

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>

## Phenomenology of Fundamental Reality

Kadić, Nino

*Awarding institution:*  
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

### END USER LICENCE AGREEMENT



Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

### Take down policy

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

# **Phenomenology of Fundamental Reality**

Nino Kadić

PhD – Philosophy

King's College London

2022

*For my aunt Ana and my grandma Jaga*

*You will live forever through the stories we share as a family*

## **Declaration**

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 55 000 words.

## **Acknowledgments**

I wish to sincerely thank my supervisors David Papineau and Matthew Soteriou for always being available, full of understanding, and helping me throughout my studies. Without them, none of this would have been possible. I want to extend my gratitude to my thesis examiners, Sam Coleman, and Henry Taylor, for enthusiastically accepting their role and expressing interest in my work. Special thanks to Bill Brewer for supervising me for a period of time, as well as for his invaluable comments on certain parts of the thesis. I am extremely grateful to Yujin Nagasawa, Philip Goff, and Anna Marmodoro for their assistance with my ideas and with the whole process of being a PhD student. I also thank everyone at the Department of Philosophy at King's College London for their help, the conversations and good times we had at the Philosophy Bar and elsewhere. In addition, the UKRI London Arts & Humanities Partnership and the Frankopan Fund have provided me with funding during the course of my studies, which I genuinely appreciate and cherish. I am thankful to the special people who have motivated and advised me: Andrea Azenić, Tim Bollands, Luca Dondoni, Ena Grabar, Emerson Green, Viktor Ivanković, Giulia Lorenzi, Goran Stambolija, and Kai-Hong Tang. Most of all, I send my love and gratitude to my parents, Katarina and Šerif, who are always by my side, in every possible way, as an endless source of support.

## **Abstract**

Panpsychism, the view that consciousness is present everywhere at the fundamental level of reality, has established itself as an increasingly popular option in the philosophy of mind. Situated between substance dualism and reductive physicalism, panpsychism aims to capture the intuitions behind both, integrating consciousness into the physical world without explaining it in terms of purely physical facts. In this thesis, I offer a defence of panpsychism.

First, I examine influential arguments against physicalism, such as Thomas Nagel's (1974, 1979) perspective-based proposal, Frank Jackson's (1982) knowledge argument, David Chalmers' (1996, 2010) conceivability argument, as well as vagueness arguments (Antony 2006a; Tye 1996). I argue that none of them offers a knockdown case against physicalism, particularly because of the threat presented by the phenomenal concept strategy (Chalmers 2007; Papineau 2007a), though they suggest panpsychism as a promising alternative that avoids many common objections.

Second, I discuss independent arguments for panpsychism, including the Russellian monist view (Goff 2017; Russell 1927/1992), the intrinsic nature argument (Seager 2006; Strawson 2006), recent developments in the literature as presented by Hedda Hassel Mørch (2018, 2020) and Luca Dondoni (2022), Tye's (2021b) representational panpsychism based on the vagueness argument, as well as cosmopsychism (Nagasawa & Wager 2016; Shani 2015) – an alternative to panpsychism that shares its aims whilst claiming to avoid its shortcomings.

Third, I address the most pressing objection to panpsychism: the combination problem or the question of how fundamental consciousness can combine to lead to complex consciousness (James 1890). Focusing especially on the subject-summing version of the combination problem or the issue of how small subjects can combine to form complex subjects, I go over various proposals for either solving or avoiding it, such as Philip Goff's (2016)

phenomenal bonding relation and William Seager's (2010) combinatorial infusion proposal, arguing against both and agreeing at the end with Sam Coleman (2014), who advocates that subject-summing is demonstrably incoherent.

However, as I argue in the fourth and final chapter, the panpsychist still has an option that avoids the subject-summing problem, while also circumventing the objections faced by the alternative view of cosmopsychism. Inspired though not reliant on the philosophy of Gottfried Wilhelm Leibniz (1902/2005), I introduce what I call *monadic panpsychism*, the view that the phenomenal character of any given consciousness-bearing physical ultimate is determined by its relations to other such ultimates featuring in the relevant causal structure.

