodied minds) and not zombies (mindless bodies) like the contemporary version. The zombie version is more popular since in current debates since it seems to do better against grounding versions of physicalism that accept multiple realizability. For more on this see Goff (2017 chapter 4.1.1).

without Q obtaining. This possibility is incompatible with physicalism. If this is a genuine possibility, Q cannot be grounded in P. If so, physicalism must be false.

An easy way to grasp the argument is to think of 'P & ~Q' in terms of *philosophical zombies*. Philosophical zombies (or simply zombies) are described in the literature as *full physical duplicates* of real humans. Zombies are observationally indistinguishable from humans. All the physical facts that obtain in humans obtain in zombies. The difference between humans and zombies is that zombies lack experiences. Zombies instantiate only P facts, while humans instantiate both P and Q facts. There is *nothing* that it is like to be a zombie.

The *Simple Conceivability Argument* (Zombie Version):

P-1. Zombies are conceivable.
P-2. If zombies are conceivable, zombies are possible.
P-3. If zombies are possible, physicalism is false.
C. Physicalism is false.

If physicalism is true, either (1) zombies should be inconceivable (meaning P-1 is false), or (2) if zombies are conceivable, they should not be possible (meaning P-2 is false).

P-1 is a redescription of Consciousness Gap in terms of conceivability. Remember:

> *Consciousness Gap*: The physical facts (as grounds) do not a priori entail the experiential facts (as groundees).

'P & ~Q' is the negation of 'P entails Q.' Moreover, 'P & ~Q' is a priori. Thus, to say that 'P & ~Q' is conceivable is the same as to say there is *no a priori entailment* from P to Q. In general, for a ground A and a groundee B, if it is conceivable that A obtains while B fails to obtain, then there is an explanatory gap between A and B (see *Table 4*).

Can we conceive of a scenario in which a ground obtains, but a groundee does not?	*Yes*	Explanatory gap
	No	No explanatory gap

Table 4: Conceivability as a guide to explanatory gaps.

Moreover, the reverse s likewise true; namely, if there is an explanatory gap between a ground A and a groundee B, it must be conceivable that A obtains while B fails to obtain. If so, conceivability gaps translate into explanatory gaps, and vice-versa, explanatory gaps translate into conceivability gaps.

The above considerations demonstrate that if P-1 is true, Consciousness Gap is true, and vice-versa, if Consciousness Gap, P-1 is true. As Levine points out: "The conceivability of zombies is [...] the principal manifestation of the explanatory gap." (2001: 79) In response, as I already mentioned, type-A physicalists reject Consciousness Gap and P-1, while type-B physicalists accept Consciousness Gap and P-1.

P-2 is founded on the assumption that *conceivability entails possibility*. Thus, it posits that a priori conceptual relations translate into metaphysical truths about the referents of the involved concepts. In contemporary debates, P-2 is controversial and is widely contested by physicalists. Type-B physicalists accept Consciousness Gap and agree that zombies are conceivable (meaning, P-1 is true). However, they deny that conceivability entails possibility and thus reject P-2. Typically, type-B physicalists are *phenomenal concept strategy* physicalists. Later in this chapter, I will argue against the phenomenal concept strategy and defend P-2.

P-3 posits that the conceivability of zombies is incompatible with physicalism. Many physicalists understand physicalism as a *modal thesis* about a necessary connection between the fundamental physical facts (P) and the dependent experiential facts (Q). However, if zombies ('P & ~Q') are possible, this necessary connection is false. Thus, P-3 is commonly accepted by both physicalists and anti-physicalists. If true, P-3 straightforwardly leads to *the conclusion* of the argument, namely that physicalism is false.

2.1.2. Imagination and Conceivability

Before proceeding to the different physicalist responses to the conceivability argument, I first clarify P-1 and the notion of conceivability. This is crucial to avoid misconceptions about the argument.

Conceivability *is not* imagination. Imagination is typically understood as the ability to create mental images and thus is dependent on *experiences*.[32] Conceivability, in contrast, is best understood as rational coherence, or as Chalmers (2002) puts it, as what *cannot be ruled out a priori*. In E. J. Lowe's words: "Conception, one might want to say, involves the deployment of concepts in thinking, whereas imagination just deploys perception-based imagery." (2012: 924)

Both imagination and conceivability seem to require rational coherence; however, only imagination seems to involve experiencing necessarily. This means that phenomenality is essential for imagination while only accidental for conceivability.[33]

The difference between conceivability and imagination is perhaps most visible in cases of complex mental scenarios. Descartes (1645/1996 Meditation VI) gives the example of a *chiliagon*, which is a thousand-sided figure. He argues that a chiliagon is *impossible to imagine* and distinguish from a similar thousand-and-one-sided figure by the human mind. Nevertheless, a chiliagon is *perfectly conceivable*. Similarly, to give a contemporary example (from quantum theory), it is impossible to imagine the high-dimensional configuration space of the wave function; nevertheless, high-dimensional spaces are perfectly conceivable.

Overall, I understand conceivability as a pure version of imagination, stripped from all the limitations of forming and maintaining mental images. The set of the conceivable

[32] I use the word 'image' here for convenience sake. I do not wish to suggest that blind people are incapable of imagination. What I properly mean by 'image' should be understood as sensory experience and the use of previous sensory experiences by the subject, to represent something.
[33] Proponents of *cognitive phenomenology* would disagree. They would say that concepts have their own unique phenomenology not reducible to the phenomenology of the senses. Nevertheless, even if we accept cognitive phenomenology, we can contrast conceivability to imagination as the ability to use concepts without any associated sensory phenomenology. On the other hand, imagination uses concepts in conjunction with sensory phenomenology.

encompasses the set of the imaginable. Imagination implies conceivability, while conceivability does not imply imagination. If so, anything we can imagine, we can also conceive of, but not vice-versa. This makes conceivability preferable over imagination for serious philosophical use.

Conceivability is not always a reliable philosophical tool. Chalmers (2002) differentiates between prima facie and ideal conceivability.

Prima facie (or at first glance) *conceivability* is the simplest form of conceivability. A scenario is prima facie conceivable if it *seems* conceivable without further reflection. Almost all scenarios are prima facie conceivable. However, further reflection often shows that many prima facie scenarios are inconceivable. This is evident in cases involving mathematics. For example, '5246 - 737 = 4510' might be prima facie conceivable. However, further reflection shows that it is not. '5246 - 737' equals 4509, and not 4510. Perhaps the only propositions that are not prima facie conceivable are the ones that are obviously contradictory, such as 'square-circles exist.' Thus, prima facie conceivability is not a good guide to what is rationally coherent.

Ideal conceivability, on the other hand, is conceivability upon ideal rational reflection. Ideal rational reflection can be understood in terms of the reflection of an ideal mind that has been given an infinite amount of time. However, as human beings, despite being capable of rational thought, we are neither ideal thinkers nor do we have infinite time at our disposal. Thus, as Chalmers (2002) suggests, a better way to define ideal conceivability is in terms of our human capacities, as the ability to reach a justification, "undefeatable by better reasoning," that a scenario cannot be ruled out a priori. Ideal conceivability includes all non-contradictory statements and entails prima facie conceivability. If something is conceivable after ideal a priori reflection, it must have also been conceivable at first glance. However, not all prima facie conceivable scenarios are ideally conceivable.

The clearest example of the superiority of ideal conceivability over prima facie conceivability comes in cases of mathematical propositions that seem to be conceivably both true and false. One such example is Goldbach's conjecture. Goldbach's conjecture states that every even number greater than two is the sum of

two prime numbers. Goldbach's conjecture is currently unproven; thus, we still do not know whether it is true or false, or even if it is at all provable. Chalmers argues that such cases are only prima facie conceivable, but not ideally conceivable: "there is no reason to believe that both Goldbach's conjecture and its negation are ideally conceivable." (2010: 145) If so, the prima facie conceivability of Goldbach's conjecture being either true, false, or not provable is only the result of our current non-ideal grasp of mathematics.

To summarize, ideal conceivability is a good guide to what is rationally coherent, while prima facie conceivability is not. Thus, unless otherwise specified, by conceivability, I mean ideal conceivability.

2.1.3. From Conceivability to Possibility

So far, I have only talked about conceivability. I discussed the capacity of humans to come up with scenarios that make strange propositions true. Nevertheless, while we are lost deep into armchair metaphysics inventing worlds where unicorns exist, an important question arises: 'why does this matter?' To understand how the conceivability arguments put conceivability to good philosophical use, I now turn to P-2 and the notion of possibility.

P-2 posits that if zombies are conceivable, then zombies are possible. It depends on an implicit assumption about a *link* between epistemology and metaphysics. Each of the epistemic arguments against physicalism expresses this link differently.[34] The conceivability argument posits this link in terms of conceivability.

> *Link* (conceivability version): Conceivability entails possibility.

Possibility is about metaphysics. It concerns what *could have been* the case in reality. Possibility in this sense, in the technical literature, is known as *metaphysical possibility*. Metaphysical possibility is different from logical possibility and

[34] See section 1.3.3.

physical/nomological possibility. Briefly, *logical possibility* is equivalent to conceivability and is only limited by the laws of logic; *physical/nomological possibility* is limited both by what could be Real (like metaphysical possibility) and by the actual laws of nature. Throughout the dissertation, by 'possibility,' unless otherwise specified, I always mean metaphysical possibility.

Possibility does not directly translate into actuality. To say something is possible does not mean that it is the case in reality, but only that it could have been the case. For example, we might say that flying cars are both conceivable and possible. This is likely true, although there are no flying cars on Earth right now to the best of my knowledge. However, there seems nothing in reality that makes it impossible that flying cars could exist.

One way to get a grip on 'possibility' is by spelling it out in terms of grounding. In reductive grounding cases, the Nominal entities are fully determined by their Real grounds. The Nominal entities could have been different iff their Real grounds were different. Thus, in reductive grounding cases, possibility is determined by (a) which entities are Real and (b) what can the Real entities ground. More generally speaking, for all Nominal groundees, the possibility of their obtaining is fully determined by the obtaining of their grounds (regardless of whether these grounds are Nominal or Real).

Why believe that Link is true? Mainly because there are plenty of examples, both from ordinary life and scientific practice, indicating that Link (or some version of it) is likely true. In what follows, I justify link via two examples.

My first example comes from ordinary life. Imagine that you own a gold coin. Now, imagine that someone makes the following request: 'give me the coin, but keep the gold.' How would you respond to this? Keep in mind, no cheating is allowed. If you decide to fulfill the request, you must give away *your* coin without any alterations.

The above request appears absurd. I believe that what explains the feeling of absurdity is that the scenario entailed by 'give me the coin, but keep the gold' is neither conceivable nor possible. First, assuming that I have not altered the gold coin in any way, a scenario where I have the coin's gold, yet I do not have the coin, seems

inconceivable. Moreover, practice confirms that such a scenario is impossible. No matter how hard I try, I cannot separate a coin from the material it is made from without destroying the coin. This case and countless similar cases demonstrate that inconceivably is often taken to entail impossibly in ordinary life.

The feeling of absurdity would likely disappear if the scenario were conceivable. Imagine that you have two gold coins, and you are asked to give up one of them. This scenario is both conceivable and possible and does not appear absurd. Putting all of this together, it becomes plain that, in ordinary life, conceivability is often assumed as a prerequisite of possibility. If so, Link is often both assumed and practically confirmed in ordinary life.

My second example comes from physics. Conceivability scenarios often play a key role in theoretical physics and science in general. Take Albert Einstein's (1905/1990) hypothesis about the quantum nature of light as an example, or in fact, any of the hypotheses in his 'annus mirabilis' papers. By pure abstract thought, with the right concepts, Einstein revealed something new about the world simply by tracing the logical consequences of the relations between concepts. Einstein's theories were empirically verified only later. The history of science is full of such examples.[35] Thus, scientific practice shows that conceivability can be a good guide to possibility and confirms Link.

Clarifying precisely when conceivability entails possibility is the key to the success of the conceivability argument. Link, as stated above, is overly vague and in need of further qualification. For one, it is not clear whether Link quantifies over *all* or over *some* proportions. Moreover, if it quantifies only over some propositions, which propositions are these, and why do these propositions have metaphysical significance? In the rest of this chapter, I intend to further qualify Link in the light of counter-examples raised by phenomenal concept strategists.

[35] Before telescopes and other sophisticated equipment, astronomical theories such as the ones by Ptolemaeus or Copernicus were by and large conceivability exercises. And even today, although we empirically know and can verify much more, a large portion of physics is pure theory. Theory often precedes experiment in science.

2.2. Against the Conceivability Argument

Physicalists typically respond to the Conceivability Argument by either (1) rejecting P-1 or (2) rejecting P-2.

Rejecting P-1 amounts to the claim that *zombies are inconceivable*. This is the type-A physicalist position. Rejecting P-2 amounts to the claim that *zombies are impossible*. This is the type-B physicalist position. I dedicate the rest of this chapter to defending the conceivability argument from the phenomenal concept strategy as the dominant kind of type-B physicalism.[36]

2.2.1. The Phenomenal Concept Strategy

Phenomenal concept strategy (PCS) physicalists think that zombies are conceivable yet impossible. Versions of PCS physicalism have been defended by Brian Loar (1990/1997), Christopher Hill (1997), David Papineau (1998, 2002, 2007), Brian McLaughlin (2001), Janet Levin (2002, 2007, 2019), Katalin Balog (2012a, 2012b), and Esa Diaz-Leon (2014, 2016), among others. All PCS physicalists agree that the conceivability of zombies is the result of the special nature of *phenomenal concepts*. Thus, all PCS physicalists are happy to embrace *conceptual dualism* while defending *ontological monism*.

The general attitude of the strategy is well captured by Loar:

> At the root of almost all weird positions in the philosophy of mind lies this rather elementary and unremarkable conceptual fact, blown up into a metaphysical problem that appears to require an extreme solution. (1990/1997: 609)

By this "rather elementary and unremarkable conceptual fact," Loar has in mind that phenomenal concepts refer to physical brain states in a special and direct way that

[36] I do not discuss type-A physicalism in this dissertation primarily because I consider type-A physicalism to be highly implausible. I find it evident that zombies are conceivable and thus disagree with type-A physicalists at a foundational level.

gives rise to Consciousness Gap. By "weird positions in the philosophy of mind," Loar has in mind all positions on the metaphysics of consciousnesses that are not versions of physicalism.

Phenomenal concepts are understood, by both anti-physicalists and PCS physicalists, as the concepts we use when thinking about experiences in terms of *how they feel*. PCS physicalists argue that both phenomenal and physical concepts refer to the same physical brain states. However, in contrast to physical concepts, which characterize their referents in physical terms (e.g., 'c-fibers firing'), phenomenal concepts do not refer via any physical descriptions. Thus, according to PCS physicalists, although experiences are physical, and physical and phenomenal concepts *co-refer* to the same physical states, there is *no a priori connection* between phenomenal and physical concepts. The lack of such an a priori connection results in Consciousness Gap and the conceivability of zombies.

PCS physicalism is sometimes called 'non-reductive physicalism.' However, it is crucial to note that this is only meant in the *epistemic* sense of 'reduction,' i.e., in the sense that PCS physicalists are type-B physicalists. Remember, in my usage, the unqualified term 'reduction' refers to *metaphysical* reduction, and PCS physicalists are reductive physicalists in the metaphysical sense.

PCS physicalists see the metaphysical relation between physical and phenomenal states as an *a posteriori necessary truth*. Metaphysically, on their view, 'P is Q' (or 'Q is grounded in P') is necessarily true in all possible worlds that are physically identical to our world. However, this necessary truth is only knowable a posteriori. Thus, although it is conceivable that P and Q can come apart, this is not possible. To better understand the notion of a posteriori necessity, I make a quick recourse to the work of Kripke, who discovered it.

2.2.2. A Posteriori Necessities

Saul Kripke introduced the novel philosophical idea of *a posteriori necessities* in *Naming and Necessity* (1972/1980). A posteriori necessities are necessarily true propositions whose truth is knowable only *a posteriori*. Before Kripke, it was

standardly assumed by philosophers that necessary truths are always a priori. Typically, these were limited to mathematical and logical truths. To cut a long history lesson short, Kripke argued that two terms could co-refer to the same entity and do so *necessarily* even if they are a priori disconnected. Metaphysically, in such cases, there is only one referent. Epistemically, however, the fact that these terms co-refer is only known *a posteriori*. Truths involving such co-referring terms are *a posteriori necessities*.

Kripke's examples of a posteriori necessities often involve a commitment to *scientific essentialism*, namely, the thesis that science discovers the essential properties of natural kinds. *Essential properties* are the properties that an entity necessarily has across all possible worlds. For example, let us assume for the sake of argument that H_2O is the essence of water.[37] *Accidental properties*, on the other hand, are the contingent properties that an object might or might not have across possible worlds. For example, liquidity and transparency are plausibly accidental properties of water. The identity between natural kinds and their essences is a paradigmatic example of a posteriori necessity.

All cases of a posteriori necessity involve two terms that are *rigid designators*. Rigid designators, in Kripke's terminology, are terms that always refer to the same entity across all possible worlds in which that entity exists. Thus, all the natural kind terms and the terms designating their essences are rigid designators. The 'water is H_2O' case is one such instance of an a posteriori necessity between a natural kind ('water') and its essence ('H_2O'). There are countless other cases: 'heat is molecular motion,' 'gold is Au,' and so on. Generalized, all a posteriori necessities between natural kinds and their essences involve two different co-referring terms as rigid designators.

Kripke's original thesis is about *words*; however, in the current debates, PCS physicalists typically understand a posteriori necessities to be about *concepts*. To illustrate the current debates, I will henceforth, for simplicity of exposition, extend Kripke's views to include concepts. With this in mind, I define two broad classes of

[37] Essence terms are best seen as *placeholders* for whatever essence a complete physics would discover. Of course, H_2O molecules are not fundamental entities. Plausibly, they are grounded in some quantum states, or whatever states the complete physics of the future might describe.

concepts that cover the Kripkean cases of a posteriori necessities. These are Appearance and Essence concepts.

> *Appearance concepts*: A concept is an Appearance concept iff it refers via the referent's accidental properties.[38]
>
> *Essence concepts*: A concept is an Essence concept iff it refers via a description of the referent's essential properties.

Appearance and Essence concepts are *different kinds of concepts*. There is *no a priori connection* between Appearance and Essence concepts. The only way to know that Appearance and Essence concepts co-refer is a posteriori. Nevertheless, if Kripke is right, both Appearance and Essence concepts *refer to the same entity*, and they do so across all possible worlds. This is because both Appearance and Essence concepts are *rigid designators*.

In the 'water is H_2O' case, the <water> concept (here and henceforth, I use the '<…>' notation to designate concepts) is an instance of an Appearance concept, while <H_2O> is an instance of an Essence concept. <water> refers to H_2O via its accidental properties (such as wet, transparent, thirst-quenching, and so on), namely by how H_2O is *experienced* by observers on Earth. <H_2O> refers to H_2O by a description of H_2O. In general, in all cases of a posteriori necessities involving natural kinds, the left-hand side ('water,' 'heat,' 'gold,' etc.) typically corresponds to an Appearance concept. The right-hand side ('H_2O,' 'molecular motion,' 'Au,' etc.) typically corresponds to an Essence concept.

Essence concepts are rigid designators since the essence is what identifies an entity across all possible worlds. Appearance concepts are likewise rigid designators since they refer to the entity that is the bearer of the accidental properties that serve as their mode of presentation. This is why, even without knowledge of H_2O (as was the case

[38] Note, Appearance concepts are not limited to visual appearances. I have named them as such only because visual experience seem to be the most common way to refer via accidental properties.

before the discovery of H_2O), if Kripke is right, <water> still rigidly designates H_2O across all possible worlds.

The different ways Appearance and Essence refer to the same entity accounts for their a priori disconnection. The lack of an a priori connection means that *it is conceivable* that the referent of co-refiring Essence and Appearance concepts refer to different entities. Thus *it is conceivable* that 'water is not H_2O,' although this is *not possible*. Thus, if this is true, in cases of a posteriori necessities involving Appearance and Essence concepts, *conceivability does not entail possibility*.

In summary, if Kripke is right, a posteriori necessities are exceptions to Link. In cases of a posteriori necessities, conceivability does not entail possibility because despite being a priori disconnected, the involved concepts nevertheless co-refer. We can rigidly refer to the same entity either via its appearance or by describing it in terms of its essence. Thus, in these cases, although it might be conceivable that the referent of an Appearance concept is not the referent of an Essence concept, this does not translate to a genuine possibility.

2.2.3. Phenomenal Concepts Are Special

By now, it might be obvious why a posteriori necessities matter for the metaphysics of consciousness. PCS physicalists want to say that, just like 'water is not H_2O' might be conceivable but not possible, 'P and ~Q' is likewise *conceivable but not possible*. Thus, according to PCS physicalists, the identity (or grounding) between Q and P is an instance of an a posteriori necessity.

It is worth noting that a posteriori necessities are not, by themselves, a counter-example to the conceivability argument. Kripke himself argued *against physicalism* by showing that 'P is Q' cannot be an a posteriori necessity. In Kripke's view, *both* physical and phenomenal concepts (P and Q) are Essence concepts. Physical concepts are Essence concepts since evidently, they refer via a physical description of the essence of their physical referents. Phenomenal concepts, on the other hand, are likewise Essence concepts since, in Kripke's view, they refer via a phenomenal description of the essence of phenomenal states. The essences of experiences, in his

view, are what phenomenal states feel like. Thus, if Kripke is right, zombies are conceivable not because we think of the same essence in two different ways but because we think of *two different essences*. Thus, Kripke uses the special nature of phenomenal concepts to argue that physicalism is false.

PCS physicalists *accept* Kripke's doctrine of a posteriori necessities and his thesis that phenomenal concepts are not Appearance concepts. However, they *reject* Kripke's doctrine of phenomenal concepts as Essence concepts, or at least his assumptions about the features of Essence concepts.

According to PCS physicalists, phenomenal concepts are special among the concepts participating in a posteriori necessities. This is because, in their view, phenomenal concepts are *neither* Appearance concepts nor typical Essence concepts.

All PCS physicalists *reject* an implicit assumption made by Kripke, an assumption that Loar (1990/1997: 600) calls the '*Semantic Premise*.' Roughly, in the terminology I am using here, the Semantic Premise can be paraphrased as follows.

> *Semantic Premise*: A posteriori necessities always involve at least one Appearance concept.[39]

All PCS physicalists reject the Semantic Premise. The rejection of the Semantic Premise is crucial for the success of PCS since otherwise 'P & ~Q' cannot be an a posteriori necessity (assuming, as PCS physicalists standardly do, that P involves Essence concepts).

All PCS physicalists agree that phenomenal concepts have three key features:

1. *Conceptual isolation:* phenomenal concepts are a priori disconnected from all physical concepts.

[39] For reference, Loar's original formulation of the Semantic Premise is: "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." (1990/1997: 600)

2. *Rigidity:* phenomenal concepts are rigid designators of physical properties in the brain.
3. *Directness:* phenomenal concepts refer without a contingent mode of presentation involving accidental properties.

Conceptual isolation accounts for the conceivability of zombies, despite their impossibility. Rigidity accounts for 'P is Q' being an instance of an a posteriori necessity. Directness accounts for phenomenal concepts not being Appearance concepts.

What makes phenomenal concepts different from Appearance and typical Essence concepts? Goff summarizes and classifies the logical geography of the debate clearly and simply. He distinguishes four types of concepts based on what they *reveal* about the essences of their referents (2017: 102):

> *Transparent Concepts: "*A concept referring to entity *e* is transparent just in case it reveals the essence of *e*."
>
> *Translucent Concepts*: "A concept referring to *e* is translucent just in case it reveals some significant aspect of the essence of *e*, but not the complete essence."
>
> *Mildly Opaque Concepts*: "A concept referring to *e* is mildly opaque just in case it is not transparent or translucent, but it transparently reveals (i.e., reveals under transparent concepts) significant accidental properties of *e*, typically properties that uniquely identify *e* in the actual world."
>
> *Radically Opaque Concepts*: "A concept referring to *e* is radically opaque just in case it does not transparently reveal any significant properties (either essential or accidental) of *e*."

Transparent and translucent concepts correspond to Essence concepts (in my terminology). Plausible examples of transparent concepts include pure physical concepts, mathematical and logical concepts, and, as Goff argues, *phenomenal concepts*. As an example of a translucent concept, Goff (2017: 102) mentions <being a sphere the same mass as the Earth>. In this concept, 'being a sphere' reveals part of the referent's essence, while 'the same mass as the Earth' does not reveal anything essential (at least not for the average user).

Mildly opaque concepts correspond to Appearance concepts. Most ordinary concepts that refer via a contingent mode of presentation, such as <water>, <heat>, <gold>, etc., are mildly opaque.

Radically opaque concepts are neither Essence nor Appearance concepts. According to Goff, radically opaque concepts serve as *blind pointers* "allowing the concept user to latch onto an entity without conceiving of it in terms of any of its features" (2017: 91). Goff (2011) mentions <that stuff>, <that guy I met at a party>, and <the property Kev is thinking about right now> as plausible examples of radically opaque concepts.

We are now finally in a position to fully understand the PCS physicalist position.[40] PCS physicalists can be divided into two camps based on their account of phenomenal concepts. Here, inspired by Goff's characterization of the debate, I call these two camps 'Blind Pointers' and 'Dual Carvers.'

Blind Pointers argue that phenomenal concepts are *radically opaque concepts*. In contrast, *Dual Carvers* argue that phenomenal concepts are *special transparent concepts* (special Essence concepts in my terminology). I analyze their views in more detail in the next two sections.

[40] There are still many fine-grain differences among the PCS physicalist's accounts of phenomenal concepts that cut across different dimensions, that I have not presented here. For a detailed overview of these differences, see Chalmers (2007: 171–3).

2.2.4. The Blind Pointers

Blind Pointers argue that phenomenal concepts are *radically opaque*. If so, phenomenal concepts refer to physical brain states without revealing *neither accidental nor essential properties* of their referents. Thus, phenomenal concepts *are not* Essence concepts.

Blind Pointers are the majority among PCS physicalists. They include Loar (1990/1997), Hill (1997), (2001), Papineau (2002, 2006, 2007), McLaughlin, and Balog (2012a, 2012b), among others.

Loar (1990/1997) (who is widely regarded as the originator of the PCS) is explicit that phenomenal concepts *are not* Appearance concepts. This is because: "Phenomenal concepts [...] do not conceive their reference via contingent modes of presentation. And so they can be counted as conceiving phenomenal qualities directly." (1990: 609)

Loar thinks that both phenomenal and physical concepts *co-refer directly* (without a contingent mode of presentation) to the same physical essences. In his view, both physical and phenomenal concepts provide a "direct grasp of essence." (1990/1997: 609) However, they do so *differently*. Physical concepts refer by revealing how their referents are internally constituted; phenomenal concepts refer demonstratively, by experiencing the essence. Loar points out that *only physical concepts are transparent* because only physical concepts reveal the internal constitution of their referents. This, in Loar's view, explains the *conceptual isolation* of phenomenal concepts from physical concepts.

In Loar's own words:

> What generates the problem is not appreciating that there can be two conceptually independent 'direct grasps' of a single essence, that is, grasping it demonstratively by experiencing it, and grasping it in theoretical terms. The illusion is of *expected transparency*: a direct grasp of a property ought to reveal how it is internally constituted, and if it is not revealed as physically constituted, then it is not so. The mistake

> is the thought that a direct grasp of essence ought to be a transparent grasp, and it is a natural enough expectation. (1990: 609)

Thus, According to Loar, phenomenal concepts are neither transparent nor translucent. Moreover, they have no contingent mode of presentation; thus, they are not mildly opaque. Therefore, phenomenal concepts *must be radically opaque*.

Other Blind Pointer physicalists share similar views about the radical opacity of phenomenal concepts. For example, Balog argues that phenomenal concepts do not "reveal anything about the metaphysical nature of phenomenality." (2012a: 31) While according to McLaughlin (2001: 324), "[phenomenal concepts] do not conceptually reveal anything about the essential nature of phenomenal properties: they simply name or demonstrate them." Papineau is perhaps the most radical among the Blind Pointer physicalists. He argues that *all* basic (atomic) concepts are radically opaque. Thus, on his view, there are no transparent and translucent basic concepts. As he puts it: "I don't recognise any way in which the mind 'captures' something, apart from simply referring to it." (Papineau, 2006: 106)

Despite their differences, all Blind Pointer physicalists agree: phenomenal concepts are radically opaque concepts. Thus, phenomenal concepts are neither Appearance nor Essence concepts. They refer directly, are conceptually isolated, and are rigid designators of physical brain states. They form a posteriori necessities with co-referring physical concepts. These a posteriori necessities are conceivably false, although their falsity is not possible. If so, the conceivability of 'P & ~Q' *does not entail possibility*. If so, the conceivability argument fails, and physicalism is true.

2.2.5. The Dual Carvers

Dual Carvers agree with anti-physicalists (such as Kripke, Chalmers, and Goff) that phenomenal concepts are *transparent*. If so, phenomenal concepts refer to physical brain states and *reveal* the *essential properties* of their referents. If so, phenomenal concepts are Essence concepts. Nevertheless, Dual Carvers argue that physicalism is true.

The Dual Carvers are a growing minority among PCS physicalists. They include Diaz-Leon (2014, 2016), Trogdon (2017), and Levin (2019), among others.

Aspects of the Dual Carvers view already present in Loar's work. As I already mentioned, Loar argued that there could be multiple ways to grasp a single essence directly. By this, he meant that there could be a posteriori necessities involving either an Essence and a radically opaque concept or two radically opaque concepts. Dual Carvers take this idea a step further. They argue that, moreover, there can be a posteriori necessities *involving two transparent concepts*.

Like the Blind Pointers, the Dual Carvers reject the Semantic Premise. However, unlike the Blind Pointers, the Dual Carvers reject the Semantic Premise because it entails that a posteriori necessities can *never* involve two transparent concepts. I refer to this thesis as follows.

> *Single Carving*: For any entity E, there is only one privileged way to refer to E via a description of its full essence.

Thus, Dual Carvers *reject both* the Semantic Premise and Single Carving. Single Carving is implicitly assumed by Kripke, Chalmers, Goff, and other anti-physicalists.[41] On their view, there can be only one transparent concept for each essence. If Single Carving is true, transparent concepts are epistemically special and always conceptually isolated, but only because two transparent concepts *can never co-refer*. In contrast, against anti-physicalists, Dual Carvers argue that some transparent concepts can both *co-refer* and *be conceptually isolated*.

For example, Diaz-Leon (2014) argues that the concept <bachelor> co-refers with concepts such as <unmarried man> and <unmarried male Homo Sapiens>. On her view, all these concepts are either transparent or translucent. Moreover, all of these

[41] Chalmers assumes both the Semantic Premise and Single Carving. Goff assumes only Single Carving.

concepts are *conceptually distinct*, meaning they are different concepts. Thus, she concludes, multiple transparent concepts can co-refer, and Single Carving is false.

Levin (2019) gives two examples involving geometric shapes against Single Carving. Levin's first example involves <sphericity> (a descriptive concept of a sphere) and <that shape> (a recognitional-demonstrative concept created by pointing at a sphere). She argues that <sphericity> and <that shape> are co-referring, transparent, and conceptually isolated concepts. Levine's second example is inspired by the famous '*Molyneux's Question*.'[42] In this example, she argues that <tactile-sphere> (a tactile recognitional concept of a sphere) and <visual-sphere> (a visual recognitional concept of a sphere), are co-referring, transparent, and conceptually isolated concepts.

To sum up, all Dual Carvers agree that phenomenal concepts are transparent concepts yet reject Single Carving. In their view, phenomenal concepts refer directly, are conceptually isolated, and are rigid designators of physical brain states. They form a posteriori necessities with co-referring physical concepts that are likewise transparent. These a posteriori necessities are conceivably false, although their falsity is not possible. Thus, in their view, the conceivability of 'P & ~Q' *does not entail a possibility*. If so, the conceivability argument fails, and physicalism is true.

[42] From Locke (1694/1975). In the original formulation, Locke asks whether a 'Man born blind, and now adult,' who has a tactile concepts of geometrical shapes, could visually recognize these shapes, if he became able to see. Locke's answer is 'no.'

2.3. Defending the Conceivability Argument

2.3.1. The Transparency Conceivability Argument

If either the Blind Pointers or the Dual Carvers are right, physicalism remains true despite Consciousness Gap and the conceivability of zombies. On both views, the specialness of phenomenal concepts results in dualism at the level of concepts while preserving physical monism at the level of ontology.

In response to the PCS, anti-physicalists typically follow Kripke and argue that phenomenal concepts *are transparent*. Chalmers (1996, 2003, 2010) and Goff (2017) have both given improved versions of the conceivability argument using this approach.

In a nutshell, the key question in the debate between PCS physicalists and anti-physicalists is: 'what do phenomenal concepts refer to?' Physicalists say *physical states*. Anti-physicalists say *non-physical experiences*.

Chalmers and Goff accept there are a posteriori necessities. However, they disagree with PCS physicalists about the nature of phenomenal concepts. According to Chalmers and Goff, phenomenal concepts cannot be both conceptually isolated from physical concepts and co-refer with physical concepts. Instead, in their view, phenomenal concepts refer to non-physical experiences. Thus, the putative a posteriori necessity 'P is Q' is false, and physicalism is likewise false.

In the rest of this section, I focus mainly on Goff's approach to the conceivability argument. I find Goff's version to be simpler and more versatile than Chalmers' version, without a loss of power. Chalmers' influential version rests on the complex two-dimensional semantic framework. Goff's version does not. Thus, since the technical complexities of the 2-D framework are not essential to the conceivability argument, I prefer Goff's approach to the argument.

In Goff's (2011, 2017) view, some a posteriori necessities *are* a counter-example to Link. However, against PCS physicalists, Goff argues that the consciousness case is

different. In the case of consciousness, namely in the case of 'P & ~Q,' conceivability does entail possibility.

Goff argues that conceivability entails possibility iff the involved concepts are transparent. This is because only transparent concepts accurately represent what it takes for their referents to exist.[43] In Goff's own words:

> When conceiving of entity e under a transparent concept, one understands (or has a priori access to) the nature of e: one understands what it is for that e to be part of reality (to exist if e is an object, to be instantiated if e is a property, to take place if e is an event, etc.). (2017: 96–7)

In Goff's view, conceivability is an imperfect guide to possibility unless we use transparent concepts. This is because, in cases involving non-transparent concepts, our concepts are imperfect representations of reality. If so, all exceptions to Link are due to "our ignorance of the nature of what we are conceiving." (Goff, 2017: 99)[44]

Both *pure physical and phenomenal concepts are transparent,* according to Goff. The transparency of pure physical concepts (the concepts that refer to physical entities via purely causal and/or structural descriptions) is fairly uncontroversial. Many physicalists are happy to accept that pure physical concepts are transparent. The transparency of phenomenal concepts is more controversial.

In Goff's view, the transparency of phenomenal concepts is the *best explanation* of the high certainty we have in introspective beliefs (what Goff calls 'super-justification'). For example, my belief in the introspective fact that I experience *something* has much greater certainty than any fact about the external world.[45] Similarly, I believe that the

[43] It is important to note that although transparent concepts reveal what it takes for their referents to exist, they do not entail the *existence* of their referents. In other words, if a concept is transparent, that does not mean that the concept is satisfied. For example, I might have a transparent concept <particle-x> revealing the essence of an imaginary fundamental particle, although no such a particle exists, has existed, or can exist in our world.

[44] The thesis that some concepts reveal the essences of their referents and therefore give us access to metaphysical knowledge is not new in philosophy. Descartes notion of *clear and distinct perceptions*, and more recently, Lowe's (2012) notion of *fully adequate concepts*, express a similar idea.

[45] Augustine and Descartes have also made this point.

experience of red is different from the experience of blue, with much greater certainty than any fact about the external world. In Goff's view, we cannot make sense of the certainty of these beliefs unless phenomenal concepts are both transparent and satisfied.

Goff (2017: 100) proposes a restricted version of Link based on the above arguments. I paraphrase his version of Link as follows.

> *Transparent Link* (conceivability version): conceivability entails possibility iff all the concepts involved in a conceivable proposition are transparent.

Transparent Link is the key premise in Goff's (2017: 101) version of the conceivability argument, or what he calls the '*Transparency Conceivability Argument*.' I paraphrase his Transparency Conceivability Argument as follows.

P-1. 'P & ~Q' is conceivable.
P-2. 'P & ~Q' involves transparent concepts.
P-3. Transparent Link is true.
C-1. 'P & ~Q' is possible. [from P-1, P-2, and P-3]
P-4. If 'P & ~Q' is possible, physicalism is false.
C-2. Physicalism is false. [from C-1 and P-4]

Given the Transparency Conceivability Argument, type-A physicalists once again reject P-1. However, PCS physicalists might reject P-2 or P-3 of the Transparency Conceivability Argument: Blind pointers reject P-2, while Dual Carvers reject P-3.

I argue against the Blind Pointers and the Dual Carvers. Against the Blind Pointers, I argue that there are no radically opaque concepts. Against the Dual Carvers, I argue that even if there are multiple co-referring transparent concepts, these cannot be conceptually isolated. If my arguments against the Blind Pointers and the Dual Carvers are sound, PCS physicalism is false.

2.3.2. Against the Blind Pointers

The Blind Pointers' account hangs on the existence of radically opaque concepts. My view is that there are no radically opaque concepts.[46] I argue that concepts must refer via either accidental or essential properties. Since radically opaque concepts are hypothesized to refer via neither accidental nor essential properties, they cannot be concepts.

My argument against radically opaque concepts is as follows.

P-1. A concept must refer either via accidental or essential properties.
P-2. Radically opaque concepts refer via neither accidental nor essential properties.
C. There are no radically opaque concepts.

The argument is simple and valid. Proponents of radically opaque concepts accept P-2 and thus must reject P-1; thus, P-1 requires further justification.

Why believe P-1? First, because it is highly plausible that a concept must refer in order to be a concept. To refer, a concept must pick out a referent. If so, 'picking out a referent' is a necessary property of concepts. 'Picking out a referent' does not entail that the concept's referent exists, but only that, were the referent to exist, the concept would successfully pick it out. If a mental state does not pick out a referent, that mental state cannot be a concept. Second, assuming the preceding is true, the only way for a concept to pick out a referent is in virtue of some of the referent's properties. An entity can have either accidental or essential properties and no other properties. If so, P-1 must be true.

Rejecting P-1entails either that (a) there are concepts that do not pick out a referent, or (b) that there are concepts that pick out their referent via neither accidental nor essential properties. Option (a) seems to go against the nature of concepts, and

[46] Chalmers (2010) has the same opinion. In his view, all concepts must have a primary and secondary intension. Radically opaque concepts have no primary intension, but only a secondary intension. Thus, they are incoherent in the context of Chalmers' two-dimensional framework.

moreover, it is not how Blind Pointers see phenomenal concepts. This leaves option (b) as the Blind Pointers' response to P-1.

Proponents of radically opaque concepts must *explain* how a concept could refer if it refers via neither accidental nor essential properties. Mildly opaque concepts pick out referents via accidental properties. Transparent and translucent concepts pick out referents via essential properties. However, if a concept such as <that stuff> is indeed radically opaque, it is not clear to me how it could pick out anything at all. What other worldly features are there that might allow this concept to latch onto a referent? In the absence of an explanation of the above, I take it as highly plausible that P-1 is true.

What about all the putative cases of radically opaque concepts? If sound, does my argument entail that all of these concepts are somehow defective? I argue that it does not.

My view is that belief in radically opaque concepts arises due to a confusion between how some concepts are expressed *in language* vs. how they are used *in thought*. I am happy to agree that some concepts are expressed in language in a way that reveals nothing about their referents. However, as I will soon show with examples, I believe that those same concepts reveal *something* about their referents when used in thought. If so, all putative radically opaque concepts refer because they contain a hidden reference to either accidental or essential properties. If so, radically opaque concepts are, in fact, either mildly opaque, translucent, or even transparent concepts whose meaning has been obfuscated by language.

Goff (2011) gives the examples of <that stuff>, <the property Kev is thinking about right now>, and <that guy I met at a party> as putative radically opaque concepts. Papineau and Goff (2014) argue that proper names might be radically opaque.

The concept <that stuff> is typically mentioned in the literature followed by a clause stating its referent. To me, this indicates that <that stuff> gives an impression of radical opacity because it is *obscurely labeled*, based on how we sometimes *talk* about things, but, in fact, is not a radically opaque concept. I could indeed refer to water (or almost anything) by calling water 'that stuff' at the *level of language*. However, at the

level of thought, whenever I call something 'that stuff,' I have *some specific stuff* in mind, instead of just any stuff. Thus, when I call something 'that stuff,' I always seem to have at least some of its properties in mind, and thus I understand which stuff I am talking about.

What about the reverse case, where I am not the concept creator, but instead, I hear about 'that stuff' from someone else? Would not 'this stuff' result in a radically opaque *for me* in such a case? I think not.

Imagine that someone (let us assume her name is Anne) were to ask me: 'what do you think about that stuff?'. Assume that from Anne's point of view, 'that stuff' means water. However, from my point of view, I find it impossible to understand Anne. From my point of view, 'that stuff' could refer to *any* stuff. If my preceding assumptions are right, at this stage, for me, 'that stuff' does not yet correspond to a concept. To try to understand Anne, I must *make a guess* about what 'this stuff' could have meant. For example, I might assume that Anne means 'some object in the close vicinity.' My guess might be completely wrong; nevertheless, it latches onto reality via accidental or essential properties and thus results in concept. This concept is not radically opaque.

The above analysis shows that linguistic expressions such as 'that stuff' might not correspond to radically opaque concepts. Instead, they might either not correspond to concepts or correspond to concepts that are not radically opaque. If this explanation is plausible and given the mystery of how radically concepts refer, it is reasonable to assume that there are no radically opaque concepts.

A similar analysis applies to proper name concepts such as <Cicero> or <Tully>. Concept users could indeed use a name *in language* even if they have little or no information about the referent. However, *in thought,* it seems that every proper name concept always has at least *some* information associated with it, even if this information is false or based on a guess. At the very least, it seems that every proper name concept contains the information that it refers to a single individual with that specific name.

What about Goff's examples of <the property Kev is thinking about right now> and <that guy I met at a party>? <the property Kev is thinking about right now> reveals that the referent is a property and that Kev is thinking about it right now. 'Is a property' seems to be an essential feature of the referent, while 'Kev is thinking about it right now' seems to be an accidental feature. Thus, this concept is at least mildly opaque and perhaps even translucent. Similarly, <that guy I met at a party> reveals that it refers to a guy. If 'being a guy' is an essential property of the person I am referring to, the concept might be translucent.

In summary, all the cases of reputed radically opaque concepts I examined can be explained via concepts that are not radically opaque. I assume that the same explanation applies to all other cases of reputed radically opaque concepts. If so, and given everything else I said in this section, I ask any readers that might be unconvinced: why believe in radically opaque concepts?

2.3.3. Against the Dual Carvers

The Dual Carvers' account hangs on (1) the *rejection* of Single Carving and (2) the *acceptance* of the stronger thesis that concepts that dual carve (namely co-referring transparent concepts) can be conceptually isolated. I call the thesis entailed by (1) 'Transparent Co-Referring.' Moreover, I call the thesis assumed in (2) 'Transparent Isolation.'

> *Transparent Co-Referring*: transparent concepts can co-refer.
>
> *Transparent Isolation*: co-referring transparent concepts can be conceptually isolated.

The Dual Carvers' account is best seen as the conjunction of Transparent Co-Referring and Transparent Isolation.

Dual Carving: the conjunction of Transparent Co-Referring and Transparent Isolation.

I argue that even if Transparent Co-Referring is plausible, Transparent Isolation is not. Since Dual-Carving needs both, Dual-Carving is false.

Single Carving is prima facie more attractive than Dual Carving. The attractiveness of Single Carving comes from the way we think of essences. Essences are *unique*. Each entity has only *one essence*. By definition, an essence is what makes an entity what it is, qua self. Transparent concepts reveal the full essence of their referent. Thus, if essences are unique, it is reasonable to think that the concepts that give a full grasp of essences are also unique. Transparent concepts carve nature at the joints, and nature *only has one set of joints*. If so, it is reasonable to assume that each essence should correspond to only one transparent concept.

Single Carvers and Dual Carvers agree that Reality has a privileged structure. Moreover, they agree that Reality's privileged structure is epistemically accessible. However, unlike Single Carvers, Dual Carvers think there are *multiple equally good ways* to carve nature at the joints conceptually. Thus, if Dual Carving is true, all (or at least some) essences can be grasped in multiple different yet equally good ways via different transparent concepts. Thus, Dual Carving entails multiple equally good ways to grasp what it is for something to exist, i.e., to be a part of Reality.

If Dual Carving is true, it likely true due to some *conceptual features*. This is because it is highly unlikely that Dual Carving is a feature of Reality itself. Dual Carving being a feature of Reality itself, entails that Reality has multiple overlapping joints. Moreover, it entails that at least some entities have multiple overlapping essences. This seems absurd and goes against the whole idea of essence. Dual Carvers accept the existence of essences; thus, I suspect they would not like to endorse the 'multiple essences' thesis.

What kind of conceptual feature could account for Dual Carving? I focus on *representation* as one candidate for such a conceptual feature. By 'representation,' I have in mind the ability of concepts to *represent* their referents *via their content*. I

assume that each concept has content, and the content represents a referent. In plain terms, the representation is how we think about the referent.

It is plausible and somewhat common to have *multiple representations* of the same entity. Each representation might reveal some essential properties of the referent. For example, there might be a statue and a painting portraying the same person. If they are any good, plausibly, both the statue and the painting reveal at least some of that person's properties. Cases such as these are best seen as analogs of translucent concepts. Although the two representations reveal some of their referents' essential properties, they fail to reveal the full essence.

If transparent concepts are representations, they are representations that reveal the full essence. If so, transparent concepts are *perfect representations*. With this in mind, the key question for the Dual Carver becomes: can two representations be *both perfect and different*?

Words from different languages with the same meaning might be representations that are both perfect and different. For the sake of example, assume that <snow> is a transparent concept. We can express the concept <snow> in different languages using different words. For example, we can say 'snow' in English or 'naige' in French. Now, for the sake of argument, suppose that there are *different concepts* corresponding to each of these words. Suppose that the concept <snow> corresponds to 'snow' while the concept <naige> corresponds to 'naige.' If this is so, then <snow> and <naige> are different concepts that are *transparent* and *co-refer* to snow. If so, Transparent Co-Reference must be true.

The above example shows that some transparent concepts might co-refer. These concepts have the same meaning yet represent the referent differently. If <snow> and <naige> are transparent and I understand them, I must also understand that they have the same meaning. Namely, I must understand they both refer to snow. This is because transparent concepts reveal the full essence of their referent. Thus, for all co-referring transparent concepts, I must know the full essence of their referent.

Belief in Transparent Isolation seems to arise due to a confusion between *representation* and *revelation*. Representation is about how the meaning is presented to us, qua concept users. Revelation, on the other hand, in the case of transparent concepts, is about grasping the meaning, namely about fully understanding what it means for the referent to exist. Even if there are multiple ways to represent the same referent (e.g., <snow> and <naige>), in transparent concepts, all representations must reveal *the same* full essence. In other words, there might be multiple ways to refer to the same referent transparently, but there can be *only one way* to grasp what it means for the referent to exist. Thus, Transparent Isolation is false.

Formally stated, my argument against Transparent Isolation is as follows.

P-1. Transparent Isolation entails that there can be different revelations of the same essence.
P-2. An essence can be revealed in only one way.
C. Transparent Isolation is false.

If sound, the above argument shows that Transparent Isolation is false even if Transparent Co-Referring is true. Dual-Carving requires both Transparent Co-Referring and Transparent Isolation. Thus, if Transparent Isolation is false, Dual-Carving is likewise false. A careful analysis of examples from the literature likewise suggests that Transparent Isolation is false.

None of Diaz-Leon's (2014) examples support Transparent Isolation. Diaz-Leon argues that many transparent concepts co-refer to bachelor: <bachelor>, <unmarried man>, <unmarried male Homo Sapiens>, etc. Diaz-Leon might be right about this; however, her examples are not sufficient to establish Dual Carving. The concepts she mentions are *not* conceptually isolated. It is clear to anyone who grasps these concepts that they refer to the same entity, namely to a bachelor. Thus, Diaz-Leon's examples only establish Transparent Co-Referring but not Transparent Isolation. Dual Carving requires both.

Levin's (2019) gives the examples of (a) <that shape> and <sphericity>, and (b) <tactile-sphere> and <visual-sphere>. She argues that the concepts in both (a) and (b)

are transparent, co-refer, and are conceptually isolated. If she is right, her examples seem to establish both Transparent Co-Referring and Transparent Isolation.

In response, I argue that, in case (a), Levin is right about Transparent Co-Referring but is wrong about Transparent Isolation. Moreover, in case (b), Levin is wrong about both Transparent Co-Referring and Transparent Isolation.

In the case of (a) <that shape> and <sphericity> must both reveal to me the full essence of sphericity in order to be transparent. Otherwise, I will not understand that <that shape> refers to a sphere instead of a triangle. However, if both <that shape> and <sphericity> reveal the full essence of sphericity, they cannot be conceptually isolated. They summon the same essence to my mind's eye. In the case of (b) <tactile-sphere> and <visual-sphere> seem to grasp different aspects of what it means to experience a sphere phenomenally. In other words, it is plausible that these concepts co-refer without revealing the essence of sphericity. If so, these concepts are mildly opaque and not transparent. If so, they are conceptually isolated simply because the visual experience and the tactile experience of a sphere are distinct. There is nothing mysterious about the conceptual isolation between mildly opaque concepts.

Before ending the chapter, I must note that Dual Carvers are not out of the water even if my argument against Transparent Isolation fails. This is because Dual Carving appears *incompatible with physicalism*. Transparent Co-referring entails that there can be multiple equally good ways to think about the same referent. This means that both <c-fibers firing> and <pain> are equally good ways to think about the same referent. If so, <c-fibers firing> and <pain> can be used *interchangeably*. The same applies to all the concepts in P and Q. If both P and Q are equally good ways to think about the world, P and Q are *interchangeable*. However, if both the experiential and physical ways of thinking are equally good, nature could be described *equally well as experiential and as physical*. If so, the physical facts cannot be more fundamental than the experiential facts. Instead, both the physical and the experiential facts must be fundamental. This is not physicalism! Thus, Dual Carving is incompatible with physicalism. If Dual Caring is true, physicalism must be false.

Chapter III:
Against Non-Reductive Physicalism

"Whatever we conceive as existent, we can also conceive as non-existent. There is no being, therefore, whose non-existence implies a contradiction. Consequently there is no being whose existence is demonstrable." (Hume, 1779/1997 part 9)

"[M]etaphysics as I understand it is about what grounds what. It is about the structure of the world. It is about what is fundamental, and what derives from it." (Schaffer, 2009a: 379)

"What kind of world could be identical to ours in every last microphysical fact but be biologically distinct? [...] Such a world seems quite inconceivable." (Chalmers, 1996: 73)

3.1. Ground Functionalism

3.1.1. Non-Reductive Physicalism

Jonathan Schaffer (2017, forthcoming) defends ground functionalism, a version of type-B physicalism where explanatory gaps are everywhere, in all instances of grounding. Like other type-B physicalists, Schaffer accepts Consciousness Gap and rejects Cosmic Scrutability.

> *Consciousness Gap*: The physical facts (as grounds) do not a priori entail the experiential facts (as groundees).
>
> *Cosmic Scrutability*: The microphysical facts (as grounds) a priori entail all other facts (as groundees).

However, unlike PCS physicalists, Schaffer is not concerned with the nature of phenomenal concepts. Instead, Schaffer argues that Consciousness Gap is not a special case. In his view, there are explanatory gaps even in assumed paradigmatic cases of a priori entailment involving only physical facts. If so, Consciousness Gap is not special since no explanatory gap is special. Thus, if ground functionalism is true, Consciousness Gap based arguments against physicalism fail.

Ground functionalism is a version of physicalism that is both *epistemically* and *metaphysically* non-reductive. Epistemically, as I already mentioned, Schaffer rejects all a priori entailments between the more and the less fundamental entities. Metaphysically, as I will soon present in more detail, Schaffer recognizes only *one sense* of 'existence.' This is the ontologically privileged, joint-carving sense of 'existence.' Thus, translated to my terminology, Schaffer's view is that *all entities,* whether they are fundamental or dependent, are Real entities.

> *Real entities*: an entity E is Real iff E exists in the ontological privileged sense of 'exist.'

In a nutshell, ground functionalism is a version of physicalism where both the fundamental and the dependent entities are Real, and there are explanatory gaps in all grounding cases.

In this chapter, I set out to defend the specialness of Consciousness Gap as a challenge for physicalism. I offer a two-fold response to Schaffer's ground functionalism. First, I argue that anti-physicalists like Chalmers and Goff (whom I call lightweight anti-physicalists) are *already safe* from Schaffer's challenge. This is because they reject Schaffer's assumption that both the fundamental and the dependent entities are Real. Second, I argue that anti-physicalists can accept Schaffer's epistemic views and still argue against physicalism based on Consciousness Gap. This is because Consciousness Gap *is special*, even if explanatory gaps are everywhere.

Before proceeding to my arguments against ground functionalism, it is important to first outline what exactly Schaffer is rejecting and what does ground functionalism posit. I dedicate the rest of this section to that task.

3.1.2. Macrophysical Scrutability

Even if Cosmic Scrutability fails,[47] many philosophers are willing to accept a weaker, restricted scrutability thesis.[48] The following such thesis is of particular importance to the philosophy of mind:

> *Macrophysical Scrutability*: the microphysical facts (as grounds) a priori entail all macrophysical facts (as groundees).

[47] Besides the counter-example of Consciousness Gap, Block and Stalnaker (1999) have famously argued that CS fails since there seems to be no a priori entailment between the microphysical truths and pre-theoretical macroscopic truths. For example, pre-theoretical truths about water (understood as the 'watery stuff' that is wet, transparent, odorless, and so on) do not seem to be a priori scrutable from the truths of H_2O.

[48] In Chalmers' (2012: 39) terminology, a scrutability thesis is restricted iff its base doesn't a priori entail all the other truths, but only a limited class of other truths.

The most thorough defense of Macrophysical Scrutability in the recent literature comes from Chalmers and Jackson (2001) and Chalmers (1996, 2012). Chalmers defines the *microphysical* truths as "truths about fundamental physical entities in the language of a completed fundamental physics." (2012: 110) On the other hand, the *macrophysical* truths are the "truths about any entities, including macroscopic entities, in the language of classical physics." (2012: 110) They involve "the structure and dynamics of the world at the macroscopic level, at least insofar as this structure and dynamics can be captured in terms of spatiotemporal structure (position, velocity, shape, etc.) and mass distribution." (Chalmers & Jackson, 2001: 330)

Macrophysical Scrutability is far more plausible than Cosmic Scrutability. Macrophysical Scrutability involves only physical truths at different scales. Thus, it involves "only a change of scale" (Chalmers & Jackson, 2001: 331) between physical truths. It entails that the physical structure and dynamics of the cosmos are intelligible at *all scales* in virtue of its fundamental building blocks. It is worth noting that we currently lack a complete theory of physics. Even given current physics, there are ongoing debates about the ontological implications of both quantum theory and the theory of relativity. In light of this, it is safe to say that based on current physics, no one certainly knows whether Macrophysical Scrutability is true or not. My goal is not to defend Macrophysical Scrutability but simply to state it as a plausible and widely held thesis.

Macrophysical Scrutability is widely accepted (explicitly or implicitly) by both anti-physicalists and physicalists. Chalmers and Jackson (2001), Chalmers (1996, 2012), and Goff (2017) defend Macrophysical Scrutability, while McQueen (2015), McLaughlin (2019), Balog (1999), and Levine (2001), amongst others, are at least open to the plausibility of Macrophysical Scrutability. For example, McQueen argues that the truth of mass additivity is a priori entailed by the principles of Newtonian microphysics, while arguably by the principles of special relativity, "with enough relativistic information one could deduce that something has non-additive mass." (2015: 1388) Moreover, McLaughlin (2019) argues that there is a priori entailment from the truths of quantum mechanics to the truths of chemistry. Finally, according to Balog (1999: 523), the 'A Priori Entailment Thesis' (which roughly translates to Cosmic Scrutability), "might be correct about all truths except phenomenal truths."

My view is that Macrophysical Scrutability is *indispensable* for the three 'epistemic arguments' against physicalism: the explanatory argument, the knowledge argument, and the conceivability argument.[49] Consciousness Gap (expressed either in terms of explanation, knowledge, or conceivability) features as an *epistemic premise* in all three of these arguments.[50] I argue that, although it is not explicitly stated, Macrophysical Scrutability likewise features in the background of these arguments as an *implicit assumption*.

Consciousness Gap alone—as an epistemic premise—is impotent against physicalism as a metaphysical thesis. This is why all three of the epistemic arguments posit a further premise that *links epistemology to metaphysics*. Link (conceivability version) and Transparent Link (conceivability version)—from the previous chapter—play the role of that premise in the context of the conceivability argument.

Link (conceivability version): Conceivability entails possibility.

All versions of Link express a connection between a priori entailment and possibility. Since all the cases of a priori entailment relevant for the epistemic arguments involve grounding, all versions of Link fundamentally express a connection between explanatory gaps and possibility. Thus, the most general version of Link that underlies all the epistemic arguments against physicalism is as follows:

Link: explanatory gaps entail possibilities.

Link entails that for any explanatory gap, *it is possible* that the ground could obtain while the groundee fails to obtain. If so, the putative grounding claim contained in that explanatory gap must be false. If so, via Link, explanatory gaps entail the falsity of the grounding claims that they contain.

[49] See section 1.3.3.
[50] See Chalmers (2003: 107–8).

I argue that Macrophysical Scrutability *justifies* Link. Macrophysical Scrutability applies ubiquitously throughout the cosmos. It provides a ubiquitous class of true grounding cases without explanatory gaps. Macrophysical Scrutability demonstrates that explanatory gaps are not everywhere. Instead, it appears that, as a general principle, the following thesis is true:

No Gaps: there are no explanatory gaps in true cases of grounding.

No Gaps *connects* Macrophysical Scrutability and Link. No Gaps entails that explanatory gaps are a good guide to locating false grounding claims. It entails that, for any explanatory gap involving a ground P and a groundee Q, it might be *possible* that P does not ground Q. If so, Link must be true.

Putting all of this together: Macrophysical Scrutability justifies No Gaps, while No Gaps justifies Link. Thus, via No Gaps, Macrophysical Scrutability justifies Link. Although roughly stated, this chain of inference demonstrates a connection between Macrophysical Scrutability and Link. In what follows, I put my hypothesis to test by applying it to Chalmers' (2010) and Goff's (2017) versions of Link. Before proceeding, keep in mind that Macrophysical Scrutability involves only physical concepts while Consciousness Gap involves physical concepts (for the macrophysical facts) and phenomenal concepts (for the phenomenal facts).

Chalmers (2010) defends a *strong* form of Link. He argues that conceivability (and thus explanatory gaps, as I have defined them) *always* entails possibility. In his view, there are no genuine counter-examples to Link; all suggested counter-examples are only apparent and due to an *ambiguity* in the involved concepts. Chalmers argues that such concepts contain two intensions of meaning.[51] Once this is cleared up, it becomes plain that both intensions correspond to distinct possibilities. Chalmers argues that physical and phenomenal concepts are not ambiguous.[52] Thus, Consciousness Gap unambiguously entails a metaphysical possibility.

[51] In the technical terminology of Chalmers' two-dimensional semantic framework, the putative counter-examples to Link involve concepts with different primary and secondary intensions.
[52] In Chalmers' terminology, these concepts have coinciding primary and secondary intensions.

Goff (2017: 96–105), as I already mentioned, defends a *weaker* form of Link. He argues that conceivability (and thus explanatory gaps, as I have defined them) entails possibility iff the involved concepts are transparent. In his view, all counter-examples to Link involve concepts that are not transparent. Goff argues that physical and phenomenal concepts are transparent. Thus, Consciousness Gap entails a metaphysical possibility.

Assume that Macrophysical Scrutability was false. If so, in Chalmers' case, there would be no ubiquitous class of grounding cases without explanatory gaps that involve non-ambiguous concepts. Similarly, in Goff's case, there would be no ubiquitous class of grounding cases without explanatory gaps that involve transparent concepts. Thus, in either case, if Macrophysical Scrutability were false, there would be little motivation to posit Link. If so, Macrophysical Scrutability, as an implicit assumption, is indispensable for Chalmers' and Goff's versions of Link.

3.1.3. The Existence Explanatory Gap

Schaffer (2017, forthcoming) accepts Consciousness Gap but rejects both Cosmic Scrutability and Macrophysical Scrutability. If ground functionalism is true, explanatory gaps are *always the rule, and there are no exceptions*.

Schaffer has defended the thesis that explanatory gaps are everywhere in a series of papers (2016, 2017, forthcoming). His most detailed defense is in "The Ground Between the Gaps" (2017), and I mainly focus on it here. There, Schaffer argues that explanatory gaps occur even in paradigmatic cases of grounding, such as between the H, H and, O atoms and H_2O molecules.

Schaffer (2017) is not perfectly clear what he means by the terms 'H, H, and O' and 'H_2O.' Before proceeding, it is important to note that 'H, H, and O' must stand for two hydrogen atoms and one oxygen atom *after* they have formed molecular bonds, and not for individual 'free-floating' atoms before they have formed molecular bonds. This is entailed by grounding being a *synchronic* relation. If H, H, and O is the ground of H_2O, both the H, H, and O atoms and the H_2O molecule must exist at the *same time*

and place, and the atoms must already be bonded in the right way. Thus, for clarity, instead of 'H, H, and O,' I use the label 'H+H+O.' I take H+H+O to be a placeholder term for any plurality of fundamental microphysical entities, grounding H_2O. Atoms are not fundamental entities; nevertheless, I speak of them as such for convenience.

Schaffer's argument for explanatory gaps everywhere is primarily based on it being "conceivable, logically possible, and not a priori knowable otherwise that there are no *derivative entities*." [emphasis mine] (2017: 13–4) By 'derivative entities,' Schaffer refers to what I have been calling 'depended entities.' Schaffer illustrates the preceding with the conceivability of what he, following Bennett (2011), calls a *'flat-world.'*

To conceive of a flat-world, we need to conceive a scenario in which the fundamental physical facts obtain, yet no concrete groundees obtain. The flat-world is a world in which only fundamental entities exist, and there are no concrete dependent entities. Thus, in the flat-world, there would be H+H+O atoms, but there would be no H_2O molecules grounded in them.

The flat-world is a perfect microphysical copy of our world. Thus, the flat-world is observationally indistinguishable from our world (2017: 9). The flat-world is not a world where oceans are empty, nor is it a world where there is some sort of quantum mush everywhere. Rather, the key difference between the flat-world and worlds with dependent entities is the number of entities. There are more entities in worlds with dependent entities than in the flat-world. To use a metaphor, if God is running an inventory of what exists, the inventory would have fewer items in the flat-world than in worlds with dependent entities.

Schaffer's flagship example involves *mereological composition*. Following van Inwagen (1990), there is an ongoing debate in the metaphysics literature regarding when, and if at all, parts compose a further entity. Roughly, that is van Inwagen's '*Special Composition Question*.' Mereological nihilism, i.e., the thesis that parts never compose anything, is true in the flat-world. The flat-world's conceivability demonstrates that the Special Composition Question cannot be answered a priori in virtue of the microphysical facts alone.

In Schaffer's own words:

> I am saying that an ideal mind, given the empirical information that there are H, H, and O atoms in a given arrangement, and given the conceptual information that an H_2O molecule is an individual composed in a given way and with a given nature, still needs more information to determine whether an H_2O molecule is present. She needs to know if those H, H, and O atoms compose anything, and the nature of what they compose. She needs substantive metaphysical information about the principles of composition. (2017: 10)

Schaffer argues that the non-trivial nature of the Special Composition Question results in a ubiquitous explanatory gap. I formulate this explanatory gap as follows: [53]

> *Existence Gap*: the microphysical facts (as grounds) do not a priori entail the existence facts of macrophysical entities (as groundees).

Schaffer concludes that: "explanatory gaps are everywhere in nature, lurking in every concrete transition from more to less fundamental." (2017: 14) His solution for closing these explanatory gaps everywhere is that "grounding bridges gaps." (2017: 2) The idea behind this slogan is that there is a need for "substantive grounding principles" (2017: 14) connecting grounds and groundees. The grounding principles are posited abductively, "by inference to the best explanation, in a holistic and fallible manner" (forthcoming). They explain how a specific ground determines a specific groundee.

Existence Gap seems to contradict Macrophysical Scrutability. Schaffer uses Existence Gap to argue that all versions of Link are false, that explanatory gaps alone never entail metaphysical possibilities. Schaffer's ground functionalism posits that possibility is always determined by the grounding principles in conjunction with the grounds. If ground functionalism is true, conceivability never entails possibility without the

[53] Schaffer (2017: 13–4) also posits the conceivability of a ghost-world where all dependent entities are epiphenomenal. This is expected to show that even if composites exist, their natures are not a priori scrutable from the microphysical truths. Everything I say here about existence— mutatis mutandis—applies to the case of nature.

inclusion of the grounding principles. Thus, if ground functionalism is true, all scrutability theses whose bases consist solely of truths about the fundamental entities are false. If so, both Cosmic Scrutability and Macrophysical Scrutability are false. The only scrutability thesis that Schaffer accepts is one that includes the grounding principles in the scrutability base alongside the microphysical truths.[54]

[54] Schaffer's (2017: 18–9) view is that the most likely candidate for a scrutability is a 'PTIG' scrutability base, namely the conjunction of the microphysical truths (P), a totality 'that's all' premise (T), the indexical truths (I), and the grounding principles (G).

3.2. Lightweight Anti-physicalism

3.2.1. War of the Worlds

We have reached an impasse between opposing worldviews at this point in the discussion. On the one hand, in the philosophy of mind debates, Macrophysical Scrutability appears to be broadly accepted by both anti-physicalists and physicalists. The standard view in the metaphysics of consciousness debates is that explanatory gaps are an odd exception to an otherwise intelligible universe and need to be addressed. According to friends of Macrophysical Scrutability, the H+H+O truths a priori entail the H_2O truths. On the other hand, metaphysicians who have dealt with the Special Composition Question and who think mereological nihilism is conceivable might agree with Schaffer that Existence Gap is true. If so, in their view, the H+H+O truths do not a priori entail the H_2O truths.

Existence Gap and Macrophysical Scrutability seem to contradict each other. The clash over Macrophysical Scrutability and Existence Gap is of crucial importance to the philosophy of mind. Existence Gap and Macrophysical Scrutability both involve truths about the macrophysical entities. Yet, Macrophysical Scrutability justifies Link, while Existence Gap undermines Link.

In the rest of this section, I offer a diagnosis of the dispute. Before proceeding, it is important to have the following two epistemic theses in mind:[55]

> *Lightweightism*: For some true grounding relations between concrete entities, the grounds a priori entail their groundees.
>
> *Heavyweightism*: For all true grounding relations between concrete entities, the grounds do not a priori entail their groundees.

[55] These two epistemic these are based on Chalmers' (2012, pp. 267–71) classification of ontological views.

Ground functionalism is a version of Heavyweightism. In contrast, anti-physicalists like Chalmers and Goff (henceforth, I restrict the label 'friends of Macrophysical Scrutability' to their views) embrace a form of Lightweightism.

3.2.2. Real and Nominal Existence

Why do friends of Macrophysical Scrutability say they cannot conceive of H+H+O without H_2O, while Schaffer says he can? I argue that this epistemic disagreement has a likely *metaphysical explanation*. My proposed diagnosis is that the disagreement arises because the two sides understand the existence of macrophysical entities in two different ways, using *two different senses of 'existence.'*

Schaffer is explicit that, in his view, both the fundamental and the dependent entities exist equally, in "*the one and only sense of existence*." (2009a: 360) He takes "entities like tables to be full-blown 'heavyweight' entities on the roster of entities." (2009a: 360) Thus, in Schaffer's view, if H_2O exists, there is one more entity on the one and only list of what exists, in addition to the H+H+O atoms.

Schaffer's one notion of existence is naturally read as the primitive sense of 'existence' deployed in serious discussion of ontology (in the 'Ontology room'). This is the sense of 'existence' we use when we ask whether God, Platonic universals, or quantum particles exist. In Sider's (2011) terminology, expressions that employ this sense of 'existence' aspire to 'carve nature at the joints.' In this dissertation, I called the entities that exist in this sense of 'exist' Real entities. Depending on whether they are in the roles of grounds or groundees, the Real entities can be further described as *Real grounds* or *Real groundees*.

Chalmers (2009: 120) is open to the idea that the fundamental entities might be Real; however, he opposes the idea that groundees might likewise be Real. In his view, insofar as the dependent entities exist, they exist only in a nominal way. In his earlier work, Chalmers (1996) argued that the dependent facts are *redescriptions* of the fundamental facts. Using the biological facts as an example, he says:

> Once God (hypothetically) made sure that all the physical facts in our world held, the biological facts came along for free. The B-facts merely redescribe what is described by the A-facts. They may be *different* facts (a fact about elephants is not a microphysical fact), but they are not *further* facts. (1996: 41)

Goff (2017: 62–3), similarly to Chalmers, accepts the Reality of the fundamental entities but opposes the Reality of the macrophysical entities. According to Goff, the macrophysical entities are *grounded by analysis* in the microphysical entities. In cases of grounding by analysis, "the grounding fact provides all that is essentially required for the entities contained in the grounded fact to be part of reality." (2017: 45) Using 'party' as an illustrative example, Goff says:

> It's not as though there are the people dancing, drinking, and so on, and then there's this extra thing—the party—that floats above their heads. There's a very intuitive sense in which the fact that there is a party is *nothing more* than the fact that there are people revelling; a world in which there are people revelling is *already thereby* a world in which there is a party. (2017: 42–3)

The sense of 'existence' that Chalmers and Goff use to describe the macrophysical entities corresponds to how 'existence' is often used in ordinary talk but not in the Ontology room. This is the sense of 'exist' that we typically use when we say that Harry Potter, universities, or parties exist. This sense of 'existence' does not carve nature at the joints. In this dissertation, I called the entities that exist in this sense of 'exist' Nominal entities. Depending on whether they are in the roles of grounds or groundees, the Nominal entities can be further described as *Nominal grounds* or *Nominal groundees*.

> *Nominal entities*: an entity E is Nominal iff E exists in the nominal sense of 'exist.'

Schaffer thinks that all groundees are Real. Chalmers and Goff seem to agree that all fundamental entities are Real; however, they think that all macrophysical groundees are Nominal. Regardless of who is right, this shows that we can conceive of macrophysical entities as either Real or Nominal groundees. If so, when we conceive of H+H+O as Real-H+H+O (as a Real ground), we can conceive of H_2O as either Real-H_2O (as a Real groundee) or as Nominal-H_2O (as a Nominal groundee) (see *Table 5*).

	Ground	**Groundee**	**Explanatory Gap?**
Schaffer	Real-H+H+O	Real-H_2O	*Yes*
Chalmers	Real-H+H+O	Nominal-H_2O	*No*
Goff	Real-H+H+O	Nominal-H_2O	*No*

Table 5: The explanatory gap between H+H+O and Real-H_2O.

The results in Table 5 make perfect sense given some plausible assumptions about the Real and the Nominal entities.

Real entities are metaphysically privileged. Real groundees are *ontologically equal* to their Real grounds. If so, it is plausible to expect that Real groundees contain information about their existence that is unique to them, over and above the information contained in their grounds. Their existence is *not fully contained* within the existence of their grounds. Thus, it is clear why—even given full information about ground and groundee—the existence of Real groundees might not be a priori scrutable from the entities in their ground.

Nominal entities are not metaphysically privileged. Nominal groundees are *ontologically inferior* to their Real grounds. If so, it is plausible to expect that Nominal groundees contain no information about their existence over and above the information contained in their grounds. Their existence is *fully contained* within the existence of their grounds. Thus, it is clear why—given full information about ground and groundee—the existence of Nominal groundees might be a priori scrutable from the entities in their ground.

Existence Gap involves Real groundees, while Macrophysical Scrutability involves Nominal groundees. This conclusion clearly follows from the above exegesis and assumptions. If so, Existence Gap and Macrophysical Scrutability do not contradict each other; instead, they involve *different kinds of entities* in the roles of groundees. Existence Gap shows that the existence of Real-H_2O (as a groundee) does not analytically follow from the existence of Real-H+H+O (as a ground). In contrast, Macrophysical Scrutability shows that the existence of Nominal-H_2O (as a groundee) does analytically follows from the existence of Real-H+H+O (as a ground).

The difference between Real groundees and Nominal groundees explains the clash between friends of Macrophysical Scrutability and Schaffer. It explains why explanatory gaps are everywhere on the ground functionalism picture, while explanatory gaps are not everywhere on the Macrophysical Scrutability picture. Friends of Macrophysical Scrutability reject Real groundees. Thus, Existence Gap does not apply in their *reductive grounding* frameworks. In contrast, Schaffer rejects the Nominal groundees. Thus, Macrophysical Scrutability does not apply in his *non-reductive grounding* framework. If so, the belief in Real groundees is an underlying assumption of ground functionalism that friends of Macrophysical Scrutability reject.

I do not take a hard stance on this dispute. I only claim that, in the context of the above dispute, the belief in Nominal groundees is what makes Macrophysical Scrutability plausible. In contrast, the belief in Real groundees makes Existence Gap plausible. Thus, if all macrophysical groundees are Real, as Chalmers and Goff seem to think, neither Existence Gap nor ground functionalism would not get off the ground.

Putting all of this together, I define the lightweight anti-physicalist position as follows.

> *Lightweight Anti-physicalism*: (1) All macrophysical entities are Nominal groundees; (2) For all true grounding relations between grounds and Nominal groundees, given full and unambiguous information about ground and groundee, the ground a priori entails the

groundee; (3) The fundamental grounds are either entirely non-physical or not entirely physical.

3.3.3. Schaffer's Response

Schaffer (2017: 21–4) anticipates his opponents might think that embracing Lightweightism would help them secure the specialness of Consciousness Gap. In response, Schaffer argues that this strategy cannot succeed. Roughly, he argues that if Lightweightism were true, it should apply equally to *all* instances of grounding. Thus, if friends of Lightweightism think Macrophysical Scrutability is true, they should also think that there is no gap between the microphysical and the phenomenal facts. In other words, Schaffer seems to be saying that a commitment to Lightweightism entails both that Consciousness Gap is false and that Cosmic Scrutability is true. In his own words:

> For if it can be "just by meanings" that the H, H, and O atoms compose something miscible, it can equally be "just by meanings" that these neurons and synapses compose someone miserable. Or at least, no relevant difference between the chemical and the phenomenal has been identified that keeps the latter specially out of reach of this stretched out notion of the analytic. (2017: 23)

He concludes, "It is hard to be a dualist if analytic connections are so easy!" (2017: 23)

According to Schaffer, friends of Lightweightism cannot explain why there is an explanatory gap between the physical and the phenomenal (and nowhere else) without falling into *circular thinking*. As he puts it: "Why is this connection specially opaque? Because there is no analytic connection. Why is there no analytic connection? Because the connection is specially opaque." (2017: 24)

I think that Schaffer, in his argument above, might be begging the question against friends of Lightweight Anti-physicalism. Lightweight Anti-physicalism does not entail that *anything goes*, that there are no explanatory gaps whatsoever. Schaffer fails to take notice of the *qualifications* that define Lightweight Anti-physicalism.

Firstly, all friends of Lightweight Anti-physicalism would agree that for any grounding connection to be a priori scrutable, the posited grounding connection *must be true*. No friend of Lightweight Anti-physicalism would expect the truths of H+H+O atoms to a priori entail the truths of wombats. Why would they when the posited grounding connection is evidently false? Secondly, Lightweight Anti-physicalism entails that, even in the cases where the posited grounding connection is true, the ground a priori entails the grounded iff the groundee is *a Nominal groundee*. Thus, the expectation for there being no explanatory gaps in true grounding cases applies only to cases of reductive grounding. Schaffer, in his arguments above, takes none of these qualifications into consideration.

In his argument above, Schaffer seems to presuppose (a) that consciousness is grounded in the physical entities and (b) that consciousness is a Nominal groundee. In other words, Schaffer appears to presuppose that some reductive and Type-A version of physicalism is true. If this were the case, then yes, Schafer's argument would be sound. However, this is precisely what friends of Lightweight Anti-physicalism want to deny. Friends of Lightweight Anti-physicalism agree that there are *correlations* between physical states and human consciousness. However, they do not start with the *further assumptions* that consciousness has a physical ground and that consciousness is a Nominal groundee. Instead, they use Consciousness Gap and Macrophysical Scrutability to point out that physicalists are wrong to believe either of these two claims. There is *no circularity* in thinking that Consciousness Gap obtains because consciousness is not grounded in the physical or because consciousness is not a Nominal groundee.

In summary, friends of Lightweight Anti-physicalism already reject one of the base assumptions behind Schaffer's heavyweightism: namely, Schaffer's commitment to Real groundees. If Chalmers' and Goff's conception of the macrophysical domain is feasible, their arguments seem to be already safe from Schaffer's Existence Gap challenge. To decisively disprove Lightweight Anti-physicalism by rejecting Macrophysical Scrutability, the onus is on Schaffer to give *a further argument* against Nominal groundees, as employed by Chalmers and Goff.

3.3. Heavyweight Anti-physicalism

3.3.1. Beyond Lightweightism

So far, I argued that anti-physicalists could reasonably reject Existence Gap by embracing Lightweight Anti-physicalism. However, despite this, I still have not fully diffused Schaffer's challenge. So construed, being an anti-physicalist seems to hinge on the success of Lightweight Anti-physicalism. Therefore, if Schaffer is right, explanatory gap arguments cannot work on a heavyweight view, such as ground functionalism.

The above is problematic since it entails that one can be an anti-physicalist iff Lightweight Anti-physicalism is true. Although versions of Lightweightism seem to be often assumed outside of the metaphysics literature, many contemporary metaphysicians strongly favor Heavyweightism. Ground functionalism seems like the best attempt yet at creating a heavyweight version of physicalism. Thus, suppose that Schaffer is right, and we do need to conceive of H_2O and dependent entities in general as Real groundees. As Schaffer (2017: 22–3) anticipates, the physicalist can simply embrace ground functionalism or another version of Heavyweightism.[56]

I think there is no reason to restrict anti-physicalism to Lightweight Anti-physicalism. It is plausible that some anti-physicalists could be sympathetic towards heavyweightism and think Existence Gap is true, although they do not agree with Schaffer's physicalism. Such anti-physicalists would be happy to bite the bullet of Existence Gap and embrace heavyweight anti-physicalism.

> *Heavyweight Anti-physicalism*: (1) All concrete groundees are Real groundees; (2) For all true grounding relations holding between grounds and Real groundees, the grounds never a priori entails their Real

[56] It's an additional problem, that's beyond the scope of this paper, whether physicalists would be motivated to embrace ground functionalism. Ground functionalism seems to be a close cousin (if not a version) of emergentism or naturalistic dualism. Thus, it might appear too semi-physicalist (or too semi-dualistic) for most physicalists.

groundees; (3) The fundamental grounds are either entirely non-physical or not entirely physical.

Friends of Heavyweight Anti-physicalism can argue against physicalism either (a) without recourse to Consciousness Gap, or (b) with recourse to Consciousness Gap, by saying that Consciousness Gap is special on heavyweight views.

Arguing against physicalism without recourse to Consciousness Gap means giving up one of the primary motivations for being an anti-physicalist. At least since Descartes, anti-physicalism has been fueled by the puzzlement of how consciousness can be physical or can exist in virtue of something physical. The idea that consciousness is physical has often evoked a feeling of absurdity even among physicalists.[57] Moreover, giving up on Consciousness Gap being true would mean giving up many argumentative advances made by anti-physicalists in the last decades, especially when it comes to the discussions of phenomenal concepts. Thus, although it might be possible to argue for Heavyweight Anti-physicalism without recourse to Consciousness Gap being true, that is certainly undermotivated and disadvantageous.

Alternatively, friends of Heavyweight Anti-physicalism can say that Consciousness Gap is special, even on heavyweight views. After all, the mysteriousness of consciousness does not seem to diminish if Existence Gap is true. Even if all groundees are Real groundees, the presence of consciousness in the physical world seems to remain as puzzling as ever. The epistemic arguments against physicalism are the best articulation of this puzzlement. In what follows, I show how to argue for Heavyweight Anti-physicalism based on Consciousness Gap.

3.3.2. Deep Opacity

Schaffer argues that if ground functionalism is true, Consciousness Gap is not unique in any way. In his own words: “I agree that the phenomenal information is extra information. My point is that the same holds for all other higher-level information.” (2017: 18)

[57] See Sundström (2018).

Against Schaffer, I argue that Consciousness Gap is significantly more mysterious than Existence Gap. In other words, the grounding of consciousness is more mysterious than the grounding of macrophysical Real groundees (henceforth: composites). I call this the 'Deep Opacity Argument:'

P-1. *Deeply Opaque:* Consciousness Gap is deeply opaque.
P-2. *Explanation*: Deeply opaque explanatory gaps require an explanation of their deep opacity.
C. Consciousness Gap requires an explanation of its deep opacity.

The Deep Opacity Argument is valid and simple. Nevertheless, both P-1 and P-2 need further clarification and justification.

First, *P-1: Deeply Opaque*. Schaffer argues that all grounding relations are opaque since it is conceivable that the grounds obtain, while the groundees do not. Against Schaffer, I argue that even if we accept Heavyweightism, not all grounding relations are equally opaque. Even if explanatory gaps are everywhere, some explanatory gaps are more mysterious than others and thus *deeply opaque*.

> *Deep Opacity*: An explanatory gap is deeply opaque iff (1) The groundee concept is transparent, and (2) The groundee concept does not a priori reveal that its referent has the ground posited by the metaphysical theory under consideration.

A few terminological remarks. First, Deep Opacity requires transparent concepts to ward off counter-examples involving concepts that are ambiguous or fail to reveal the essences of their referents. Like Goff, I assume that phenomenal, physical, and logical concepts are transparent. Second, I use 'reveal' in a purely epistemic way to indicate what appears to be the nature of a concept's referent upon a priori reflection. Thus, revelation by itself has no metaphysical implications. Third, 'the metaphysical theory under consideration,' unless otherwise specified, is ground functionalism.

I claim that Consciousness Gap *is deeply opaque* because phenomenal concepts *do not a priori* reveal that experiences have physical grounds. Both PCS (phenomenal concept strategy) physicalists and anti-physicalists univocally agree on this; they agree that phenomenal concepts a priori reveal *only* the phenomenal characters of experiences. Take the experience of pain as an example. A priori reflection on the phenomenal concept <pain> only reveals the feeling of pain, that pain *hurts*. It does not reveal any putative physical grounds that pain might have. Plausibly, this is the reason why, at least since Descartes, it has been a forte of dualists and, in general, of anti-physicalists to argue that disembodied minds are conceivable. Thus, in the case of consciousness, it is neither a priori evident that experiences have physical grounds nor that experiences are grounded.

In contrast, Existence Gap *is not deeply opaque* because concepts of composites *a priori reveal* that composites have physical grounds. By definition, molecules are physical entities grounded in atoms bonded in the right way. Real-H_2O, even if epiphenomenal, is a molecule grounded in H+H+O. Thus, the <Real-H_2O> concept a priori reveals that its referent Real-H_2O is grounded in H+H+O. Because of this, unlike in the consciousness case, a 'disembodied' molecule (or any macrophysical composite in general), without parts and a physical ground, appears inconceivable. Remember, H+H+O and Real-H_2O are observably indistinguishable. Thus, although H+H+O might be conceivable without Real-H_2O, Real-H_2O seems inconceivable without H+H+O. Thus, in the case of composites, it is a priori both that composites are grounded and that they have physical grounds.

To sum up, purely on a priori considerations of their groundees, Consciousness Gap is deeply opaque, while insofar as Existence Gap can be called opaque, it must be *regularly opaque*.

> *Regular Opacity*: An explanatory gap is regularly opaque iff (1) the groundee concept is transparent, and (2) the groundee concept a priori reveals that its referent has the ground posited by the metaphysical theory under consideration.

In deep opacity cases, we presume a ground, although this is not evident from a priori reflection on the groundee concept. In contrast, in cases of regular opacity, the ground is not simply presumed. Instead, it is plain upon a priori reflection on the groundee concept (assuming we have a transparent groundee concept).

By itself, deep opacity does not translate into a difference of metaphysical significance. Nevertheless, deep opacity appears *more mysterious* than regular opacity. In the case of Existence Gap, both composites and their grounds have physical natures, and there is an a priori connection from groundee to ground. Thus, there appears to be explanatory work done by *both* the ground and the grounding principles. This is why the grounding of composites in physical grounds is unsurprising and does not seem like a cosmic accident. In contrast, in cases of deep opacity, the ground and the groundee are a priori disconnected. The ground does no obvious explanatory work. Thus, if ground functionalism is true in the consciousness case, all the explanatory work appears to be done *only* by the grounding principles. This is why the grounding of consciousness in a physical ground is surprising and seems like a cosmic accident.

This brings us to *P-2: Explanation*. I argue that if the ground functionalist cannot remove the mysteriousness of deep opacity, she must *at least* explain the presence of deep opacity.

The first reason in support of P-2 is that mysteriousness, by itself, asks for an explanation. Ground functionalism promises to alleviate mysteriousness. In Schaffer's (forthcoming) own words: "a satisfying metaphysics should be *explanatory*." Thus, since deep opacity is mysterious and Consciousness Gap is deeply opaque, the ground functionalist should explain its deep opacity.

The second reason in support of P-2 is that, on the heavyweight view, regular opacity seems to be *the rule*, while deep opacity seems to be *the exception to the rule*. Existence Gap is regularly opaque and is everywhere in all potential instances of composition. In contrast, Consciousness Gap is deeply opaque and only obtains in cases where there is consciousness. Thus, Consciousness Gap is an exception, and exceptions, in general, deserve an explanation.

Schaffer might object and say that even if Consciousness Gap is deeply opaque, it is *not the only* deeply opaque explanatory gap. Schaffer (forthcoming) mentions potential explanatory gaps involving sets, value, and particle locations. The *set gap* is a putative explanatory gap between the members of sets (as grounds) and sets (as groundees). The *value gap* is a putative explanatory gap between the grounds of value and the value facts (as groundees). Finally, the *particle gap* assumes Albert's (1996) reading of Bohmian quantum theory. It is a putative explanatory gap between the one fundamental particle (the 'world particle,' in Albert's terminology) in 3N-dimensional configuration space (as a ground) and the locations of the many dependent particles in ordinary 3-dimensional space (as groundees). I go over these cases one by one.

The set gap does not appear to be deeply opaque. The set singleton {Socrates} (as a groundee)—when conceived under a transparent, logical concept—clearly a priori reveals that it has Socrates (as its ground) as a member.

The value gap might be deeply opaque. The moral property 'goodness' (as a groundee) does not seem to a priori reveal that it has a physical ground. This might simply be because moral concepts are not transparent. Alternatively, it might be because the posited grounding relation is false. The value gap posits that moral properties are Real. Traditionally, this has been the position of moral realists who oppose naturalism. If so, moral realists could argue against ground functionalism based on the deep opacity of the value gap in a way analogous to how I use Consciousness Gap to argue against it. If so, everything I say here about consciousness—*mutatis mutandis*–likewise applies to the case of value. Thus, it seems that recourse to the value gap at best provides an unstable shelter for the ground functionalist.

The particle gap does not appear to be deeply opaque. When conceived under a transparent physical concept, each dependent particle (as a groundee) does seem to a priori reveal *something* about the world particle (as its ground). After all, the one fundamental particle and the many dependent particles *are both essentially particles*.

Thus, they share the *same essential properties* of what it means to be a particle. Thus, the particle gap cannot be deeply opaque.[58]

This ends my defense of the Deep Opacity Argument. Assuming that the preceding reasoning is sound, anti-physicalists can use the Deep Opacity Argument to argue for Heavyweight Anti-physicalism. Anti-physicalists could argue that *panpsychism* can do better than ground functionalism. Panpsychism is roughly the thesis that all fundamental physical entities are *intrinsically conscious*. Goff (2017) defends a version of panpsychism, while Chalmers (1996, 2003) is greatly sympathetic towards it.[59] The major selling point of panpsychism is its compatibility with the causal closure of the physical and all the other scientific evidence that physicalism is compatible with. Thus, if panpsychists are right about this, both ground functionalism and panpsychism can work as attempts at a best explanation of consciousness. I call the conjunction of panpsychism and Heavyweight Anti-physicalism—*'heavyweight panpsychism.'*

Heavyweight panpsychism promises to *remove* the deep opacity of Consciousness Gap. Given heavyweight panpsychism, it seems that human experiences (as groundees)—when conceived under transparent phenomenal concepts—do a priori reveal *something* about the experiences of the fundamental entities (as their grounds). Both human experiences and the experiences of the fundamental entities are *essentially experiences*. Thus, although we might not know what exactly the fundamental experiences feel like, we grasp what is essential to them as a kind. This makes Consciousness Gap *regularly opaque,* and thus the grounding of consciousness no longer looks like a cosmic accident.[60]

Schaffer ignores the panpsychism option. I do not claim that heavyweight panpsychism is true. My claim is that all other things being equal—heavyweight panpsychism paints a more cohesive picture of reality than ground functionalism.

[58] There's a lot more to be said about potential explanatory gaps between the quantum and the macrophysical scales, but that is beyond the scope of this paper. For a discussion see Ney and Albert (2013).

[59] I will explore panpsychism I much greater detail in Chapter 4.

[60] Another advantage of heavyweight panpsychism is that it seems to avoid the combination problem, which is roughly the problem of how the fundamental consciousness gives rise to dependent consciousness. If heavyweight panpsychism is true, phenomenal combination would be explained via the grounding principles.

Unlike ground functionalism, heavyweight panpsychism removes the deep opacity of Consciousness Gap. Thus, if heavyweight panpsychism fits the empirical data equally well as ground functionalism, yet explains more, Schaffer's adoption of ground functionalism needs further justification.

The ground functionalist could explain the deep opacity of Consciousness Gap by appeal to *the nature of phenomenal concepts*. Doing so entails walking in the footsteps of the PCS physicalist. As I already discussed, Blind Pointer physicalists typically think that Consciousness Gap obtains because phenomenal concepts are *radically opaque*. In their view, consciousness is physical, yet it appears mysterious due to the special features of phenomenal concepts. This strategy might offer a way out for the ground functionalist (assuming they can also find a way to respond to my argument against radically opaque concepts from Chapter 2). However, it also entails a return to the old (pre-Existence Gap) debate about phenomenal concepts. In effect, to resolve the problem, the ground functionalist must acknowledge that Consciousness Gap is special.

I leave it up to the reader to choose their preferred solution in the phenomenal concepts debate. I showed that even if explanatory gaps are everywhere, Consciousness Gap stands out as deeply opaque. The mysteriousness of Consciousness Gap demands an explanation. Whether the reader agrees with the anti-physicalist or the physicalist, Consciousness Gap regains its dialectical significance. There is a lot more to be said about the details. Nevertheless, the message is clear: it is viable to be a heavyweight anti-physicalist, based on the specialness of Consciousness Gap.

In conclusion, I defined two metaphysical positions that anti-physicalists can take in response to Schaffer's Existence Gap challenge. One, they can *reject* Existence Gap and *embrace* Lightweight Anti-physicalism. Two, they can *accept* Existence Gap and *embrace* Heavyweight Anti-physicalism. I argued that in either case, they could be anti-physicalists in virtue of Consciousness Gap. The number of explanatory gaps does not decrease the mysteriousness of consciousness.

Chapter IV:
Against Panpsychism

"It seems that to find a place for consciousness within the natural order, we must either revise our conception of consciousness or revise our conception of nature." (Chalmers, 2003: 102)

"Realistic materialists—realistic anybodies—must grant that experiential phenomena are real, concrete phenomena, for nothing in this life is more certain." (Strawson, 2015: 164)

"I have a clear and distinct idea of myself, in so far as I am simply a thinking, non-extended thing; and on the other hand I have a distinct idea of body, in so far as this is simply an extended, non-thinking thing." (Descartes, 1645/1996: 54)

4.1. Panpsychism

4.1.1. What is Panpsychism?

Panpsychism is the metaphysical thesis that fundamental physical reality is *experiential*. Panpsychists aspire to *redefine* the physicalist conception of fundamental reality. They argue that consciousness is not only a property of humans, animals, and other complex beings, but moreover, it is an essential property of the cosmos at its most fundamental level. If so, there is something that it is like to be a fundamental physical entity, such as a quark, an electron, or spacetime. Moreover, panpsychists argue, human experiences are grounded in (or in some way produced by) the experiences at the fundamental level. If panpsychists are right about this, the promised result is a metaphysical picture where human experiences and the physical world cohere better than in the physicalist picture.

Panpsychism is rising in popularity.[61] In contemporary philosophy, versions of panpsychism have been defended by David Chalmers (1996, 2003, 2015), Galen Strawson (2006b, 2015, 2020), Gregg Rosenberg (2004), Itay Shani (2015), William Seager (2016), Philip Goff (2017), Hedda Hassel Mørch (2018a, 2020), Bernardo Kastrup (2018), Luke Roelofs (2019), and Miri Albahari (2020), among others. Contemporary panpsychists largely build their views upon the works of Bertrand Russell (1927) and the physicist Arthur Eddington (1928). Moreover, it is worth noting that panpsychism is not a new theory but has a long and venerable history in both Western and Eastern philosophy.[62]

My focus in this chapter is on one specific version of panpsychism: *pure Russellian constitutive panpsychism*. As I will soon argue, I consider this version of panpsychism to be the most metaphysically simple and elegant and the most explanatory powerful when it comes to solving the problem of consciousness.

[61] Panpsychism was almost forgotten and largely seen as ridiculous by analytic philosophers throughout most of the 20th century. 'Reductio ad panpsychism' arguments, namely argument against philosophical theses based on their entailment of panpsychism, were not uncommon in the literature. For an example, see Madden and Hare (1971).

[62] See Skrbina (2005) for a in depth overview of the history of panpsychism in the West.

Pure panpsychism is the metaphysical thesis that fundamental reality is *entirely experiential*. Panpsychism, so defined, is a thesis of *mental monism*.[63] It is the view that "mind is all there is to reality: mind is the stuff of reality, the ('categorical') *stuff being* of reality." (Strawson, 2020: 317–8) According to pure panpsychists, both the physical facts and the facts of human experiences are fully grounded in the facts of the fundamental experiences.

Russellian panpsychists[64] argue that there is more to fundamental physical reality than a pure structure. These panpsychists agree with physicalists that physics accurately describes the *observable structure* of the fundamental entities. However, they moreover argue that it is implausible for a physical structure to exist autonomously, without having some *deeper qualitative ground*. In their view, experiences are the perfect candidates for this deeper qualitative ground. Thus, Russellian panpsychists conclude that the observable physical structure of the cosmos *is* the structure of the fundamental experiences. In their view, physical entities are the observable manifestation of the fundamental experiences. Humans observe their structure but not their essential phenomenal characters (see *Fig. 1*).

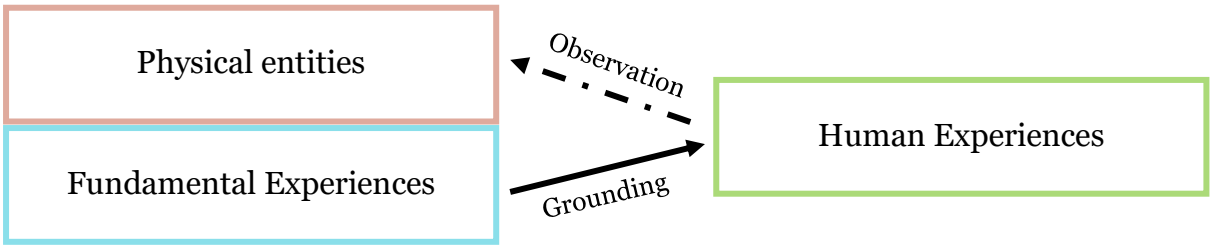

Figure 1: The Russellian panpsychist framework.

Goff expresses the motivation behind Russellian panpsychism particularly well:

[63] Is panpsychism synonymous with idealism? Traditionally, idealism has likewise been defined as mental monism. Pure panpsychism is a kind of mental monism. Thus, I agree with Chalmers (2020) that pure panpsychism can accurately be classified as idealism. I use the term 'panpsychism' primarily to avoid the many misconceptions and misrepresentations surrounding idealism and its many versions.

[64] The adjective 'Russellian' is in honor of Bertrand Russell who held similar views in some of his works, most notably in *The Analysis of Matter* (1927).

> All we get from physics is this big black and white abstract structure, which we metaphysicians must somehow color in with concrete categorical nature. Assuming the falsity of substance dualism, we know how to color in one bit of it: the brains of organisms are colored in with consciousness. How to color in the rest? The most elegant, simple, sensible option is to color in the rest of the world with the same pen. (2017: 171)

Constitutive panpsychism (as I understand it) is the thesis that the hypothetical fundamental experiences reductively ground and intelligibly explain everything else.[65] Metaphysically, constitutive panpsychists posit that only the fundamental experiences are Real, while all dependent physical entities and all dependent experiences (including human experiences) are Nominal. If constitutive panpsychism is true, all that God (metaphorically speaking) needs to do to recreate the cosmos is to recreate the fundamental experiences; everything else would follow, metaphysically 'for free.' Epistemically, they posit that there should be *no explanatory gaps* between the fundamental experiences and the rest of reality.

Putting the above together, we arrive at pure Russellian constitutive panpsychism. Panpsychism, so defined, is a version of Lightweight Anti-physicalism.[66] It is the most metaphysically simple version of panpsychism since it posits only experiences as fundamental. It is metaphysically elegant because it identifies the fundamental physical structure of the cosmos with the structure of the fundamental experiences. It is explanatory powerful since it promises a metaphysical picture where fundamental reality and human experiences fit together seamlessly, without explanatory gaps.

Panpsychists and philosophers sympathetic to panpsychism—as I have defined it above—include Chalmers (1996, 2015, 2020), Strawson (2006b, 2006a, 2015, 2020), Goff (2017, 2019, 2020), and Roelofs (2019), among others. Henceforth, unless

[65] Strictly speaking, in the literature, constitutive panpsychism is often defined only as a thesis of metaphysical reduction. My definition of constitutive panpsychism conflates constitutive panpsychism and type-A panpsychism. This is because I see little motivation to be a constitutive panpsychist without being a type-A panpsychist, and I do not know of anyone defending one without the other.
[66] See Chapter 3.

otherwise specified, by 'panpsychism,' I refer to pure Russellian constitutive panpsychism.

In the rest of this chapter, I argue that panpsychism is false. I argue that panpsychism faces an under-discussed explanatory gap between the fundamental experiences and at least some of the physical facts. I call this 'the missing objects problem.' I focus on *spacetime* as one aspect of the missing objects problem. I argue that, based on introspection, no human experience essentially has the metric structure of spacetime. Moreover, I argue that human experiences are a good model for the putative fundamental experiences. I conclude that there is an explanatory gap between the fundamental experiences and spacetime. If panpsychists cannot close this explanatory gap, panpsychism is false.

4.1.2. The Motivations for Panpsychism

I see three main motivations for panpsychism:

1. Taking introspection seriously.
2. Taking physical science seriously.
3. The belief that there is more to fundamental Reality than physical science describes.

First, all panpsychists take introspection seriously since they see introspection as *epistemically privileged*. Here, I use 'introspection' in a broad sense that covers (a) the act of *unmediated acquaintance* (the act of experiencing itself, as it happens), (b) the act of *mediated acquaintance* involving phenomenal concepts (thinking about experiences as they happen), and (c) the act of *reflection* on phenomenal concepts (thinking about experiences after they have happened). Despite their differences,[67] all panpsychists agree that introspection, either in one or more of its modes, can lead to substantive *metaphysical knowledge* about the essences of experiences. Taking introspection seriously leads panpsychists to accept the following thesis:

[67] For example, Strawson (2006a: 251–2) prefers to focus the (a), while Goff (2017) prefers to focus on (b) and (c).

> *The Cartesian Thesis*: Introspection reveals that the essences of experiences are experiential.

I have named the Cartesian Thesis in honor of Descartes, its first and most famous defender. It is both an epistemic and a metaphysical thesis. Epistemically, the Cartesian Thesis posits that introspection is a guide to metaphysical truth when it comes to experiences. Metaphysically, the Cartesian Thesis posits that experiences *are* essentially experiential, that experiences are essentially what they feel like.

Panpsychists are typically quite explicit in their acceptance of the Cartesian Thesis. For example, Goff (2017 chapter 5) explicitly defends it by positing that phenomenal concepts are transparent. Typically, the Cartesian Thesis is backed by arguments such as the conceivability argument. Strawson goes a step further and takes the Cartesian Thesis to be an "immovable" (2015: 169) and "fundamental given natural fact." (2006b: 4) In his view: "In having experience in the way we do, we are directly acquainted with certain features of the ultimate nature of reality." (2015: 188) Russell and Eddington seem to have shared Strawson's sentiment. In Russell's own words: "as regards the world in general, both physical and mental, everything that we know of its intrinsic character is derived from the mental side." (1927: 402)

Second, taking physical science seriously amounts to accepting the Causal Argument (from Chapter 1). Panpsychists are happy to agree with physicalists (a) that the physical universe is a causally closed whole, and (b) that physics describes the fundamental structure of the universe. Thus, in the panpsychist picture, just like in the physicalist picture, the Real fundamental entities are the entities that physics describes.

Third, the belief that there is more to fundamental Reality than physical science describes is the core posit of the Russellian metaphysical framework. The Russellian framework is based upon Russell's views on physical structure. In his own words:

> "Physics, in itself, is exceedingly abstract, and reveals only certain mathematical characteristics of the material with which it deals. It does

> not tell us anything as to the intrinsic character of this material." (1927: 10)

Eddington was one of the early adopters of the Russellian framework and expressed the same idea when he wrote: "The physical atom is, like everything else in physics, a schedule of pointer readings. The schedule is, we agree, attached to some unknown background." (1928: 130–1)

In the current literature, the Russellian framework is standardly characterized as the thesis that physics describes only the 'dispositional,' or 'extrinsic,' or 'quantitative,' properties of fundamental Reality, but is silent about its 'categorical,' or 'intrinsic,' or 'qualitative' properties. Following Montero (2015), I moreover refer to these 'hidden' fundamental properties as *inscrutables*. I elucidate this characterization by focusing on the relation between dispositional and categorical properties.

Dispositional properties are, roughly, the properties corresponding to *what things do*. Dispositional properties are individuated by their relations to other dispositional properties, typically via stimulus conditions and manifestations. Physics seems to characterize the world exclusively in dispositional terms. Simon Blackburn expressed this well in a famous passage:

> Resistance is par excellence dispositional; extension is only of use, as Leibniz insisted, if there is some other property whose instancing defines the boundaries; hardness goes with resistance, and mass is knowable only by its dynamical effects. Turn up the magnification and we find things like an electrical charge at a point, or rather varying over a region, but the magnitude of a field at a region is known only through its effect on other things in spatial relations to that region. (1990: 62–3)

Blackburn concludes that "science finds only dispositional properties, all the way down." (1990: 63) Proponents of the Russellian framework agree with Blackburn and understand all physical properties as purely dispositional.

Categorical properties, on the other hand, are roughly the properties corresponding to *what things are*. They are individuated by what they are in themselves and not by their relations to any other properties. Thus, they neither have stimulus conditions nor manifestations. There is a broad agreement in contemporary philosophy that phenomenal properties are categorical properties (or at least have a categorical aspect).

The Russellian framework, by itself, is neutral on the nature of the inscrutables. It is compatible with many different theories, positing different kinds of inscrutables. Panpsychism is perhaps the most popular and well-motivated such theory.[68] As entailed by the Cartesian Thesis, Russellian panpsychists think phenomenal properties are the one clear example of categorical properties that we transparently grasp. Putting together this and other motivations I discussed above, panpsychists posit that *inscrutables are experiential*.

4.1.3. Further Nuances

Panpsychism comes in many different forms. Even with all of the above stipulations, panpsychists still disagree on many other points. In this section, I briefly outline *three* such points of disagreement. This section is largely expository. Although I hope that readers will find it interesting and illuminating, taking a stance on any of the nuances I outline here is not essential to my arguments.

The first question I want to discuss is: what kind of experiences are the fundamental experiences? The simple answer is that *nobody knows*. Moreover, it seems that nobody can know with absolute certainty. As I have already said back in Chapter 1, experiences seem to be *epistemically private*, in the sense that a direct grasp of experiences is only available to the subjects undergoing those experiences. It is hard enough to infer or guess what other people and animals are experiencing, let alone what the cosmos itself might be experiencing.

[68] I explore non-panpsychists Russellian views in Chapter 5.

With that being said, we can still give an educated guess about what the fundamental experiences might be like. Goff, Seager, & Allen-Hermanson (2020) mention two options: (1) panexperientialism, the thesis that non-cognitive experiences (sensations, feelings, emotions, etc.) are fundamental, and (2) pancognitivism, the thesis that cognitive experiences (thoughts, beliefs, etc.) are fundamental. I agree with them that panexperientialism is the more plausible option. One, because it is somewhat controversial whether pure thought has a unique phenomenality of its own.[69] Second, and more importantly, our observations of conscious beings in nature suggest that the further down the evolutionary tree we go, the simpler the experiences. Animals appear to have simpler forms of consciousness than humans, while amoeba, if conscious at all, plausibly have even simpler experiences. Based on this, it appears it is most reasonable to assume that fundamental experiences are *very simple non-cognitive experiences*.

Goff gives a particularly vivid depiction of panexperientialism:

> [T]here seems nothing incoherent with the idea that consciousness might exist in very simple forms. We have good reason to think that the conscious experience of a horse is much less complex than that of a human being, and the experiences of a chicken less complex than those of a horse. As organisms become simpler perhaps at some point the light of consciousness suddenly switches off, with simpler organisms having no experience at all. But it is also possible that the light of consciousness never switches off entirely, but rather fades as organic complexity reduces, through flies, insects, plants, bacteria and amoeba. For the panpsychist, this fading-whilst-never-turning-off continuum further extends into inorganic matter, with fundamental physical entities – perhaps electrons and quarks – possessing extremely rudimentary forms of consciousness, to reflect their extremely simple nature. (Goff, 2019: 113–4)

[69] This amounts to there being cognitive phenomenology. My opinion is that there is cognitive phenomenology. Moreover, the philosophical consensus seems to be shifting towards acceptance of cognitive phenomenology.

The second question I want to discuss is: which entities are conscious if panpsychism is true? Most panpsychists agree that *all* the fundamental entities are conscious. However, most panpsychists do not think that all dependent entities are conscious. Instead, they typically argue that only *some* dependent entities are conscious.

Assuming the above is right, this second question for the panpsychist becomes: which of the dependent entities are conscious? We know that humans are conscious. Moreover, based on observation, we have good reasons to think that animals are conscious. In contrast, based on observation, many panpsychists would argue—and I am happy to agree with them—that molecules and tables do not seem to be not conscious. In general, most inorganic macrophysical entities do not appear to be conscious. Thus, it appears that the reasonable thing to suppose is that consciousness among the dependent entities is associated with living beings. Another reasonable assumption might be that all dependent entities that realize structures isomorphic to the nervous systems of living beings are conscious.[70]

The third and final question I want to discuss is: is there one fundamental subject or many? Here, panpsychists have two general options: micropsychism and cosmopsychism.

Micropsychism is the thesis that there are *many* fundamental subjects. It is based on a bottom-up approach (smallism) in terms of part-whole relations according to which *the parts ground the whole*. In the micropsychist picture, the many fundamental subjects combine, like Lego bricks, to give rise to the dependent subjects and, in general, to the world as we know it. Thus, if micropsychism is true, human minds are grounded in a plurality of fundamental parts, each of which is a subject in its own right.

Cosmopsychism is the thesis that there is *one* fundamental subject. The one fundamental subject is *the universe as a whole*. This view is based on a top-down approach in terms of part-whole relations according to which *the whole grounds the parts*. Independent of panpsychism, this metaphysical approach has an ancient

[70] A further and important question for any panpsychist theory is: what exactly determines which dependent entities are conscious. So far, only the panpsychist interpretation of the Integrated Information Theory (IIT) might have a precise answer to this question. See Mørch (2018b).

history, and in recent times has been most prominently defended by Johnathan Schaffer (2010) under the name of *priority monism*. Priority monism does not entail that the universe has no parts, but only that the universe is a whole that is *prior to* its parts. The parts are grounded in the whole as *aspects*,[71] like semicircles grounded in a circle or shades of red grounded in red. Thus, if cosmopsychism is true, human subjects are grounded as aspects of the whole universe as the one ubiquitous and fundamental subject.

I take cosmopsychism to be superior to micropsychism, based on both empirical and philosophical grounds. Empirically, micropsychism seems to fit better with the atomistic and corpuscular physics of the past. However, current physics seems to point to the existence of one Real and ubiquitous physical entity. The theory of relativity, on the super-substantivalist interpretation, posits that this entity is spacetime.[72] Everettian and GRW quantum theory posit that all matter is the manifestation of one physical entity (that corresponds to the wave function). Moreover, philosophically, cosmopsychism is simpler than micropsychism. It is the *simplest* conceivable version of panpsychism.[73]

[71] Therefore, priority monism should be distinguished from the stronger view that only the whole exists and has no parts. Such a view is called existence monism. See Schaffer (2010: 66).
[72] For a defense of super-substantivalism, see Schaffer (2009b).
[73] Due to similar reasons, many contemporary panpsychists seem to either endorse or lean towards cosmopsychism. Philosophers who sympathize with cosmopsychism include Chalmers (2020), Goff (2017), Kastrup (2018), Albahari (2020), Strawson (2006b, 2015, 2020), Shani (2015), and others.

4.2. Explanatory Gaps Once Again

4.2.1. Two Explanatory Gaps

In the two previous chapters, I examined and ultimately rejected both reductive and non-reductive versions of physicalism. I rejected these views primarily because they failed to sufficiently explain human experiences in virtue of fundamental Reality. In both cases, Consciousness Gap was crucial to my arguments.

Panpsychists promise to do better than both PCS physicalists and ground functionalists. They reject the assumption that Consciousness Gap obtains either because of the special features of phenomenal concepts or because explanatory gaps are everywhere. Instead, panpsychists argue that Consciousness Gap obtains because the physicalist conception of fundamental Reality is wrong. If so, physicalists have been putting the cart before the horse all along.[74] Panpsychists propose to Consciousness Gap by redefining our conception of fundamental Reality. However, as I will show, there are *two* serious problems with the panpsychist maneuver.

The panpsychist framework involves three essential elements: (1) fundamental experiences, (2) dependent experiences, and (3) physical entities. The fundamental experiences ground both the dependent experiences and the physical entities. Thus, *two metaphysical seams*—involving the relation between the fundamental and the dependent facts—are at the core of the panpsychist framework:

> *Seam-I*: The metaphysical relation between the fundamental and dependent experiences.
>
> *Seam-II*: The metaphysical relation between the fundamental experiences and the physical entities.

[74] In Chapter 5 I examine a different kind of physicalism, that I call extended physicalism, that cannot be accused of 'putting the cart before the horse' in this way.

Panpsychists face putative explanatory gaps at both Seam-I and Seam-II. Gap-I corresponds to Seam-I, while Gap-II corresponds to Seam-II.

> *Gap-I:* An explanatory gap between the fundamental and dependent experiences.
>
> *Gap-II:* An explanatory gap between the fundamental experiences and some physical entities.

Gap-I is typically explored in the literature as an aspect of the well-known '*combination problem*.' In contrast, Gap-II is either ignored or under-discussed. It corresponded to what I call the '*missing objects problem*.' My main goal in this chapter, after I briefly outline the combination problem, is to elucidate and put to use the missing objects problem as an argument against panpsychism. I do this by focusing on spacetime as one aspect of the missing objects problem.

4.2.2. The Combination Problem

The combination problem is roughly the problem of how the fundamental experiences give rise to dependent experiences. In Chalmers' words, it is the problem of: "How do the experiences of fundamental physical entities such as quarks and photons combine to yield the familiar sort of human conscious experience that we know and love?" (2016: 179)

The combination problem is well-known and well-discussed. The problem was named by William Seager (1995), although it has a long history, and versions of it can already be found in William James' *The Principles of Psychology* (1890). In contemporary debates, Goff (2006, 2009) and Sam Coleman (2014) have given influential formulations of the problem. Chalmers (2016) provides what is perhaps the most extensive discussion of the problem.

Some opponents of panpsychism take the combination problem to entail an incoherence in the very idea that experiences could be grounded in other

experiences.[75] This devastating metaphysical conclusion might be avoidable.[76] Nonetheless, at the very least, the combination problem threatens to undermine the motivation for panpsychism by undermining its explanatory merits. Gap-I entails that panpsychism fails to intelligibly explain human experiences, that panpsychism leaves human experiences shrouded in mystery. If so, there is a danger that panpsychism does no better at explaining human experiences than physicalism. If so, there might be no reason to prefer panpsychism over physicalism.[77]

Following Goff (2009) and Chalmers (2016: 187), the combination problem can be formulated in the form of a conceivability argument against panpsychism. I present a version of that argument as follows.

The *Anti-panpsychism Conceivability Argument*:

P-1. 'PP & ~Q' is conceivable.
P-2. If 'PP & ~Q' is conceivable, 'PP & ~Q' is possible.
P-3. If 'PP & ~Q' is possible, panpsychism is false.
C. Panpsychism is false.

In the argument above, I take 'PP' to stand for all the fundamental experiential facts, while 'Q' stands for some dependent experiential fact. Thus, P-1 of the Anti-panpsychism Conceivability Argument corresponds to Gap-I, expressed in terms of conceivability. The argument shows that even given the panpsychist redefinition of fundamental Reality, a duplicate of fundamental Reality lacking dependent experiences is both conceivable and possible. As Goff puts it:

> No matter what weird and whacky conscious states we attribute to your micro-level bits, it seems conceivable that those micro subjects exist in the absence of some further subject at the macro-level. (2017: 182)

[75] See Coleman (2014) for a powerful argument of this kind.
[76] See Chalmers (2016) for an overview of the panpsychist responses.
[77] Some panpsychists are happy to bite the 'mysteriousness' bullet. Strawson (2006a) is a notable proponent of this position. In his view, despite the explanatory worries highlighted in Gap-I, panpsychism still makes human experiences *less mysterious* than physicalism, and thus should be preferred over physicalism.

If so, despite posting fundamental experiences, panpsychism seems unable to integrate dependent experiences (such as our human experiences) within nature.

There have been several attempts by panpsychists to solve the combination problem.[78] Nevertheless, none of these proposed solutions seem to have been decisive. Thus, the combination problem remains a serious problem for panpsychism.

Discussing the nuances of the combination problem is beyond the scope of this dissertation. My goal here is different. Even if panpsychists manage to solve the combination problem, there is one more equally serious problem they have to solve. I call that problem the 'missing objects problem,' and I dedicate the rest of this chapter to it.

4.2.3. The Missing Objects Problem

Gap-II has so far received little attention in the literature. Since it currently lacks a name, I name it the *missing objects problem*. The missing objects problem is best seen as the hard problem of consciousness in reverse.[79] Roughly, it is the problem of whether the panpsychist can account for the physical structure of the cosmos in terms of the hypothetical fundamental experiences.

Against the panpsychist, I argue that if fundamental reality were purely experiential, some physical objects would lack a metaphysical explanation in terms of the fundamental ontology. If so, were we to rebuild the cosmos from pure experiences, these objects would be *missing* from our reconstruction of reality. The missing objects problem assigns the panpsychist the task of intelligibly explaining the existence of these objects.

Howell (2015), Pereboom (2011: 115), Chalmers (2020: 361–2, 365), Adams (2007), Goff (2017: 181–6), and Strawson (2020: 330) discuss versions or aspects of the missing objects problem. Howell frames the problem as a causal exclusion argument. Pereboom and Chalmers are naturally read as positing an explanatory gap. Chalmers

[78] For an overview see Chalmers (2016 section 7.6).
[79] In indebted to Miri Albahari for coming up with this expression.

is the most specific. He locates two important aspects to the problem: one involving spatiotemporal relations as groundees, and the other involving physical dispositions as groundees. Adams, Goff, and Strawson discuss space (or spacetime) as a potential problem for the panpsychist. They acknowledge that space (or spacetime) might not be grounded in experiences, and if so, this might lead to an impure version of panpsychism where both experiences and space (or spacetime) are fundamental.

None of the philosophers above frames the problem as a formal argument against the panpsychist conception of fundamental reality. In this chapter, I aim to do better. I focus exclusively on the grounding of spacetime as an aspect of the broader missing objects problem. Besides spacetime, there might be other ways, that involve different physical entities or their properties, to express the problem. The essential requirement of the missing objects problem is that the panpsychist must provide an, at least in principle, intelligible explanation of *all* the currently available physical facts, without exceptions.

Both the combination problem and the missing observables problem entail cracks in the elegant theoretical structure of panpsychism. The missing objects problem strikes at the very core of panpsychism, at the panpsychist conception of fundamental reality as wholly experiential. So far, no panpsychist has given a convincing and detailed story of how spacetime could be essentially experiential. I challenge panpsychists to give such a story and take the missing objects problem seriously as a whole.

4.3. The Spacetime Argument

4.3.1. The Spacetime Gap and Argument

The general theory of relativity (GTR)—on its most straightforward ontological interpretation—entails the existence of *spacetime*.[80] GTR, so understood, is a *substantivalist* theory of spacetime. As Schaffer (2009: 132) reports, spacetime substantivalism is the consensus view among philosophers of physics. Moreover, Albert Einstein himself, despite some early doubts, explicitly endorsed spacetime substantivalism.[81] Spacetime substantivalism entails that spacetime is a *fundamental entity* that is ubiquitous and dynamic. We live in spacetime,[82] and we observe the effects of spacetime on all material objects in the form of gravity.

GTR is a theory of the geometry of spacetime.[83] It characterizes spacetime as a *four-dimensional manifold* structured by the *spacetime metric*. [84] In geometry, a metric is a structure that determines the *shortest continuous path* between any two points in a set. In simple terms, the metric determines *distances*. The spacetime metric is the metric of spacetime. Thus, it determines the distance between any two spacetime points.[85]

The spacetime metric is an obvious candidate for an essential property of spacetime. It is ubiquitous and obtains in all regions of spacetime. It is invariant for all observers in all frames of reference, and it determines the causal structure of spacetime.[86] In the

[80] The same is true of the special theory of relativity.
[81] See Einstein (1920/2007).
[82] Moreover, if super substantivalism is true, we are spacetime. See Schaffer (2009b).
[83] See Maudlin (2012) for an accessible introduction to both the special and general theory of relativity.
[84] Specifically, spacetime is a Lorentzian manifold, a type of a pseudo-Riemannian manifold with a metric signature of (1, 3) or (3, 1), depending on the convention.
[85] Strictly speaking there are many spacetime metrics. All spacetime metrics are calculated via the Einstein field equations. What unites all the spacetime metrics and makes them a unique kind of metric is that they all share the *same metric signature*. In light of this, I use 'spacetime metric' to refer to any metric that has the $(1, n-1)$ metric signature.
[86] In the technical jargon, the spacetime metric determines the *light-cone* structure of spacetime.

literature, the metaphysical position postulating that the spacetime metric is an essential property of spacetime is known as *metric essentialism*.[87]

I will argue that the spacetime metric is not an essential property of human experience (section 4.3.3). Moreover, assuming that panpsychism is true, I argue that human experiences are a good model for the putative fundamental experiences (section 4.3.6). If so, panpsychism faces the following explanatory gap:

> *Spacetime Gap*: There is an explanatory gap between the fundamental experiences (as grounds) and spacetime (as a groundee).

I will use Spacetime Gap to argue that panpsychism is false. If panpsychism is true, the structure of the fundamental experiences *must be isomorphic* to the physical structure of the cosmos. However, Spacetime Gap entails that the structure of spacetime is *not isomorphic* to the structure of the fundamental experiences. Spacetime Gap entails that a scenario in which all the facts of the fundamental experiences obtain, yet no spacetime metric facts obtain is conceivable. In this scenario, the cosmos would lack the spacetime metric. Thus, Spacetime Gap entails that the fundamental experiences are not the ground of spacetime. If so, it is not enough that God creates the fundamental experiences to recreate our world. God would have to do more: God would need to create spacetime in addition to the fundamental experiences. If so, panpsychism is false.

I call the resulting argument the '*Spacetime Argument':*

P-1. *The Ground of Spacetime:* If panpsychism is true, some fundamental experience grounds spacetime.
P-2. *No Spacetime Metric:* Human experiences do not essentially have the spacetime metric.
P-3. *Spacetime Metric:* Spacetime essentially has the spacetime metric.

[87] See Maudlin (1989, 1990) for canonical defenses of metric essentialism.

P-4. *Explanatory Gap:* There is an explanatory gap between grounds that lack the spacetime metric and groundees that have the spacetime metric.

P-5. *Similarity:* If the fundamental experiences have an essential metric, they have the same essential metric as human experiences.

C-1. *Spacetime Gap:* There is an explanatory gap between the fundamental experiences (as grounds) and spacetime (as a groundee).

P-6. *No Gaps:* There are no explanatory gaps in true cases of grounding.

C-2. Panpsychism is false.

The Spacetime Argument has two parts. The first part establishes an epistemic conclusion in C-1 based on premises 1 through 5. The second part builds upon C-1 and, via premises 1 and 6, establishes a metaphysical conclusion in C-2.

The Spacetime Argument is a contemporary extension of an old idea—that goes back to at least Descartes—that no experience is essentially spatial. The Spacetime Argument extends this idea to spacetime and puts it to use against panpsychism. If sound, the argument refutes panpsychism. In the rest of the paper, I offer a further defense of each of the argument's premises.

4.3.2. Premise I: The Ground of Spacetime

P-1, The Ground of Spacetime, states: If panpsychism is true, some fundamental experience grounds spacetime. I take P-1 to follow from the panpsychist conception of fundamental Reality. All panpsychists agree that the physical structure of the cosmos is the structure of the fundamental experiences, as observed by humans (and dependent subjects in general).

To give an analogy, let us assume that panpsychism is true, and the only fundamental entity in the universe besides me is my coffee mug. According to panpsychists, my coffee mug undergoes experiences. Moreover, my coffee mug has all of its physical properties (shape, mass, etc.) in virtue of its experiences. The physical structure of my coffee mug *is* the observable structure of its experiences.

Given spacetime substantivalism, spacetime is a fundamental entity. Thus, spacetime must be grounded in some fundamental experience. If so, the geometric structure of spacetime, the structure that GTR describes, must be the indirectly observable structure of some fundamental experience. Thus, I take it that P-1 is uncontroversial for all panpsychists sympathetic to spacetime substantivalism.

4.3.3. Premise II: No Spacetime Metric

P-2, No Spacetime Metric, states: Human experiences do not essentially have the spacetime metric. I base P-2 on the evidence from introspection. In light of the Cartesian Thesis (from section 4.1.2), panpsychists agree that introspection reveals the essences of experiences. I will defend P-2 *inductively* via an introspective analysis of a sample of human experiences. This analysis can be repeated on any experience that I have left out.

An experience essentially has the spacetime metric iff its phenomenal character necessarily instantiates the spacetime metric. My defense of P-2 is two-fold and involves a separate analysis of simple and complex human experiences. First, I argue that simple experiences do not essentially have any metric. Second, I argue that although complex experiences might have an essential metric, no complex experience essentially has the spacetime metric. I conclude that neither simple nor complex human experiences have the spacetime metric essentially. If so, P-2 is true.

The standard practice in philosophy is that *conceivability* is a good guide to determining facts about essences. As I already mentioned,[88] conceiving an entity under a transparent concept entails conceiving of its full essence. The consensus among panpsychists seems to be that phenomenal concepts are transparent.[89] Thus, in what follows, I assume the availability of transparent phenomenal concepts for all human experiences.

[88] In Chapter 2.

[89] Even panpsychists like Strawson, who deny the full grasp of the essences of experiences, agree that we grasp some of the essential properties of experiences. Thus, even these panpsychists should agree that we can form transparent concepts of the essential properties of experiences that we do grasp.

With the above in mind, I demonstrate my claims with a simple conceivability exercise. An experience essentially has the spacetime metric iff that experience is inconceivable without the obtaining of spacetime metric facts.

I begin my analysis with *simple experiences*. By 'simple experiences,' I refer to all tokens of human experiential types that appear to have no proper parts.

Colors appear to be a paradigmatic case of simple experiences. Moreover, at least prima facie, colors appear to have some geometric structure. Colors can neither be imagined nor perceived unless they are extended in some way. Even a colored point has some extension, takes up some space in the visual field (or the field of imagination). That might be so, yet, I argue, individual colors have no essential metric structure. I lustrate this with the example of RED (I use capitals to refer to experiences in terms of their phenomenal characters).

Imagine an infinite two-dimensional RED plane (or simply, imagine that everything you see is RED). What is the metric of this plane? To answer this question, we must find the shortest continuous path between two points on the plane. If the plane is Euclidean, the path between two points will appear straight. If the plane is non-Euclidian, this path might appear curved[90] or even jagged.[91] If the plane is purely topological, there will be no fixed fact about such a path. Whatever turns out to be the case, these instances of the RED plane are *visually indistinguishable*. In all these cases, I experience the same RED. This example shows that RED alone reveals nothing about the putative metric of the plane. Whether or not the plane has a metric is a *further fact* not entailed by RED. Thus, RED is conceivable without the obtaining of any facts about distance. Thus, RED has *no* essential metric. The same is true, mutatis mutandis, for all the other colors.

The case is even clearer when it comes to other simple experiences beyond colors. Take any instance of the other *sensory experiences*, or take any *pain* or *pleasure*, or any *emotion*, *desire*, or *will* as an example. All tokens of these experiential types seem conceivable without the obtaining of facts about distance. In all of these cases, it

[90] For example, if the plane is hyperbolic and has the Poincaré metric.
[91] For example, if the plane is flat and has the 'taxicab metric.'

appears to be a category mistake even to apply the concept of distance. If I am right about this, it should be inductively obvious that simple human experiences have no essential metric, let alone the spacetime metric.

Next, onto *composite experiences*. By 'composite experiences,' I have in mind all experiences that can be broken down into simple experiences. Composite experiences include pluralities of simple experiences at a given time (think of a painting) and pluralities of simple experiences that evolve over time (think of a movie).

We ordinarily perceive reality as a composite experience, a conglomerate of different sensory experiences. Metaphorically speaking, the manifest image is more like a painting than paint, more like a movie than a frame. Perhaps we might have never developed a concept for distance if we only experienced isolated simple experiences (like RED or C#). However, we ordinarily perceive composite experiences, and we have a concept for distance that applies perfectly well to reality. Consider the experience of a Euclidean triangle or the experience of a walk in the park. Experiences such as these appear inconceivable without the obtaining of facts about distances. Thus, it seems, at least some composite experiences might have an essential metric. Yet, I argue, no composite experience essentially has the spacetime metric.

Visual experiences are the paradigmatic example of composite experiences associated with distance perception. What is the metric structure of visual space? The historical consensus was that visual space is Euclidean.[92] However, growing empirical data has shattered this consensus. As Farid Masrour (2015: 1821–2) reports, Walter Blumenfeld's (1913) experiments were especially influential in this. Blumenfeld's experiments were among the first to show that visual space violates Euclid's parallel lines postulate. Since then, there has been an ongoing debate about the exact geometry of visual space.

Many contemporary researchers argue that visual space has no fixed geometry but can instantiate many different geometries depending on the external stimuli.[93] Mark

[92] Kant was perhaps the most radical in this. He famously claimed that Euclidean geometry is both necessary and a priori.
[93] Wagner (2006) is a notable proponent of this view.

Wagner (2006) is a notable proponent of this view. As he puts it: "Our goal should be to find not *the* geometry of visual space, but the *geometries* of visual space." (2006: 230) Some of these geometries (some subfields of visual space, in some contexts) might be Euclidean, while others are non-Euclidean. If so, it is reasonable to assume that some visual experiences might essentially have a Euclidean metric, while others might essentially have various non-Euclidean metrics or even no metric.

There is no empirical evidence that visual space, in any context, instantiates the spacetime metric. This should come as no surprise, given the properties of the spacetime metric. As I will now explain in more detail, the spacetime metric has properties that are unlike the properties of any metric we can visualize.

The spacetime metric is compatible not only with *positive distances* (such as the distances we are familiar with from experience) but moreover with *null distances* and *negative distances*. The distinction between positive, negative, and null distances gives spacetime its unique 'light-cone' structure. In the technical jargon, if the spacetime distance between two events is positive, those two events are *timelike* separated. If the spacetime metric is negative, those two events are *spacelike* separated. If the spacetime metric is null, those two events are *lightlike* separated.[94] [95]

For any spacetime point, there is a perfectly natural reference frame that distinguishes between events that are at a positive, a negative, and a null spacetime distance from the origin. This is true at any scale and for any spacetime region, from the cosmos as a whole to the sub-atomic level. For example, consider two events A and B occurring in different regions of the brain. If A and B occur at the same time, A and B are spacelike separated in the instant they occur. This is because light cannot traverse the spatial distance between A and B instantaneously. However, if we know that A and B are causally connected, then light must have been able to traverse the spatial distance between A and B. In this case, A and B must be either lightlike or timelike separated.

[94] Here I am assuming the 'West-coast' convention: (+, -, -, -).

[95] The association of timelike separation with positive distances and of spacelike separation with negative distances is purely a matter of mathematical convention. However, the distinction between positive, negative, and null distances, and thus the light-cone structure of spacetime is *not a convention*. It is a real property of spacetime.

There is no counterpart to the light-cone structure in visual experience. We never visually experience null or negative distances. That is why visual models of spacetime, although very useful as heuristic tools, fail to convey the true structure of spacetime. All this clearly shows that visual experiences are conceivable without the obtaining of the spacetime metric. If so, visual experiences do not essentially have the spacetime metric.

Beyond visual composite experiences, the same appears to be true for all other composite experiences. We could try to add and mix experiences to conjure up a composite experience that might essentially have the spacetime metric. Yet, I am doubtful that such an attempt would be fruitful. I can think of no case of a composite experience that is inconceivable without the spacetime metric. Although rationally coherent, the null and negative distances of the spacetime metric seem *impossible to experience*. If so, the spacetime metric is beyond the scope of human experience. In the absence of a counter-example, inductively, I conclude that neither simple nor composite human experiences essentially have the spacetime metric. Thus, no human experience essentially has the spacetime metric. Thus, P-2 is true.

4.3.4. Premise III: Spacetime Metric

P-3, Spacetime Metric, states: Spacetime essentially has the spacetime metric. P-3 follows from metric essentialism. If spacetime substantivalism is true, the spacetime metric is certainly the most obvious candidate for an essential property of spacetime. The spacetime metric is ubiquitous, invariant for all observers, and determines the causal structure of spacetime. Thus, I take P-3 to be uncontroversial for anyone sympathetic to spacetime substantivalism.

Beyond plain obviousness, P-3 can be more formally defended via a conceivability exercise. If the spacetime metric is essential to spacetime, spacetime should be inconceivable without the spacetime metric for a sufficiently informed thinker.[96] It seems that this is clearly the case. Namely, for anyone who knows enough about GTR

[96] Our current <spacetime> concept is likely translucent (in Goff's terminology) given that we still do not know the full geometry of spacetime. Remember (from Chapter 2), a translucent concept reveals a *part* of its referent's essence.

(or even the special theory of relativity), spacetime is inconceivable without the obtaining of the spacetime metric. If spacetime's metric structure were different, GTR would no longer apply to spacetime; spacetime would no longer be the same entity, it would no longer *be* spacetime. If so, P-3 is true.

4.3.5. Premise IV: Explanatory Gap

P-4, Explanatory Gap, states: There is an explanatory gap between grounds that lack the spacetime metric and groundees that have the spacetime metric. I have already defined explanatory gaps as follows:

> *Explanatory Gaps*: an explanatory gap obtains iff there is *no* a priori entailment between a ground and a groundee in a putative case of grounding.

Moreover, as I have previously argued (in Chapter 1), conceivability is an excellent guide to locating explanatory gaps. If it is conceivable that the ground could obtain without the groundee obtaining, there is an explanatory gap between that ground and that groundee.

I already demonstrated the conceivability of human experiences without the obtaining of the spacetime metric (in section 4.3). Thus, it is clear that there is an explanatory gap between human experiences (as grounds) and the spacetime metric (as a groundee). Thus, the explanatory gap entailed by P-2 already confirms P-4.

P-4 is moreover justified by the fact that different metrics describe different structures. Take the 3D Euclidean metric as an example. For any two points in 3D Euclidean space, with coordinates (x_1, y_1, z_1) and (x_2, y_2, z_2), the Euclidean distance 'ΔE^2' between them is:

$$\Delta E^2 = \Delta x^2 + \Delta y^2 + \Delta z^2$$

In contrast, the counterpart spacetime metric 'ΔM^2' for a flat 3D Minkowski spacetime (in a vacuum), for any two spacetime points with coordinates (x_1, y_1, t_1) and (x_2, y_2, t_2) is:

$$\Delta M^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2$$

The two equations are clearly dissimilar. For one, only ΔM^2 is determined by the constant 'c' (standing for the speed of light), while ΔE^2 is not. Second, the time coordinate 't' in ΔM^2 is of special significance, while no coordinate in ΔE^2 has a special significance over the other coordinates. Contrasting the 3D Minkowski metric to the 3D Euclidean metric clearly shows that the spacetime metric is not the Euclidean metric plus a time coordinate. Instead, the two metrics describe very different structures. An entity that satisfies ΔE^2 does not tell us anything about ΔM^2. The Euclidean metric ΔE^2 is clearly conceivable without the obtaining of the spacetime metric ΔM^2. Thus, there is nothing in ΔE^2 that would indicate that it a priori entails ΔM^2. Similar considerations seem to apply to cases involving any other non-spacetime metric. Thus, P-4 is true.

4.3.6. Premise V: Similarity

P-5, Similarity, states: If the fundamental experiences have an essential metric, they have the same essential metric as human experiences. I base P-5 on the fact that panpsychists posit the fundamental experiences as a metaphysical explanans of human experiences. The putative fundamental experiences are unknown, while human experiences are intimately known. Thus, human experiences are our only guide to what the fundamental experiences might be like. P-5 is an instance of the following broader thesis:

> *Good Model*: the putative fundamental experiences must be sufficiently similar to human experiences; at least in principle, there should be no explanatory gap between them and human experiences.

Good Model entails that human experiences are a good model for what the fundamental experiences might be like. I take it that this requirement is imposed on panpsychism by one of its primary motivations, the motivation to explain human experiences better than physicalism.

Gap-I and the combination problem only further reinforce this requirement. Imagine a scenario where all fundamental experiences are sounds and all human experiences are colors. Although both grounds and groundees are experiences in this scenario, there seems to be no intelligible connection between the grounds and the groundees. It is hard to see how Gap-I could ever be closed if such a scenario were true. The rejection of Good Model amounts to accepting the plausibility of this and similar scenarios by the panpsychist. Good Model puts methodological constraints on panpsychism and safeguards the theory from such scenarios.

Good Model entails that it is not sufficient to posit that the fundamental experiences feel in *some* way. Instead, the fundamental experiences must feel a *specific* way. They must feel in a way that is intelligibly connected to the way human experiences feel. P-5, as an instance of Good Model, highlights this requirement in terms of the structure of individual experiences. P-5 entails that the fundamental experiences must have phenomenal characters with structures of the *same kind* as the structures of human experiences. Human experiences do not essentially have the spacetime metric structure. If so, it is reasonable to posit that the fundamental experiences do not essentially have the spacetime metric structure.

The above pattern of explanation is already proven to work in other cases of grounding. For example, to explain a wall's existence in grounding terms, it is not sufficient to posit that the wall is grounded in bricks arranged in *some* structure. Instead, the bricks must be arranged in a *specific* structure, the kind of structure that could intelligibly explain the wall's existence.

Accepting Good Model and P-5 makes panpsychism less hand-wavy. Moreover, it makes closing Gap-I and solving the combination problem a reasonable goal. In contrast, rejecting Good Model and P-5 threatens to transform Gap-I (and with it, the

combination problem) into an *uncrossable chasm*. Thus, panpsychists should be happy to accept both Good Model and P-5.

Some panpsychists might find P-5 controversial. For example, Strawson (2006a, 2020) explicitly defends the view that the combination problem (and with it, Gap-I) does not diminish the explanatory value of panpsychism. Nevertheless, Strawson strongly opposes 'radical emergence.'[97] In his view, the emergence of experiences from non-experiential entities is a prime example of radical emergence. In his earlier work (2006b: 17), he uses the emergence of spatial from nonspatial properties as another clear example of radical emergence. Strawson does not use the terminology of grounding. I take it that roughly, in grounding terms, what he means by 'radical emergence' is a grounding relation where the groundee has radically different properties from the ground. Thus, radical emergence appears to entail an explanatory gap between the involved ground and groundee. If so, the grounding of non-metric entities in metric entities is a plausible example of radical emergence. If so, I assume that even Strawson should be sympathetic towards Good Model and with it towards P-5.

In a nutshell, the price of rejecting P-5 is the acceptance of radical emergence. Perhaps some panpsychists—the so-called emergentist panpsychists—would be happy with this option. However, the panpsychists that my argument targets—the so-called constitutive panpsychists—would certainly not be happy with this result.

4.3.7. Premise VI: No Gaps

P-6, No Gaps, states: There are no explanatory gaps in true cases of grounding. I have already argued in support of P-6 in Chapter 3. Moreover, as I have already mentioned, many philosophers might consider P-6 controversial. PCS physicalists would argue that a posteriori necessities involving phenomenal concepts are one counter-example to P-6. Schaffer wholly and explicitly rejects P-6. Nevertheless, I take it that panpsychists must accept a version of P-6 since otherwise, panpsychism would be undermotivated.

[97] Strawson's notion of radical emergence roughly matches the more commonly used notion of *strong emergence*.

Remember, it was Consciousness Gap that took us to panpsychism from physicalism in the first place. Chalmers (2015: 249–50), Goff (2017: 147–9), and Strawson (2006b: 12–21, 2020: 331–2) agree on this. Thus, without accepting a premise in the vicinity of P-6, the panpsychists' rejection of physicalism might be undermotivated. With that, panpsychism itself would be undermotivated. If so, proponents of panpsychism must accept a version of P-6.

4.4. The Spacetime Functionalism Objection

4.4.1. Objections

I consider all premises of the Spacetime Argument, except for P-1: The Ground of Spacetime, to be either uncontroversial or sufficiently justified. P-2: No Spacetime Metric is justified by an introspective analysis of human experience. P-3: Spacetime Metric is uncontroversial since it states an obvious assumption about the putative essence of spacetime. P-4: Explanatory Gap is justified by the examples I outlined in defense of P-2 and by the explanatory gap between the Euclidean and Minkowski metrics. P-5: Similarity is justified by the fact that, if panpsychism is true, the putative fundamental experiences must intelligibly explain human experiences. P-6: No Gaps is justified since, without it, panpsychism would be undermotivated.

With the above in mind, there is one powerful objection to P-1: The Ground of Spacetime that requires an in-depth response. My opponent could contest P-1 by rejecting the fundamentality of spacetime and spacetime substantivalism. I dedicate this section to the defense of P-1, The Ground of Spacetime, from potential objections.

4.4.2. Spacetime Functionalism

The rejection of P-1 amounts to the acceptance of *spacetime functionalism*. Spacetime functionalism is roughly the thesis that spacetime is whatever plays the spacetime role. If spacetime functionalism is true, spacetime is not fundamental. Instead, spacetime is a dependent entity grounded in some more fundamental physical structure. If so, spacetime is not directly grounded in any fundamental experience. Instead, only the ground of spacetime is essentially experiential. If so, I have failed to disprove panpsychism with the Spacetime Argument. In the literature on panpsychism, Chalmers (forthcoming) explicitly endorses spacetime functionalism, while Strawson (2020: 325) appears sympathetic towards it.

Spacetime functionalism comes in many forms. The different forms are the result of positing different physical structures as spacetime's putative ground. The two most

common approaches posit (a) the wave function from quantum theory or (b) structures from quantum gravity theories. Approach (a) assumes that time is fundamental and aspires to explain only the space aspect of spacetime functionally. Approach (b) is more radical and typically posits grounds that are neither spatial nor temporal. In what follows, I illustrate my response to spacetime functionalism by focusing on approach (a). This is because approach (a) is more common than (b), and because quantum theory has plenty of empirical support (unlike quantum gravity). My main point applies—mutatis mutandis—even if one of the candidate theories of quantum gravity turns out to be true.

The wave function is the central mathematical device of quantum theory; it is the function that describes the dynamical properties of quantum systems and is behind a lot of the impressive empirical successes of the theory. [98] The domain of the wave function is standardly called a *configuration space*. It is typically characterized as *3N dimensional*, where N stands for the number of particles it describes. The universe is estimated to have at least 10^{80} particles. If so, the configuration space of the universe's wave function has at least *three* × 10^{80} *dimensions*. I refer to this conception of space as follows:

HD-Space: the 3N-dimensional space of the wave function.

Proponents of *realist* quantum theories agree that the wave function describes a real physical entity.[99] Following Tim Maudlin (2019), I call this object the '*quantum state*.' However, the properties of the quantum state are a matter of heated debate. Specifically, there is much disagreement about whether the space of the quantum state is the three-dimensional space of spacetime (henceforth, 3D-Space) or HD-Space.

David Albert (1996, 2015, 2019) and Alyssa Ney (2012, 2020), among others, argue that HD-Space is more fundamental than 3D-Space. In their view, HD-Space grounds

[98] For a detailed introduction to quantum theory, see Maudlin (2019).
[99] Realist quantum theories include the Bohmian, Everettian, and GRW theory.

3D-Space.[100] I call their position *high-dimensionalism*. In contrast, Maudlin (2007, 2013) and Eddy Chen (2017), among others, argue that 3D-Space is more fundamental than HD-Space. In their view, 3D-Space grounds HD-Space. I call their position *three- dimensionalism*.

If high-dimensionalism is true, space functionalism might likewise be true. If so, experiences need not ground spacetime, but instead, must ground HD-Space (and perhaps time, as a distinct entity). If so, even if Spacetime Gap is true, there might be no explanatory gap between experiences and HD-Space. If so, via HD-Space, experiences could indirectly ground spacetime, and panpsychism would be true.

4.4.3. Defending Premise I: The Ground of Spacetime

My response to space functionalism is two-pronged. First, there seems to be an explanatory gap between HD-Space and 3D-Space, and it is highly uncertain whether this gap can be closed. Second, it is highly uncertain whether experiences can ground HD-Space. Together, these two claims entail that adopting space functionalism against P-1 provides at best a tentative defense of panpsychism.

First, there appears to be *an explanatory gap* between HD-Space and 3D-Space. HD-Space—if it can even be called 'space'—is *vastly different* from 3D-Space. High-dimensionalists and three-dimensionalists agree on this point. As Albert puts it, there is a "particularly radical coming-apart" (2019: 95) between 3D-Space and HD-Space. Similarly, Chen (2017: 338) compares HD-Space to a 'Platonic Heaven' in relation to 3D-Space. Moreover, as Ney points out, "[n]o three dimensions of the configuration space of quantum theory correspond to the three dimensions of our manifest image" (2012: 540).

The many dimensions of HD-Space are not simply extra dimensions in addition to the three dimensions of 3D-Space. Instead, these are different dimensions, expressing different information, with no correlates among the dimensions of 3D-Space. In

[100] In his early work, Albert (1996) posits that 3D-Space is an illusion. Ney (2012) is sympathetic to this view. In their later work, Albert (2015, 2019) and Ney (2020) defend views where 3D-Space (in the case of Albert) or 3D objects (in the case of Ney) are ontologically real yet grounded in HD-Space.

Maudlin's (2007, 2019) and Bell's (1987) terminology, high-dimensionalists posits no *'local beables,'* i.e., posits no 3D-Space phenomena in their fundamental ontology.

Despite a few valiant attempts,[101] at present, space functionalists are yet to convincingly close this explanatory gap. This is a serious problem for the high-dimensionalist. As Maudlin (2007, 2019) argues, the explanatory gap between HD-Space and 3D-Space threatens to undermine the *empirical salience* of high-dimensionalism. Quantum theory is a theory about local beables in 3D-Space. High-dimensionalism posits no local beables in its fundamental ontology. Thus, the only way for high-dimensionalism to make contact with local beables is by closing the explanatory gap towards 3D-Space. Given these worries and given panpsychists' commitment to no explanatory gaps in true cases of grounding, space functionalism is, at present, a less than ideal ally for the panpsychist.

Second, even assuming that the above explanatory gap can be closed, it is highly uncertain whether experiences can ground HD-Space. HD-Space might be *too complex* to be grounded in experiences. If high-dimensionalism is true, panpsychists must explain how experiences could ground a structure that is three times 10^{80} dimensional. Clearly, no human experience is three $\times$ 10^{80} dimensional. It is impossible to imagine anything even remotely close to HD-Space. If so, an argument analogous to the Spacetime Argument, where the high-dimensional quantum state plays the role of spacetime, can be put to use against panpsychism. If so, the quantum state, on the high-dimensionalist conception, is one more aspect of the missing objects problem for the panpsychist. The panpsychist *must* tell a detailed story of how the fundamental experiences might ground HD-Space. Otherwise, the promise of space functionalism amounts to nothing but wishful thinking for the panpsychist.

[101] See Albert (1996, 2015) and Ney (2020).

4.5. Is Panpsychism Worth the Trouble?

Panpsychism promises a lot. However, it also faces serious difficulties in the form of the combination problem and my missing objects problem. The combination problem indicates that there might be a metaphysical gap between the fundamental and the dependent experiences. The missing objects problem indicates that there also might be a metaphysical gap between the fundamental experiences and some physical entities. If these two challenges are sound, they show that panpsychism can account neither for human consciousness nor for all physical entities.

I am very sympathetic towards panpsychism. I support and understand the motivation behind the view. Pure physicalism posits pure structure as fundamental, and as I have argued, struggles to metaphysically explain human consciousness. Dualism seems to break physical causal closure. Given these problems with physical and dualism, posting the Russellian framework makes perfect sense. Thus, in this regard, I am inclined to agree with Strawson that: "Physics gives the structure, but not the structure-transcendent nature, of the thing that has the structure." (2015: 172)

The lesson from this chapter is that panpsychism opens more problems than it closes. Nevertheless, the failure of panpsychism does not entail the failure of the Russellian framework. Neither the combination problem nor the missing objects problem generalizes beyond panpsychism. These problems give valuable insights into the features that any good candidate for a fundamental inscrutable must have. Thus, although panpsychism is not itself the solution, it might show us the way forward to a potential solution to the problem of consciousness.

With this chapter, I hope to have shown that a good Russellian theory must metaphysically explain both:

1) Dependent minds (and thus resolve the combination problem).
2) All physical entities (and thus resolve the missing objects problem).

There might be other candidates for fundamental inscrutables that are apt to satisfy these requirements. Some version of *neutral monism*, *panprotopsychism*, *impure panpsychism* (the view that the inscrutables are only partially experiential), or *impure*

physicalism (physicalism that posits physical inscrutables) might be true. Moreover, I still have not given *substance dualism* its proper due. Therefore, before drawing my final conclusions, I set out to explore all of these theories in the dissertation's final chapter.

Chapter IV:
The Way Forward

"Without consciousness the mind-body problem would be much less interesting. With consciousness it seems hopeless." (Nagel, 1974: 436)

"Philosophy is still young, and the human capacity for reasoning is strong. In a scrutable world, truth may be within reach." (Chalmers, 2012: p. xxiii)

"The truth about physical objects must be strange." (Russell, 1912/1959: 19)

5.1. The Story So Far

I examined pure physicalism and pure panpsychism as potential solutions to the hard problem of consciousness. Pure physicalism and pure panpsychism both face *explanatory gaps* that undermine their *metaphysical structures*. Pure physicalism (of both the reductive and non-reductive variety) faces Consciousness Gap. Pure panpsychism faces the combination and missing objects problems (specifically, Spacetime Gap).

> *Consciousness Gap*: The physical facts (as grounds) do not a priori entail the experiential facts (as groundees).
>
> *Spacetime Gap*: There is an explanatory gap between the fundamental experiences (as grounds) and spacetime (as a groundee).

If both pure physicalism and pure panpsychism are false, what hope, if any, is there for solving the problem of consciousness? As it turns out, *there is still plenty of hope*. Pure physicalism and pure panpsychism do not exhaust all the available options. In this chapter, I will briefly consider and assess *seven* other potential solutions to the problem of consciousness. These are naturalistic dualism, supernatural dualism, epiphenomenal dualism, extended physicalism, panprotopsychism, neutral monism, and impure panpsychism. I will evaluate these theories using the same standards as before. Views positing that human consciousness is dependent must be apt to close Consciousness Gap from the metaphysical ground up (or at least reduce its opacity). Views positing that consciousness is fundamental must be apt to solve both the combination and the missing objects problem.

My explanatory gap centered approach naturally favors some views over others. If the reader rejects my emphasis on explanatory gaps, I expect they will find my assessment unsatisfactory. If this is the case, I hope that the reader will at least find my insights interesting.

A few points on terminology. I will examine *four Russellian views*. These are Russellian physicalism, Russellian panprotopsychism, Russellian neutral monism, and Russellian impure panpsychism. In the philosophy of mind literature, the conjunction of Russellian panpsychism and Russellian panprotopsychism is typically called *Russellian monism* (Alter & Nagasawa, 2015; Chalmers, 2015). However, I disagree with this classification. First, because as I will soon show, panpsychism and panprotopsychism face distinct philosophical problems. Second, because there are Russellian views that are *neither* panpsychist nor panprotopsychist. The versions of neutral monism and impure physicalism that I examine here are such Russellian theses. Broadly speaking, I find it unreasonable to restrict the Russellian label only to views that deal exclusively with consciousness. Thus, I will use the label Russellian monism to refer to the conjunction of *all Russellian monist views* and not just the ones that are motivated by the problem of consciousness.

In the discussion that follows, I *will not discuss idealism*. Idealism is roughly the thesis that all Real fundamental entities are minds. Thus, so construed, idealism is a thesis of *mental monism*. As I already mentioned in the previous chapter, idealism, so construed, is *synonymous with pure panpsychism*. Thus, everything I said so far about pure panpsychism applies to idealism.

With the above terminological considerations aside, I would like to stress that my goal in this chapter is not to solve the problem of consciousness. All of the theories that I will discuss in this chapter, just like the theories I have discussed so far, face serious problems. I agree with Chalmers that:

> No position on the mind–body problem is plausible. Materialism is implausible. Dualism is implausible. Idealism is implausible. Neutral monism is implausible. None-of-the-above is implausible. But the probabilities of all of these views get a boost from the fact that one of them must be true. (2020: 370)

My ambition is only to show that some theories of consciousness are more promising than others. Those theories offer *hope for a solution*.

5.2. Alternatives

5.2.1. Dualism

Throughout the thesis, I mentioned dualism at a few points, usually in a negative light, as a view that struggles with the causal closure of the physical. This characterization does not do justice to dualism. Despite its problems with causal closure, dualism is a valuable theory of consciousness. Historically, dualism was the major competitor to physicalism, and as Papineau (2001) points out, physicalism has taken the lead only recently.[102] In light of this, I will now give dualism its fair share and examine it as a serious theory of consciousness.

Dualism is the metaphysical thesis that both the physical entities and minds are Real. Dualists typically *agree* that both physical entities and minds are fundamental. However, dualists *disagree* on many points, most notably with regards to the causal efficacy of minds. *Interactionist dualists* say that minds are both Real and causally effective. *Epiphenomenal dualists* say that minds are Real; however, they are not causally effective.[103] Interactionist dualism can be further divided into naturalistic dualism and supernatural dualism. *Naturalistic dualists* think that mental events can be integrated within physics. *Supernatural dualists* think that mental events cannot be integrated within physics. (see Table 6)

[102] This should not be surprising given dualism intuitive appeal. Anthropological research indicates that we are all born as 'natural dualists' (Astuti, 2001; Gell, 1998: 127). Similarly, the psychologist Paul Bloom (2004) has argued that the dualist thinking is present even in early childhood.
[103] At least not in relation to any physical entities. Epiphenomenalism might leave open the possibility for causality between the mental entities.

<table>
<tr><td colspan="3">Are minds causally effective?</td></tr>
<tr><td colspan="2">Yes: Interactionist Dualism</td><td rowspan="3">No: Epiphenomenal Dualism</td></tr>
<tr><td colspan="2">Are mental events integrated within physics?</td></tr>
<tr><td>Yes: Naturalistic Dualism</td><td>No: Supernatural Dualism</td></tr>
</table>

Table 6: Versions of dualism.

Naturalistic dualism is the thesis that (1) there are Real physical entities and Real minds, (2) Real minds are causally effective, and (3) mental events can be integrated within physics. In this view, minds are seen as *Real natural entities*, in the same way that spacetime might be a Real natural entity. According to naturalistic dualists, minds causally interact with the rest of the physical world in a regular and predictable manner. Thus, mental events should be accountable by natural science. Thus, we could expect future laws of physics predicting mental to physical interactions in the same way that current laws of physics are predicting purely physical interactions. Nevertheless, unlike the rest of nature, minds instantiate experiences. This makes minds dissimilar from the rest of nature. In the literature, versions of naturalistic dualism have been explored by Chalmers (1996) and Martine Nida-Rümelin (2007).

Supernatural dualism is the thesis that (1) there are Real physical entities and Real minds, (2) Real minds are causally effective, and (3) mental events cannot be integrated within physics. In this view, minds are seen as *Real supernatural entities* that causally interact with the physical world. Minds causally affect physicals systems; however, these physical effects have no physical causes. The physical effects of the mental causes are not regular and predictable; thus, mental events cannot be integrated within physics. Supernatural dualism entails that the effects of the mind on the physical world resemble *miracles*. Historically, supernatural dualism has been famously defended by Descartes, while more recently, its defenders include Richard Swinburne (1986) and W.D. Hart (1988).

Epiphenomenal dualism is the thesis that (1) there are Real physical entities and Real minds, and (2) the Real minds *are not* causally effective. Thus, epiphenomenal

dualism stands in contrast to naturalistic and substance dualism (as varieties of interactionist dualism). Since epiphenomenal minds lack physical causal efficacy, there can never be any mental to physical causal interactions. Thus, by all accounts, if epiphenomenal dualism is true, mental events could never be an integral part of physics. Versions of epiphenomenal dualism have been defended by Chalmers (1996) and Jackson (1982)

In the literature, one often finds accusations that interactionist dualism is empirically unfounded. One worry is that there is no empirical evidence in support of interactionist dualism. Another, that I already brought up at multiple places, is that interactionist dualism breaks the causal closure of the physical. Here, I will put these worries aside. This is because I want to assess a *best-case scenario* version of dualism against the explanatory standards I have set. This should not be taken to mean that I assume a scenario in which dualism has been empirically proven. If this were the case, there would be nothing further to discuss here: it would be evident that dualism wins. Instead, I only assume that the lack of evidence or the potential inconsistency with casual closure is insufficient to disprove dualism as a theory of consciousness.

Dualists neither can nor want to close Consciousness Gap. Instead, dualists fully embrace Consciousness Gap by saying that minds are Real. Moreover, dualism does not face the combination problem nor the missing objects problem. Dualists do not think that minds ground other minds nor the physical world.

Now, onto problems. All versions of dualism posit *two explanatory primitives*:

1) *The existence thesis*: both minds and physical entities are Real.
2) *The presence thesis*: minds are present only in some physical systems.

I will argue that these two explanatory primitives put dualism at a disadvantage compared to monist views such as physicalism and panpsychism.

First, dualism is *less simple* than monism. This is entailed by the existence thesis. All metaphysical views take some existence claims as explanatory primitive. Dualists think there are Real minds and Real physical entities. In contrast, monists, such as physicalists or panpsychists, think of all entities as belonging to one kind: either Real

physical entities or Real minds. Thus, the simplicity in question here does not involve the number of Real entities but instead the *number of kinds* of Real entities.

A monist might say there are countless entities of the same kind. In contrast, a dualist might say that there are only one Real mind and one Real physical entity. Nevertheless, because the dualist posits different kinds of entities, even this barren version of dualism is less simple than the most entity-rich versions of monism.

The reader might wonder: why having more kinds of entities is worse than having more entities of the same kind? Why would having one elephant and one giraffe be worse (in terms of simplicity) than having a thousand giraffes? The answer to this, in my opinion, is that more kinds entail more different entities with radically different properties. Thus, the more kinds there are, there will be more properties that all individuals do not share. Thus, the more kinds there are, the more divided the cosmos becomes in terms of properties. If this is so, dualism is always at a disadvantage in terms of metaphysical parsimony compared to monism.

Second, dualism is *less explanatory* than reductive monism. This problem is entailed by the presence thesis. On the dualist picture, by definition, there can never be a reductive explanation of the presence of minds in the physical world. Dualists cannot *explain why* minds causally interact with only *some* physical systems, such as the brains of living humans. Despite the added metaphysical complexity, both consciousness itself and its relation to the physical entities are left mysterious on the dualist picture. In this regard, dualism seems to be at an explanatory disadvantage compared to both physicalism and panpsychism.

With the above in mind, which version of dualism offers the most promising solution to the problem of consciousness? First, in terms of metaphysics, interactionist dualism seems to have an advantage over epiphenomenal dualism. All versions of dualism postulate Real minds; however, not all Real minds are made equal. Real epiphenomenal entities, in general, seem metaphysically bizarre. It is quite controversial to say that something without physical causal influence Really exists.[104]

[104] For more on similar worries against epiphenomenalism, see chapter 1.

Second, in terms of explanation, naturalistic dualism seems to have an advantage over the competition. Naturalist dualism is the only dualist view that allows for new laws of physics describing mental events. Thus, naturalistic dualism seems to promise the most in terms of explaining consciousness.

In summary, I prefer naturalist dualism over both supernatural and epiphenomenal dualism, while I prefer supernatural dualism over epiphenomenal dualism.

5.2.2. Emergentism

Emergentism about minds is the thesis that (1) there are Real physical entities and Real minds, and (2) the Real minds are *non-reductively grounded* in the Real physical entities. Moreover, Real minds *are* causally effective, and mental events *can be* integrated into physics. So construed, emergentism about minds is the same thesis as naturalistic dualism. In the previous section, I argued in favor of naturalistic dualism as my preferred dualist solution to the problem of consciousness. Everything I said about naturalistic dualism applies to emergentism about minds. Thus, instead of repeating myself, in this section, I want to examine a different question. Namely, how does emergentism about minds (i.e., naturalistic dualism) relate to physicalism? I find this to be an interesting question, and answering it will help us illuminate the relation between physicalism, emergentism, and dualism.

Emergentists about chemical compounds or living organisms posit a picture where *Real physical entities* are grounded in the Real fundamental physical entities. However, when it comes to emergentism about minds, emergent minds are Real yet are *non-physical*. Thus it seems, not all versions of emergentism are made equal: there is emergentism about Real physical entities and emergentism about Real non-physical entities.

> *Physical Emergentism*: the metaphysical thesis that all Real emergent entities are physical.

> *Non-physical Emergentism*: the metaphysical thesis that some Real emergent entities are non-physical.

Tim Crane (2010) has argued that *all* versions of emergentism should be distinguished from physicalism. Thus, Crane makes a sharp distinction between physicalism and all versions of emergentism. According to him, for a thesis to count as a version of physicalism, it must be either ontologically reductive or explanatory reductive (or both). He defines *ontological reduction* as the thesis that: "All entities (objects, properties, relations, facts, etc.) belong to a subclass of the class of physical entities." (2010: 30) On the other hand, *explanatory reduction* is the thesis that "All truths (particular truths, or general theoretical truths or laws) can be explained in principle in terms of broadly physical truths." (2010: 30) I agree with Crane that all versions of physicalism must fulfill at least one of these reductionist theses. Moreover, I agree with Crane that all versions of emergentism fail the explanatory reduction thesis. However, I disagree with Crane that *all* versions of emergentism fail the ontological reduction thesis.

Crane's (2010) view is that no version of emergentism can be physicalism, and vice versa, no version of physicalism can be emergentism. If so, according to Crane, physicalism, in all its versions, is incompatible with the existence of Real dependent entities (i.e., Real groundees). I will refer to this thesis as follows:

> *Strong Physicalism*: the metaphysical thesis that all Real entities are physical and fundamental.

Strong Physicalism entails that both Physical Emergentism and Non-physical Emergentism *are not* versions of physicalism. Strong Physicalism *fulfills* Crane's requirement for ontological reduction. If Strong Physicalism is true, there are no Real dependent entities. Instead, all Real entities are physical, and all Nominal entities reduce to the Real physical entities.

However, Strong Physicalism is not the only way to construe physicalism. Many emergentists would be dissatisfied to have their view classified as non-physicalism.

Emergentists about chemistry or biology might say that despite being emergent, chemical compounds and living organisms are nevertheless physical, and thus their view should count as a version of physicalism. Moreover, many physicalists who embrace multiple realizability of the dependent entities (such as role functionalists) would be happy with an alternative, weaker characterization of physicalism that only focuses on the physicality of the fundamental level.

> *Weak Physicalism*: the metaphysical thesis that all fundamental Real entities are physical.

Weak Physicalism entails that both Physical Emergentism and Non-physical Emergentism *are* versions of physicalism. Weak Physicalism *fails* Crane's requirement for ontological reduction. If Weak Physicalism is true, all Real fundamental entities would be physical; however, if there are any Real dependent entities, these need not be physical. Weak Physicalism allows for Real non-physical entities.

The problem with Weak Physicalism is that it might be too permissive. It allows for Physical Emergentism to count as physicalism; however, it moreover allows for Non-physical Emergentism to count as physicalism. Perhaps, some naturalistic dualists would be happy to be counted as physicalists; however, Non-physical Emergentism might include many other types of exotic Real entities. For example, views on which there are Real emergent universals, numbers, value, composites, or even—angels or ghosts—would count as physicalist according to Weak Physicalism. This is problematic since, plausibly, we would like to differentiate between views that exclude such entities and ones that do not. Allowing such views to count as physicalist is at best counter-intuitive, and at worst, trivializes the notion of physicalism.

Considering the problems of both Strong Physicalism and Weak Physicalism, I suggest a third middle way, which I believe avoids these concerns.

> *Moderate Physicalism*: the metaphysical thesis that all Real entities are physical.

Moderate Physicalism entails that Physical Emergentism *is* physicalism, but Non-physical Emergentism *is not* physicalism. Moderate Physicalism *fulfills* Crane's requirement for ontological reduction. If moderate physicalism is true, all Real fundamental entities would be physical, and there would be Real physical dependent entities. Thus, all irreducible entities would be physical, and all that reduces would reduce to the Real physical entities. Moderate Physicalism, unlike Strong Physicalism, allows for Physical Emergentism to count as physicalism. Moderate Physicalism, unlike Weak Physicalism, denies that Non-physical Emergentism can count as physicalism. I believe these considerations give Moderate Physicalism an advantage over Strong Physicalism and Weak Physicalism as the best way to construe physicalism.

In summary, it seems that whether or not emergentism is physicalism depends on how strongly or weakly we define physicalism. I believe this is not purely a matter of terminology. Real physical dependent entities (such as Real chemical compounds) are of *the same kind* as the Real fundamental physical entities; however, they are of *a different kind* than any Real non-physical dependent entities (such as Real minds or Real angels). In light of this, I offered a definition of physicalism that includes physical varieties of emergentism yet excludes non-physical varieties of emergentism. Thus, in my characterization, emergentism about minds *is not* physicalism, while emergentism about chemical compounds *is* physicalism. Thus, despite the integration of minds within nature, naturalistic dualism *is not* physicalism.

5.2.3. Extended Physicalism

Extended (or structure-plus) physicalism is the thesis that (1) all the fundamental entities are Real physical entities, and (2) the fundamental physical entities are *structure-plus physical entities*. Structure-plus physical entities are *not purely structural* (like the entities of pure physicalism). Instead, all of their extrinsic properties are grounded in their *intrinsic physical natures*. Thus, so construed, structure-plus physicalism is a version of Russellian monism that posits *physical inscrutables*.

Versions of structure-plus physicalism have been defended by Stoljar (2001), Montero (2010, 2015), and Pereboom (2011, 2015), among others. Stoljar describes his view as 'object-based' physicalism or 'o-physicalism,' while Montero describes it as 'Russellian physicalism.'

The specific version of extended physicalism I will be examining is *reductive Russellian physicalism*. In this view, there are no Real dependent entities, and minds are not Real. Instead, minds are reductively grounded in the Real structure-plus physical entities. Thus, all Real entities are structure-plus physical entities. This is a type-A view on which the truths of the physical inscrutables a priori closes Consciousness Gap.

Extended physicalism is a very attractive view. It promises to do better than both pure physicalism and panpsychism. First, extended physicalism is 'tailor-made' to fix the failings of pure physicalism. It aspires to provide a metaphysically and explanatory parsimonious monist framework, on which everything Real is physical, and Consciousness Gap is closed. Second, extended physicalism has an advantage over panpsychism since it appears to avoid both the combination and the missing objects problem. It avoids the combination problem since physical inscrutables do not engage in 'mental combination.' It avoids the missing objects problem since physical inscrutables are designed to reductively ground any physical structure.

A question that naturally arises for the extended physicalist is: 'what are the physical inscrutables?' This question is a serious challenge for the extended physicalist. In the case of pure physicalism, it was straightforward to define an entity as physical. However, by definition, when it comes to inscrutables, we are dealing with entities beyond the vocabulary and methods of physics (at least, of physics as currently conceived). In light of this, what makes an inscrutable physical?

One way to define any inscrutable is by employing the *via negativa*. We could stipulate that the physical inscrutables have no properties that involve value, experiences, divinity, or any supernatural aspects. Additionally, we could define any inscrutable by stating what it is *apt to ground*.

Montero (2010: 77, 2015: 216) argues that what makes the physical inscrutables unique is that they can explain *everything physical* (including experiences). In Montero's own words:

> If inscrutables are in this way the substance of the world—if they are, to use Stephen Hawking's words, what 'breathes fire into the equations [of any possible grand unified theory of physics] and makes a universe for them to describe'—they are not uniquely important to the mental and so a world with them should be perfectly acceptable to a physicalist. (2015: 217)

With the above in mind, we can now get some grasp on the physical inscrutables by specifying some of their properties and roles:

1) *Physical*: The physical inscrutables must be physical.
2) *Reductively ground everything physical*: The physical inscrutables must be apt to reductively ground everything physical, including experiences.
3) *Close Consciousness Gap:* The facts of the physical inscrutables (as grounds) a priori entail the experiential facts (as groundees).

Arguably, extended physicalism is an ideal version of physicalism, the best that a physicalist could hope for. It avoids the problems of both pure physicalism and pure panpsychism. It promises to close all explanatory gaps and does all of this in virtue of the Real physical entities.

I worry that extended physicalism might be *too good to be true*. Physical inscrutables sound a bit too much like a *deus ex machina*. They are mysterious, they are beyond everyone's grasp, and they almost miraculously solve all the problems of pure physicalism.

The physical inscrutables are posited primarily to ground extrinsic physical properties. The promise of closing Consciousness Gap is a fortunate side-effect. And this side-effect is doubtful since physical inscrutables are not experiential. If pure structure and dynamics failed to close Consciousness Gap, how can the grounds of pure structure

and dynamics do so? Thus, I believe that the extended physicalist owns us *a further justification* of how exactly the physical inscrutables close Consciousness Gap.

5.2.4. Panprotopsychism

Panprotopsychism is the thesis that all the fundamental entities are Real *proto-conscious entities*. Proto-conscious entities have extrinsic physical properties (like the entities of pure physicalism); however, these are grounded in their *intrinsic proto-conscious natures*. Thus, so construed, panprotopsychism is a version of Russellian monism that posits *proto-conscious inscrutables*.

Proto-conscious inscrutables are not experiential yet are the grounds of experiences in dependent entities such as humans. In Chalmers' words: "they have certain special properties that are precursors to consciousness and that can collectively constitute consciousness in larger systems." (2015: 248) To use a metaphor, proto-conscious inscrutables are *the seeds of future experiences* in dependent entities.

The specific version of panprotopsychism I will examine is *reductive Russellian panprotopsychism*. In this view, experiences are reductively grounded in the Real proto-conscious entities. Thus, all Real entities are proto-conscious entities. Thus, minds are not Real. Moreover, this is a type-A view and promises that the proto-conscious truths a priori close Consciousness Gap.

As was the case with extended physicalism, panprotopsychism seems motivated by the failings of pure physicalism and panpsychism. Moreover, just like extended physicalism, panprotopsychism promises a monistic solution to the problem of consciousness. Panprotopsychism promises to close Consciousness Gap. Moreover, panprotopsychism might avoid both the combination and the missing objects problem. It avoids the combination problem since the proto-conscious inscrutables are designed to ground experiences. It might avoid the missing objects problem since the proto-conscious inscrutables may have properties that can reductively ground any physical structure.

Extended physicalism and panprotopsychism appear to be very similar. It is far from clear what, if anything, makes them distinct. To makes matters worse, the two views are often conflated in the literature, and it is difficult to tell them apart.[105]

The panprotopsychist must answer: 'what are the proto-conscious inscrutables?' One obvious answer to this question is that proto-conscious inscrutables are *not physical*. If the inscrutables *are physical*, we are dealing with extended physicalism. Otherwise, if the inscrutables *are not physical*, we are dealing with panprotopsychism.

Moreover, as Montero (2010: 77, 2015: 216) argues, proto-conscious inscrutables are posited with the sole purpose of metaphysically explaining experiences. Similarly, Goff et al. (2020) say that "protophenomenal properties have a kind of indirect definition, in terms of their propensity to ground consciousness." Based on these insights, if the inscrutables *only* reductively ground experiences, then the view is not physicalist but panprotopsychist. Otherwise, if the inscrutables *do not only* reductively ground experiences, but instead ground everything physical (including experiences), then the view is physicalist.

We can now finally shed light on the nature of the proto-conscious inscrutables. Their properties and roles include:

1) *Neither physical nor experiential*: Proto-conscious inscrutables are neither physical nor experiential.
2) *Reductively ground experiences*: Protophenomenal entities either fully or partially reductively ground dependent experiences.
3) *Close Consciousness Gap*: The facts of the proto-conscious inscrutables (as grounds) a priori entail the experiential facts (as groundees).

Table 7 below illustrates the contrast between extended physicalism and panprotopsychism.

[105] For example, in the literature, Goff et al. (2020) list McGinn (1989), Stoljar (2001), Holman (2008), Montero (2010), McClelland (2013), Pereboom (2011, 2015), and Coleman (2012, 2014, 2015, 2016) as panprotopsychists. However, as I have already stated in the previous section, from these authors, at least Stoljar, Montero, and Pereboom would describe their views as versions of physicalism.

	Extended physicalism	**Panprotopsychism**
Inscrutables	Physical inscrutables.	Proto-conscious inscrutables.
A priori entailment	The facts of the physical inscrutables (as grounds) a priori entail the experiential facts (as groundees).	The facts of the proto-conscious inscrutables (as grounds) a priori entail the experiential facts (as groundees).
Grounding	The inscrutables reductively ground everything physical (including experiences).	The inscrutables reductively ground experiences.

Table 7: Extended physicalism and panprotopsychism.

Is panprotopsychism, so construed, a sensible view? Extended physicalism faced the problem of being potentially unable to reductively ground experiences. In contrast, panprotopsychism must answer both (1) how the proto-conscious inscrutables could ground experiences, as well as the further question of (2) what grounds the extrinsic properties of physics.

Regarding the problem of metaphysically explaining experiences, as Chalmers points out: "panprotopsychism faces a version of the combination problem that does not arise for panpsychism: How can nonexperiences constitute experiences?" (2016: 184). Nevertheless, on my definition here, panprotopsychism seems to have an advantage over extended physicalism, at least in this regard. The proto-conscious inscrutables, unlike the physical inscrutables, are 'tailor-made' to explain experiences. Still, I agree with Chalmers that the panprotopsychist owns us as a more detailed explanation about the exact details of this process.

Panprotopsychism is at a disadvantage compared to extended physicalism when it comes to solving the missing objects problem. The proto-conscious inscrutables are posited specifically to ground experiences. However, if this is the case, then what grounds the physical entities? The panprotopsychist must answer this question *while keeping her view distinct* from extended physicalism. Here, I leave it an open question whether or not the panprotopsychist can do this successfully.

5.2.5. Neutral Monism

Neutral monism is the thesis that all the fundamental entities are Real *neutral entities*. Neutral entities are intrinsically *neither experiential nor physical*.[106] Thus, so construed, neutral monism is a version of Russellian monism that posits *neutral inscrutables*. The neutral inscrutables ground both the extrinsic properties of physics and experiences. Thus, in this view, both the physical entities and experiences are grounded in the Real neutral entities.

Neutral monism is a thesis with a long and venerable history. Philosophers such as Spinoza and Hume have been interpreted as neutral monists. Arguably, many ancient philosophers such as Plato or Plotinus might likewise be considered neutral monists.[107] Contemporary neutral monism is usually traced to some of the works of Ernst Mach, William James, and Bertrand Russell. For example, here is Russell describing his version of the view:

> The stuff of which the world of our experience is composed is, in my belief, neither mind nor matter, but something more primitive than either. Both mind and matter seem to be composite, and the stuff of which they are compounded lies in a sense between the two, in a sense above them both, like a common ancestor. (1921: 10–1)

In recent years, Chalmers (1996: 129) has been sympathetic to the idea of neutral monism.

The specific version of neutral monism I will examine here is *reductive Russellian neutral monism*. Briefly, this is the thesis that the fundamental neutral entities reductively ground both minds and matter. Thus, on this view, there are Real neutral

[106] Another way to construe the neutral monist view is to say that fundamental reality is *both* experiential and physical (as opposed to neither). However, as Stubenberg (2018) highlights, upon inspection, this "both" interpretation of neutral monism appears to collapse into dualism, or either physical or mental monism. Thus, since the 'both' view seems unsustainable, I will not consider it here.

[107] Platonists seem to posit that Reality is neither physical nor experiential (in the sense of pure experience). Plausibly, Plato postulates a duality between the divine realm of the Forms (as fundamental) and the realm of appearances (as dependent), but not between experiences and physical entities.

entities, but there are neither physical entities nor Real experiences. Moreover, this is a type-A view, on which the truths of the neutral inscrutables a priori entail all other truths, and thus a priori close Consciousness Gap.

Like extended physicalism and panprotopsychism, it seems neutral monism has an advantage over both pure physicalism and pure panpsychism. Its advantage over pure physicalism comes from its promise to close Consciousness Gap. Its advantage over pure panpsychism comes from its promise to avoid the combination and the missing objects problems. Like extended physicalism and panprotopsychism, neutral monism offers a metaphysically and explanatory parsimonious monist framework as a solution to the problem of consciousness.

Neutral monists must answer: 'what are the neutral inscrutables?' It is not prima facie clear what makes the neutral inscrutables distinct from the physical inscrutables and the proto-conscious inscrutables. In what follows, I will attempt to give neutral monism a distinct conceptual space alongside extended physicalism and panprotopsychism.

First, what makes neutral monism different from panprotopsychism? Both the neutral inscrutables and the proto-conscious inscrutables are neither physical nor experiential. If so, are panprotopsychism and neutral monism the same view?[108] I think there is a key difference between the two views. The proto-conscious inscrutables are apt to ground *only experiences*, while the neutral inscrutables are apt to ground *everything*. Thus, even if we classify panprotopsychism as a version of neutral monism, it would still be a *unique version* of neutral monism. Neutral inscrutables are apt to ground experiences *and more*. Thus, neutral inscrutables could have properties that go beyond the properties of proto-conscious inscrutables.

Second, what makes neutral monism distinct from extended physicalism? By definition, neutral inscrutables are not physical. However, upon reflection, it seems that both the physical and the neutral inscrutables reductively ground both experiences and physical entities. Thus, it is not enough for the neutral monist to

[108] Kind (2015: 411–2) argues that protophenomenal monism might collapse into a type of neutral monism based on similar reasons.

stipulate that her inscrutables are not physical. Kind (2015: 413) embraces this line of thought and uses it to argue that neutral monism collapses into extended physicalism.[109] In response, the neutral monist must say what exactly makes her inscrutables non-physical.

I think that the best way for the neutral monist to motivate her view is by defining the neutral inscrutables as apt to reductively ground *more* than only experiences and physical entities. Defined as such, the neutral inscrutables might be apt to reductively ground dependent entities that are neither experiential nor physical. What might be some examples of such neutral entities (or their properties)? Value, mathematical and logical entities, universals, and free will are just a few examples that immediately come to mind. All these entities (or properties) have been problematic for physicalists and are plausibly likewise problematic for all the views I am considering here. The neutral inscrutables, when defined in this way, are the *ultimate inscrutables*. They are apt to ground all the entities that the extended physicalist, the panprotopsychist, and all the other theories might have trouble grounding.

In light of the above, based on their metaphysical and epistemic roles, neutral inscrutables can be characterized as:

1. *Neither physical nor experiential*: The neutral inscrutables are neither physical nor experiential.
2. *Reductively ground everything*: Neutral inscrutables reductively ground everything (experiences, physical entities, and everything else there might be).
3. *Close Consciousness Gap:* The facts of the neutral inscrutables (as grounds) a priori entail all other facts (as groundees).

Table 8 below illustrates the contrasts between extended physicalism, panprotopsychism, and neutral monism.

[109] In her terminology 'Russellian physical monism.'

	Extended physicalism	**Panprotopsychism**	**Neutral monism**
Inscrutables	Physical inscrutables.	Proto-conscious inscrutables.	Neutral inscrutables.
A priori entailment	The facts of the physical inscrutables (as grounds) a priori entail the experiential facts (as groundees).	The facts of the proto-conscious inscrutables (as grounds) a priori entail the experiential facts (as groundees).	The facts of the neutral inscrutables (as grounds) a priori entail all other facts (as groundees).
Grounding	The inscrutables reductively ground everything physical (including experiences).	The inscrutables reductively ground experiences.	The inscrutables reductively ground everything.

Table 8: Extended physicalism, panprotopsychism, and neutral monism.

Defined as above, neutral monism is an improvement over both panprotopsychism and extended physicalism. Panprotopsychists posit inscrutables apt to ground primarily experiences. Extended physicalists posit inscrutables apt to ground both experiences and physical entities. However, neutral monists posit inscrutables apt to ground experiences, physical entities, and any other entities that might exist yet are neither physical nor experiential.

I have given only a rough approximation of neutral monist metaphysics. Many more details need to be fleshed out before we can have a usable theory. For one, the neutral monism, as I have defined it, is certainly open to the 'deus ex machina' objection I raised against extended physicalism. Nevertheless, I think this approximation is enough to characterize natural monism as a unique thesis among the other candidate theories of consciousness.

5.2.6. Impure Panpsychism

Impure panpsychism is the thesis that all the fundamental Real entities are intrinsically *both experiential and neutral*. A recent and well-articulated example of this thesis, on which I will focus here, is Goff's (2017) 'consciousness+' panpsychism.

So construed, impure panpsychism is a version of Russellian monism that posits *conscious+ inscrutables*.

Fully spelled out, the view I will be focusing on is *reductive Russellian consciousness+ panpsychism*. I will understand this as the view that intrinsic natures of the Real fundamental entities are conscious+ properties, and they reductively ground everything. Thus, in this view, dependent experiences are not Real. Moreover, this is a type-A view, on which the truths of the Real consciousness+ properties a priori entail all other truths and thus close Consciousness Gap. This view is not an exact portrayal of Goff's views but is best seen as a model for any version of impure panpsychism, inspired by Goff's ideas.

Impure panpsychism inherits all the virtues of standard panpsychism and strives to fix its weakness. Primarily, it does this by promising to avoid the combination problem. Perhaps, if pure experiences cannot combine into other experiences, conscious+ properties can combine. Moreover, impure panpsychism seems to avoid the missing objects problem. Perhaps, if pure experiences cannot ground all physical entitles, conscious+ properties can (plausibly, in virtue of their neutral aspects). Concerning the other Russellian views, it is clear that impure panpsychism has a unique conceptual space. Impure panpsychism is not Russellian physicalism since its inscrutables are not physical. Impure panpsychism is neither panprotopsychism nor neutral monism since its inscrutables are part experiential. In light of all this, impure panpsychism seems like an especially promising solution to the problem of consciousness.

The impure panpsychist must answer: 'what are the conscious+ inscrutables?' According to Goff, the conscious+ properties are "enfolding experiential and non-experiential aspects into a single nature." (2017: 230) Thus, the conscious+ properties are part experiential and part non-experiential. But then, what exactly are the non-experiential (neutral) aspects of the conscious+ properties? In other words, what exactly is the 'plus' in 'consciousness+'?

In Goff's view (2017: 179), the plus aspect of consciousness+ properties is entirely unknown. Nevertheless, we can state a few characteristics of the plus aspect. First, the

plus aspect is *not experiential*. Second, as defined by Goff, the plus aspect is posited primarily *to solve the combination problem*. In his own words: "the unknown part of consciousness+ is crucial to phenomenal combination, and our inability to make sense of phenomenal combination is the result of our ignorance of this aspect." (2017: 180) Based on this, impure panpsychism can be seen as a middle path between pure panpsychism and full-blown panprotopsychism.

Impure panpsychism might seem like an *ad hoc* attempt to save panpsychism. After all, the plus aspect is added primarily to solve the combination problem. In response, the impure panpsychist could say that her view is not ad hoc but instead is motivated by the theoretical benefits of positing fundamental experiences. The impure panpsychist might say that positing fundamental experiences is the *only way* to close Consciousness Gap from the metaphysical ground up. Thus, if fundamental experiences are necessary to close Consciousness Gap and pure panpsychism is false, impure panpsychism is not an ad hoc view. Instead, it is the most viable alternative to pure panpsychism.

In conclusion, the question of whether impure panpsychism is well-motivated hangs on our theoretical stance towards the explanatory value of fundamental experiences. If we think that Consciousness Gap can only be closed by positing fundamental experiences, impure panpsychism makes perfect sense. If, however, we think that experiential inscrutables are not necessary to close Consciousness Gap, then impure panpsychism might seem undermotivated.

5.3. Conclusion

I considered *seven alternatives* to pure physicalism and panpsychism. From these, three are versions of dualism: naturalist dualism, supernatural dualism, and epiphenomenal dualism. The other four are versions of Russellian monism: extended physicalism, panprotopsychism, neutral monism, and impure panpsychism. All the Russellian views I explored posit different species of inscrutables; nevertheless, they all posit that human experiences are reductively grounded and are type-A. In the case of all seven views, I explored idealized versions of the theories, and I assumed that they could avoid some of the challenges posited against them. My main interest is in the theoretical promise of these theories, assuming that all other things are equal.

In what follows, I offer a ranking of these alternatives and the two views I already explored in the thesis: pure physicalism and pure panpsychism. This ranking is not conclusive. First, because it is primarily my own ranking, based on my own standards. I hope to have given good reasons to justify these standards, but ultimately, they can be questioned. Second, because I did not solve the problem of consciousness, but I only hint at where the solution *may* be found.

Nevertheless, the ranking is not arbitrary. It is based on the theoretical virtues of the different theories as potential solutions to the problem of consciousness. Many of the views I ranked highly are currently significantly under-represented and under-researched in the literature. Thus, this dissertation should be seen as a *call to action* for a future shift in research interests towards these (currently) alternative views.

Some readers might wonder if the inconclusive nature of this examination and the persistence of the problem of consciousness is not indicative of the need for a radical change. Perhaps, this project—and any similar attempt to solve the problem of consciousness—are doomed to fail since they employ flawed concepts. These readers might believe that the very notions of ‘experiences’ and ‘physical entities’ are deeply flawed and should be abandoned. Perhaps, these readers might think that we are dealing with a false dichotomy, that we are trapped in the prison of our culture and history, and that our views have been corrupted by Descartes and by reading too much

philosophy. I do not have much to say to these critics except to invite them *to do better*. So far, I have not seen a satisfactory execution of their approach. At the bottom line, despite the many critiques, the mind-body distinction does seem to capture *something* about reality. The ongoing debates in the philosophy of mind show that it is impossible to simply theoreticize the mind-body dichotomy away.

My *final ranking* of the theories of consciousness is as follows:

1) Neutral monism
2) Extended physicalism
3) Panprotopsychism
4) Impure panpsychism
5) Naturalistic dualism
6) Supernatural dualism
7) Epiphenomenal dualism
8) Pure panpsychism
9) Pure physicalism

Neutral monism wins because it promises the most in the most coherent and parsimonious way. In general, I gave the Russellian monist views an advantage over the dualistic alternatives. The Russellian views are all monistic and thus postulate only one kind of Real entity. Moreover, they employ reductive grounding, and thus, at least in principle, promise to close all explanatory gaps.

From the Russellian views, I favored neutral monism since it promises to solve problems involving the grounding of non-physical entities beyond just consciousness. Second, I favored Russellian physicalism over panprotopsychism. This is because the physical inscrutables seem to have a more ubiquitous grounding role than the protophenomenal inscrutables. Finally, I ranked impure panpsychism the lowest from the Russellian views. This is because impure panpsychism posits dual-aspect inscrutables; thus, it appears metaphysically less simple than the other Russellian alternatives.

I want to stress that my classification of the Russellian monist views is flexible. For example, if I did not think that value or free will need a metaphysical explanation, I

would not have prioritized neutral monism. Moreover, impure panpsychism could have *easily* ranked higher. Only impure panpsychism, from all the Russellian candidates, posits fundamental experiences. It is highly plausible that any dependent experiences could only be grounded in more fundamental experiences. Here, I remained neutral on this important issue. However, were I to accept that thesis, impure panpsychism would have ranked higher and possibly would have been the overall winner.

I gave dualism an advantage over pure physicalism and pure panpsychism. Dualism is not threatened by Consciousnesses Gap, the combination problem, nor the missing objects problem. I ranked naturalistic dualism the highest among the different dualist views, and I ranked supernatural dualism over epiphenomenal dualism. This is because naturalistic dualists posit that minds are integrated within the physical world. Epiphenomenal dualism, on the other hand, is ranked the lowest since it posits Real entities whose existence is highly controversial.

I stayed neutral on the issue of whether or not experiences are Real. This approach entailed that reductive views would rise higher up the list, while emergentist views would fall down the list. Had I not done this, views that posit Real experiences (regardless of whether they are moreover fundamental) would have ranked higher. Thus, dualism would have ranked higher than the Russellian views (except maybe impure panpsychism). However, on my current standards, emergentism (and dualism in general) naturally falls out.

Finally, I gave the advantage to pure panpsychism over pure physicalism. Both of these views fail to intelligibly explain human experiences. Nevertheless, pure panpsychism dares to posit Real experiences (at the fundamental level). In contrast, pure physicalism cannot move past the ambition of grounding experiences in a pure causal structure. As I have argued and as many philosophers have argued before me, this ambition seems to lead to a dead-end. The panpsychist, on the other hand, at least move the debate forward. Thus, I give pure physicalism the last place.

My final ranking does not solve the problem of consciousness. Nevertheless, I hope it shows where to look for the answer. Ultimately, this dissertation, as a whole, should

be seen as a push away from pure physicalism. Contemporary philosophy of mind has largely anchored itself on the shore of physical science. This shore is an indispensable haven, and we should never downgrade its importance. However, having found a safe port does not entail we should not sail the ocean. There are a plethora of wonderous lands beyond the horizon waiting to be fully explored. We must dare to view consciousness as a potentially Real phenomenon. More generally, beyond just consciousness, we should dare to accept the possibility that the cosmos is more than a meaningless structure. This thesis is *an invitation* for such an exploration.

REFERENCES

Adams, Robert M. (2007). Idealism Vindicated. In Peter van Inwagen & Dean Zimmerman (Eds.), *Persons: Human and Divine* (35–54). Oxford University Press.

Albahari, Miri. (2020). Beyond Cosmopsychism and the Great I Am: How the World Might Be Grounded in Universal 'Advaitic' Consciousness. In William Seager (Ed.), *The Routledge Handbook of Panpsychism* (119–130). Routledge.

Albert, David Z. (1996). Elementary Quantum Metaphysics. In J. Cushing, Arthur Fine, & Sheldon Goldstein (Eds.), *Bohmian Mechanics and Quantum theory: An Appraisal* (277–284). Kluwer.

Albert, David Z. (2015). *After Physics*. Harvard University Press.

Albert, David Z. (2019). Preliminary Considerations on the Emergence of Space and Time. In Alberto Cordero (Ed.), *Philosophers Look at Quantum Mechanics* (87–95). Springer.

Alter, Torin, & Nagasawa, Yujin (Eds.). (2015). *Consciousness in the Physical World: Perspectives on Russellian Monism*. Oxford University Press.

Armstrong, David Malet. (1968). *A Materialist Theory of the Mind*. Routledge and Kegan Paul.

Armstrong, David Malet. (1979). *A Theory of Universals* (Vol. 2). Cambridge University Press.

Armstrong, David Malet. (1997). *A World of States of Affairs*. Cambridge University Press.

Astuti, Rita. (2001). Are We All Natural Dualists? A Cognitive Developmental Approach. *The Journal of the Royal Anthropological Institute*, *7*(3), 429–447.

Balog, Katalin. (1999). Conceivability, Possibility, and the Mind-Body Problem. *Philosophical Review*, *108*(4), 497–528. https://doi.org/10.2307/2998286

Balog, Katalin. (2012a). Acquaintance and the Mind-Body Problem. In Christopher S. Hill & Simone Gozzano (Eds.), *New Perspectives on Type Identity: The Mental and the Physical* (16–43). Cambridge University Press.

Balog, Katalin. (2012b). In Defense of the Phenomenal Concept Strategy. *Philosophy and Phenomenological Research*, *84*(1), 1–23. https://doi.org/10.1111/j.1933-1592.2011.00541.x

Barnes, Elizabeth. (2012). Emergence and Fundamentality. *Mind*, *121*(484), 873–901. https://doi.org/10.1093/mind/fzt001

Bell, John Stewart. (1987). The Theory of Local Beables. In *Speakable and Unspeakable in Quantum Mechanics* (52–62). Cambridge University Press.

Bennett, Karen. (2011). By Our Bootstraps. *Philosophical Perspectives*, *25*(1), 27–41. https://doi.org/10.1111/j.1520-8583.2011.00207.x

Blackburn, Simon. (1990). Filling in Space. *Analysis*, *50*(2), 62–65. https://doi.org/10.1093/analys/50.2.62

Block, Ned, and Robert Stalnaker. (1999). Conceptual Analysis, Dualism, and the Explanatory Gap. *Philosophical Review*, *108*(1), 1–46. https://doi.org/10.2307/2998259

Bloom, Paul. (2004). *Descartes Baby*. Basic Books.

Blumenfeld, Walter. (1913). Untersuchungen über die scheinbare Grösse im Sehraume. *Zeitschrift Für Psychologie*, *65*, 241–404.

Broad, Charlie Dunbar. (1925). *The Mind and its Place in Nature*. Kegan Paul.

Chalmers, David J. (1995). Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies*, *2*(3), 200–19.

Chalmers, David J. (1996). *The Conscious Mind*. Oxford University Press.

Chalmers, David J. (2002). Does Conceivability Entail Possibility. In Tamar Szabo Gendler & John Hawthorne (Eds.), *Conceivability and Possibility* (145–200). Oxford University Press.

Chalmers, David J. (2003). Consciousness and its Place in Nature. In Stephen P. Stich & Ted A. Warfield (Eds.), *Blackwell Guide to the Philosophy of Mind* (102–42). Blackwell.

Chalmers, David J. (2007). Phenomenal Concepts and the Explanatory Gap. In Torin Alter & Sven Walter (Eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism* (167–94). Oxford University Press.

Chalmers, David J. (2009). Ontological Anti-Realism. In David Manley, David J. Chalmers, & Ryan Wasserman (Eds.), *Metametaphysics: New Essays on the Foundations of Ontology* (77–129). Oxford University Press.

Chalmers, David J. (2010). The Two-Dimensional Argument Against Materialism. In *The Character of Consciousness* (141–206). Oxford University Press.

Chalmers, David J. (2012). *Constructing the World*. Oxford University Press.

Chalmers, David J. (2015). Panpsychism and Panprotopsychism. In Torin Alter & Yujin Nagasawa (Eds.), *Consciousness in the Physical World: Perspectives on Russellian Monism* (246–276). Oxford University Press.

Chalmers, David J. (2016). The Combination Problem for Panpsychism. In Godehard Brüntrup & Ludwig Jaskolla (Eds.), *Panpsychism: Contemporary Perspectives* (179–214). Oxford University Press.

Chalmers, David J. (2020). Idealism and the Mind-Body Problem. In William Seager (Ed.), *The Routledge Handbook of Panpsychism* (353–373). Routledge.

Chalmers, David J. (forthcoming). Finding Space in a Nonspatial World. In Christian Wüthrich, Baptiste Le Bihan, & Nick Huggett (Eds.), *Philosophy Beyond Spacetime*. Oxford University Press.

Chalmers, David J., and Frank Jackson. (2001). Conceptual Analysis and Reductive Explanation. *Philosophical Review*, *110*(3), 315–61. https://doi.org/10.1215/00318108-110-3-315

Chen, Eddy Keming. (2017). Our Fundamental Physical Space: An Essay on the Metaphysics of the Wave Function. *Journal of Philosophy*, *114*(7), 333–365. https://doi.org/10.5840/jphil2017114724

Coleman, Sam. (2012). Mental Chemistry: Combination for Panpsychists. *Dialectica*, *66*(1), 137–166. https://doi.org/10.1111/dltc.2012.66.issue-1

Coleman, Sam. (2014). The Real Combination Problem: Panpsychism, Micro-Subjects and Emergence. *Erkenntnis*, *79*(1), 19–44. https://doi.org/10.1007/s10670-013-9431-x

Coleman, Sam. (2015). Neuro-Cosmology. In Sam Coleman & Paul Coates (Eds.), *Phenomenal Qualities* (66–102). Oxford University Press.

Coleman, Sam. (2016). Panpsychism and Neutral Monism. In Godehard Brüntrup & Ludwig Jaskolla (Eds.), *Panpsychism: Contemporary Perspectives* (249–282). Oxford University Press.

Crane, Tim. (2001). *Elements of Mind: An Introduction to the Philosophy of Mind*. Oxford University Press.

Crane, Tim. (2010). Cosmic Hermeneutics vs. Emergence: The Challenge of the Explanatory Gap. In Cynthia Macdonald & Graham Macdonald (Eds.), *Emergence in Mind* (22–34). Oxford University Press.

Descartes, René. (1996). Meditations on First Philosophy. In John Cottingham (Ed.), *Meditations on First Philosophy with Selections from the Objections and Replies*. Cambridge University Press. (Original work published 1645)

Diaz-Leon, Esa. (2014). Do A Posteriori Physicalists Get Our Phenomenal Concepts Wrong? *Ratio*, *27*(1), 1–16. https://doi.org/10.1111/rati.12018

Diaz-Leon, Esa. (2016). Phenomenal Concepts: Neither Circular nor Opaque. *Philosophical Psychology*, *29*(8), 1186–99. https://doi.org/10.1080/09515089.2016.1244817

Dorr, Cian. (2005). What We Disagree About When We Disagree About Ontology. In Mark Eli Kalderon (Ed.), *Fictionalism in Metaphysics* (234–86). Oxford University Press.

Eddington, Arthur. (1928). *The Nature of the Physical World*. Cambridge University Press.

Einstein, Albert. (1990). Concerning an Heuristic Point of View Toward the Emission and Transformation of Light. In Anna Beck (Trans.), *The Collected Papers of Albert Einstein*: *Vol. Volume 2: The Swiss Years: Writings, 1900-1909* (86–103). Princeton University Press. (Original work published 1905)

Einstein, Albert. (2007). Ether and the Theory of Relativity. In M. Janssen, J. D. Norton, J. Renn, T. Sauer, & J. Stachel (Eds.), *The Genesis of General Relativity*. *Boston Studies in the Philosophy of Science* (Vol. 250, 1537–1542). Springer. (Original work published 1920)

Feigl, Herbert. (1958). The 'Mental' and the 'Physical'. *Minnesota Studies in the Philosophy of Science*, *2*, 370–497.

Fine, Kit. (2001). The Question of Realism. *Philosophers' Imprint*, *1*, 1–30.

Fodor, Jerry Alan. (1974). Special Sciences, or the Disunity of Science as a Working Hypothesis. *Synthese*, *28*(2), 97–115. https://doi.org/10.1007/BF00485230

Galilei, Galileo. (2008). The Assayer. In Maurice A. Finocchiaro (Ed. & Trans.), *The Essential Galileo* (179–189). Hackett Publishing Company. (Original work published 1623)

Gell, Alfred. (1998). *Art and Agency: An Anthropological Theory*. Oxford University Press.

Gillett, Carl, and Barry Loewer. (2001). Preface. In Carl Gillett & Barry Loewer (Eds.), *Physicalism and its Discontents* (ix–x). Cambridge University Press.

Goff, Philip. (2006). Experiences Don't Sum. *Journal of Consciousness Studies*, *13*(10–11), 53–61.

Goff, Philip. (2009). Why Panpsychism Doesn't Help Explain Consciousness. *Dialectica*, *63*(3), 289–311. https://doi.org/10.1111/j.1746-8361.2009.01196.x

Goff, Philip. (2011). A Posteriori Physicalists Get Our Phenomenal Concepts Wrong. *Australasian Journal of Philosophy*, *89*(2), 191–209. https://doi.org/10.1080/00048401003649617

Goff, Philip. (2017). *Consciousness and Fundamental Reality*. Oxford University Press.

Goff, Philip. (2019). *Galileo's Error: Foundations for a New Science of Consciousness*. Rider.

Goff, Philip. (2020). Cosmopsychism, Micropsychism and the Grounding Relation. In William Seager (Ed.), *The Routledge Handbook of Panpsychism* (144–156). Routledge.

Goff, Philip, and David Papineau. (2014). What's Wrong With Strong Necessities? *Philosophical Studies*, *167*(3), 749–62. https://doi.org/10.1007/s11098-013-0195-6

Goff, Philip, William Seager, and Sean Allen-Hermanson. (2020). Panpsychism. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition). Retrieved from https://plato.stanford.edu/archives/sum2020/entries/panpsychism/

Hart, W. D. (1988). *The Engines of the Soul*. Cambridge University Press.

Hempel, Carl Gustav. (1969). Reduction: Ontological and Linguistic Facets. In Sidney Morgenbesser, Patrick Suppes, & Morton White (Eds.), *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel* (179–199). St. Martin's Press.

Hill, Christopher S. (1997). Imaginability, Conceivability, Possibility, and the Mind-Body Problem. *Philosophical Studies*, *87*(1), 61–85. https://doi.org/10.1023/A:1017911200883

Holman, Emmett. (2008). Panpsychism, Physicalism, Neutral Monism and the Russellian Theory of Mind. *Journal of Consciousness Studies*, *15*(5), 48–67.

Horgan, Terence. (1984). Supervenience and Cosmic Hermeneutics. *Southern Journal of Philosophy*, *22*(S1), 19–38. https://doi.org/10.1111/j.2041-6962.1984.tb01546.x

Howell, Robert. (2015). The Russellian Monist's Problems with Mental Causation. *The Philosophical Quarterly*, *65*(258), 22–39. https://doi.org/10.1093/pq/pqu058

Hume, David. (1997). *Dialogues Concerning Natural Religion* (J. C. A. Gaskin, Ed.). Oxford University Press. (Original work published 1779)

Huxley, Thomas Henry. (1866). *Lessons in Elementary Physiology*. Macmillan & Co.

Jackson, Frank. (1982). Epiphenomenal Qualia. *The Philosophical Quarterly*, *32*(127), 127–136. https://doi.org/10.2307/2960077

James, William. (1890). *The Principles of Psychology* (Vol. 1). Henry Holt.

Kastrup, Bernardo. (2018). The Universe in Consciousness. *Journal of Consciousness Studies*, *25*(5–6), 125–155.

Kind, Amy. (2015). Pessimism about Russellian Monism. In Torin Alter & Yujin Nagasawa (Eds.), *Consciousness in the Physical World: Perspectives on Russellian Monism* (401–421). Oxford University Press.

Kirk, Robert. (1974a). Sentience and Behaviour. *Mind*, *83*(329), 43–60. https://doi.org/10.1093/mind/LXXXIII.329.43

Kirk, Robert. (1974b). Zombies vs Materialists. *Proceedings of the Aristotelian Society (Supplementary Volume)*, *48*(1), 135–52. https://doi.org/10.1093/aristoteliansupp/48.1.135

Kripke, Saul. (1980). *Naming and Necessity*. Harvard University Press. (Original work published 1972)

Laplace, Pierre-Simon. (1951). *A Philosophical Essay on Probabilities* (F. L. Emory & F. W. Truscott, Trans.). Dover. (Original work published 1814)

Leibniz, Gottfried Wilhelm. (1998). Monadology. In Roger S. Woolhouse & Richard Francks (Eds. & Trans.), *G. W. Leibniz: Philosophical Texts* (267–81). Oxford University Press. (Original work published 1714)

Levin, Janet. (2002). Is Conceptual Analysis Needed for the Reduction of Qualitative States? *Philosophy and Phenomenological Research*, *64*(3), 571–91. https://doi.org/10.1111/j.1933-1592.2002.tb00161.x

Levin, Janet. (2007). What Is a Phenomenal Concept? In Torin Alter & Sven Walter (Eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism* (87–110). Oxford University Press.

Levin, Janet. (2019). Once More Unto the Breach: Type B Physicalism, Phenomenal Concepts, and the Epistemic Gap. *Australasian Journal of Philosophy*, *97*(1), 57–71. https://doi.org/10.1080/00048402.2018.1447587

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

Levine, Joseph. (2001). *Purple Haze*. Oxford University Press.

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

Lewis, David Kellogg. (1972). Psychophysical and Theoretical Identifications. *Australasian Journal of Philosophy*, *50*(3), 249–58. https://doi.org/10.1080/00048407212341301

Lewis, David Kellogg. (1980). Mad Pain and Martian Pain. In Ned Block (Ed.), *Readings in the Philosophy of Psychology* (216–222). Harvard University Press.

Lewis, David Kellogg. (1983). New Work for the Theory of Universals. *Australasian Journal of Philosophy*, *61*(4), 343–377. https://doi.org/10.1080/00048408312341131

Lewis, David Kellogg. (1986). *On the Plurality of Worlds*. Blackwell.

Loar, Brian. (1990). Phenomenal States. *Philosophical Perspectives*, *4*, 81–108. https://doi.org/10.2307/2214188

Loar, Brian. (1997). Phenomenal States. In Ned Block, Owen Flanagan, & Güven Güzeldere (Eds.), *The Nature of Consciousness* (597–616). The MIT Press. (Original work published 1990)

Lowe, Edward Jonathan. (2012). What is the Source of Our Knowledge of Modal Truths? *Mind*, *121*(484), 919–50. https://doi.org/10.1093/mind/fzs122

Madden, Edward H., and P. H. Hare. (1971). The Powers That Be. *Dialogue*, *10*(1), 12–31. https://doi.org/10.1017/s0012217300032364

Martin, Charles B., and John Heil. (1999). The Ontological Turn. *Midwest Studies in Philosophy*, *23*(1), 34–60. https://doi.org/10.1111/1475-4975.00003

Masrour, Farid. (2015). The Geometry of Visual Space and the Nature of Visual Experience. *Philosophical Studies*, *172*(7), 1813–1832. https://doi.org/10.1007/s11098-014-0390-0

Maudlin, Tim. (1989). The Essence of Space-Time. *PSA: Proceedings of the 1988 Biennial Meeting of the Philosophy of Science Association*, *2*, 82–91.

Maudlin, Tim. (1990). Substances and Space-time: What Aristotle Would Have Said to Einstein. *Studies in History and Philosophy of Science Part A*, *21*(4), 531–561. https://doi.org/10.1016/0039-3681(90)90032-4

Maudlin, Tim. (2007). Completeness, Supervenience and Ontology. *Journal of Physics A: Mathematical and Theoretical*, *40*(12), 3151–3171. https://doi.org/10.1088/1751-8113/40/12/S16

Maudlin, Tim. (2012). *Philosophy of Physics: Space and Time*. Princeton University Press.

Maudlin, Tim. (2013). The Nature of the Quantum State. In Alyssa Ney & David Z. Albert (Eds.), *The Wave Function: Essays on the Metaphysics of Quantum Mechanics* (126–153). Oxford University Press.

Maudlin, Tim. (2019). *Philosophy of Physics: Quantum Theory*. Princeton University Press.

McClelland, Tom. (2013). The Neo-Russellian Ignorance Hypothesis: A Hybrid Account of Phenomenal Consciousness. *Journal of Consciousness Studies*, *20*(3–4), 125–151.

McGinn, Colin. (1989). Can We Solve the Mind-Body Problem? *Mind*, *98*(391), 349–366. https://doi.org/10.1093/mind/XCVIII.391.349

McLaughlin, Brian. (2001). In Defence of New Wave Materialism: A Response to Horgan and Tienson. In Carl Gillett & Barry Loewer (Eds.), *Physicalism and Its Discontents* (319–30). Cambridge University Press.

McLaughlin, Brian. (2019). British Emergentism. In Sophie Gibb, Robin Findlay Hendry, & Tom Lancaster (Eds.), *The Routledge Handbook of Emergence* (23–35). Routledge.

McQueen, Kelvin. (2015). Mass Additivity and A Priori Entailment. *Synthese*, *192*(5), 1373–92. https://doi.org/10.1007/s11229-014-0627-7

Montero, Barbara Gail. (2010). A Russellian Response to the Structural Argument Against Physicalism. *Journal of Consciousness Studies*, *17*(3–4), 70–83.

Montero, Barbara Gail. (2015). Russellian Physicalism. In Torin Alter & Yujin Nagasawa (Eds.), *Consciousness in the Physical World: Perspectives on Russellian Monism* (209–223). Oxford University Press.

Montero, Barbara Gail, and David Papineau. (2016). Naturalism and Physicalism. In Kelly James Clark (Ed.), *The Blackwell Companion to Naturalism* (182–195). Blackwell.

Moore, G. E. (1903). *Principia Ethica*. Cambridge University Press.

Mørch, Hedda Hassel. (2018a). Does Dispositionalism Entail Panpsychism? *Topoi*. https://doi.org/10.1007/s11245-018-9604-y

Mørch, Hedda Hassel. (2018b). Is the Integrated Information Theory of Consciousness Compatible with Russellian Panpsychism? *Erkenntnis*, *84*(5), 1065–1085. https://doi.org/10.1007/s10670-018-9995-6

Mørch, Hedda Hassel. (2020). The Argument for Panpsychism from Experience of Causation. In William Seager (Ed.), *The Routledge Handbook of Panpsychism* (269–283). Routledge.

Nagel, Thomas. (1974). What Is It Like to Be a Bat? *Philosophical Review*, *83*, 435–50. https://doi.org/10.2307/2183914

Ney, Alyssa. (2008). Defining Physicalism. *Philosophy Compass*, *3*(5), 1033–1048. https://doi.org/10.1111/j.1747-9991.2008.00163.x

Ney, Alyssa. (2012). The Status of our Ordinary Three Dimensions in a Quantum Universe. *Noûs*, *46*(3), 525–560. https://doi.org/10.1111/j.1468-0068.2010.00797.x

Ney, Alyssa. (2020). Finding the World in the Wave Function: Some Strategies for Solving the Macro-Object Problem. *Synthese*, *197*, 4227–4249. https://doi.org/10.1007/s11229-017-1349-4

Ney, Alyssa, & Albert, David Z. (Eds.). (2013). *The Wave Function: Essays on the Metaphysics of Quantum Mechanics*. Oxford University Press.

Nida-Rümelin, Martine. (2007). Dualist Emergentism. In Brian McLaughlin & Jonathan Cohen (Eds.), *Contemporary Debates in Philosophy of Mind* (269–86). Blackwell.

Papineau, David. (1998). Mind the Gap. *Philosophical Perspectives*, *12*(S12), 373–89.

Papineau, David. (2001). The Rise of Physicalism. In Carl Gillett & Barry Loewer (Eds.), *Physicalism and its Discontents* (3–36). Cambridge University Press.

Papineau, David. (2002). *Thinking About Consciousness*. Oxford University Press.

Papineau, David. (2006). Comments on Galen Strawson. *Journal of Consciousness Studies*, *13*(10–1), 100–09.

Papineau, David. (2007). Phenomenal and Perceptual Concepts. In Torin Alter & Sven Walter (Eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism* (111–44). Oxford University Press.

Pereboom, Derk. (2011). *Consciousness and the Prospects of Physicalism*. Oxford University Press.

Pereboom, Derk. (2013). Précis of Consciousness and the Prospects of Physicalism. *Philosophy and Phenomenological Research*, *83*(3), 715–727. https://doi.org/10.1111/phpr.12027

Pereboom, Derk. (2015). Consciousness, Physicalism, and Absolutely Intrinsic Properties. In Torin Alter & Yujin Nagasawa (Eds.), *Consciousness in the Physical World: Perspectives on Russellian Monism* (300–323). Oxford University Press.

Place, Ullin T. (1956). Is Consciousness a Brain Process? *British Journal of Psychology*, *47*(1), 44–50. https://doi.org/10.1111/j.2044-8295.1956.tb00560.x

Putnam, Hilary. (1967). The Nature of Mental States. In David J. Chalmers (Ed.), *Philosophy of Mind: Classical and Contemporary Readings, ed*. Oxford University Press.

Raven, Michael. (2015). Ground. *Philosophy Compass*, *10*(5), 322–33.

Robinson, Howard. (1982). *Matter and Sense: A Critique of Contemporary Materialism*. Cambridge University Press.

Roelofs, Luke. (2019). *Combining Minds: How to Think about Composite Subjectivity*. Oxford University Press.

Rosen, Gideon. (2010). Metaphysical Dependence: Grounding and Reduction. In Bob Hale & Aviv Hoffmann (Eds.), *Modality: Metaphysics, Logic, and Epistemology* (109–36). Oxford University Press.

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

Russell, Bertrand. (1921). *The Analysis of Mind*. George Allen and Unwin.

Russell, Bertrand. (1927). *The Analysis of Matter*. Kegan Paul.

Russell, Bertrand. (1959). *The Problems of Philosophy*. Oxford University Press. (Original work published 1912)

Schaffer, Jonathan. (2009a). On What Grounds What. In David Manley, David J. Chalmers, & Ryan Wasserman (Eds.), *Metametaphysics: New Essays on the Foundations of Ontology* (347–83). Oxford University Press.

Schaffer, Jonathan. (2009b). Spacetime the One Substance. *Philosophical Studies*, *145*(1), 131–148. https://doi.org/10.1007/s11098-009-9386-6

Schaffer, Jonathan. (2010). Monism: The Priority of the Whole. *Philosophical Review*, *119*(1), 31–76. https://doi.org/10.1215/00318108-2009-025

Schaffer, Jonathan. (2016). Grounding in the Image of Causation. *Philosophical Studies*, *173*(1), 49–100. https://doi.org/10.1007/s11098-014-0438-1

Schaffer, Jonathan. (2017). The Ground Between the Gaps. *Philosophers' Imprint*, *17*(11), 1–26.

Schaffer, Jonathan. (forthcoming). Ground Functionalism. In *Oxford Studies in Philosophy of Mind*.

Seager, William. (1995). Consciousness, Information, and Panpsychism. *Journal of Consciousness Studies*, *2*(3), 272–288.

Shani, Itay. (2015). Cosmopsychism: A Holistic Approach to the Metaphysics of Experience. *Philosophical Papers*, *44*(3), 389–437. https://doi.org/10.1080/05568641.2015.1106709

Sider, Theodore. (2009). Ontological Realism. In David Manley, David J. Chalmers, & Ryan Wasserman (Eds.), *Metametaphysics: New Essays on the Foundations of Ontology* (384–423). Oxford University Press.

Sider, Theodore. (2011). *Writing the Book of the World*. Oxford University Press.

Skrbina, David. (2005). *Panpsychism in the West*. The MIT Press.

Smart, J. J. C. (1959). Sensations and Brain Processes. *The Philosophical Review*, *68*(2), 141–56. https://doi.org/10.2307/2182164

Stoljar, Daniel. (2001). Two Conceptions of the Physical. *Philosophy and Phenomenological Research*, *62*(2), 253–281. https://doi.org/10.1111/j.1933-1592.2001.tb00056.x

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

Strawson, Galen. (2006a). Panpsychism?: Reply to Commentators with a Celebration of Descartes. *Journal of Consciousness Studies*, *13*(10–11), 184–280.

Strawson, Galen. (2006b). Realistic Monism: Why Physicalism Entails Panpsychism. *Journal of Consciousness Studies*, *13*(10–11), 3–31.

Strawson, Galen. (2015). Real Materialism (with new postscript). In Torin Alter & Yujin Nagasawa (Eds.), *Consciousness in the Physical World: Perspectives on Russellian Monism* (161–208). Oxford University Press.

Strawson, Galen. (2020). What Does 'Physical' Mean? A Prolegomenon to Physicalist Panpsychism. In William Seager (Ed.), *The Routledge Handbook of Panpsychism* (327–339). Routledge.

Stubenberg, Leopold. (2018). Neutral Monism. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition). Retrieved from https://plato.stanford.edu/archives/fall2018/entries/neutral-monism/

Sundström, Pär. (2018). How Physicalists Can—and Cannot—Explain the Seeming 'Absurdity' of Physicalism. *Philosophy and Phenomenological Research*, *97*(3), 681–703. https://doi.org/10.1111/phpr.12394

Swinburne, Richard. (1986). *The Evolution of the Soul*. Oxford University Press.

Tennant, Neil. (2007). Mind, Mathematics and the Ignorabimusstreit. *British Journal for the History of Philosophy*, *15*(4), 745–73. https://doi.org/10.1080/09608780701605036

Trogdon, Kelly. (2017). Revelation and Physicalism. *Synthese*, *194*, 2345–66. https://doi.org/10.1007/s11229-016-1055-7

van Inwagen, Peter. (1990). *Material Beings*. Cornell University Press.

Wagner, Mark. (2006). *The Geometries of Visual Space*. Lawrence Erlbaum Assoc.

Wilson, Jessica M. (2014). No Work for a Theory of Grounding. *Inquiry: An Interdisciplinary Journal of Philosophy*, *57*(5–6), 535–579. https://doi.org/10.1080/0020174x.2014.907542