The theory consists of two key claims. The first is that phenomenal qualities are *relational* rather than intrinsic. This view has precedent in literature on the philosophy of mind, such as in Rudolf Carnap (1928), Gottlob Frege (1956), Moritz Schlick (1959), as well as more recently in David Hilbert and Mark Kalderon (2000) and Austen Clark (2000), though not in the context of panpsychism.

The second claim is that physical ultimates only possess a very basic form of consciousness necessary for experiencing *as such*. There are many ways to conceptualise this basic consciousness: non-relational or intrinsic, rudimentary and non-specific, a point of view without experiences, the minimal subject, a mere conduit for experience, phenomenal space, etc. Russellian panpsychists offer a more detailed proposal in which every physical ultimate has an appropriate associated *type* of experience. For example, all quarks have the same quark-type of experience, assuming for the sake of discussion that quarks are ultimates. Still, arguing for a more minimal notion, as I do, requires making less claims and should not be controversial to any panpsychist.

So, what follows from these two claims? I propose that the relational constitution of phenomenal qualities has its endpoint at the fundamental level rather than at the level of a combined complex or emergent subject. Put simply, physical ultimates guarantee the basic consciousness or the point of view that I occupy, while the relational structure constitutes my rich and full human experience. So to speak, relationally constituted phenomenal qualities ‘anchor’ themselves in physical ultimates. As long as these relations result in a phenomenally unified experience, which is an uncontroversial thesis, then no matter where that process reaches the end of its causal chain, it will necessarily be experienced from a subjective point of view.

Since only qualities but not subjects need to combine, my proposal avoids the subject-summing version of the combination problem, which is arguably an insurmountable challenge to other forms of panpsychism. Because of this, I conclude that my theory contributes to the field, serving as a good starting point for further developments of neither combinatory nor emergent panpsychism with regard to subjects.



## **Table of Contents**

<b>General Introduction</b> .....	<b>11</b>
<b>Chapter I: Physicalism and Anti-Physicalism</b> .....	<b>17</b>
Introduction.....	17
1. Physicalism .....	18
1.1 Defining the Physical .....	19
1.2 Hempel’s Dilemma .....	20
1.3 Alternatives to Theory-Based Conceptions of the Physical.....	21
2. Knowledge-Based Arguments Against Physicalism .....	23
2.1 Nagel’s Bat Argument.....	24
2.2 Jackson’s Knowledge Argument.....	27
3. Objections to Arguments Against Physicalism.....	29
3.1 The Ability Hypothesis .....	29
3.2 Mary vs. Marianna .....	31
3.3 The Phenomenal Concept Strategy .....	33
4. Zombies and Possible Worlds.....	36
4.1 Chalmers’s Conceivability Argument.....	36
4.2 Conceivability and Possibility.....	40
4.3 Chalmers’s Master Argument .....	42
5. The Vagueness Argument.....	44
5.1 Vagueness and Consciousness .....	44
5.2 Phenomenal Concepts are Vague.....	47
5.3 Replies to Papineau’s Objection .....	49
Conclusion .....	52
<b>Chapter II: Arguments for Panpsychism</b> .....	<b>54</b>
Introduction.....	54
1. Russellian Monism.....	55
1.1 The Standard Russellian Argument.....	56
1.2 Objections to Russellian Monism .....	60
1.2.1 Dispositionalism and Dispositional Essences .....	60
1.2.2 Russellian Monist Properties and Causality.....	62
1.2.3 The Collapse of Russellian Monism .....	65
2. A Unified Approach.....	66
2.1 Mørch’s Dispositional Argument.....	67

2.2 Issues with Mørch’s view.....	70
2.3 Reconciliation – Why Not Both? .....	72
2.4 Dondoni’s Solution .....	74
3. Vagueness and Panpsychism .....	77
3.1 Tye’s New Argument.....	78
3.2 Clarifying Consciousness* .....	79
3.3 Tye on Russellian Monism.....	80
4. Discussion.....	83
4.1 Thoughts on Reconciliation .....	83
4.2 A Categorical Formulation of Panpsychism?.....	85
Conclusion .....	86
<b>Chapter III: The Combination Problem for Panpsychism .....</b>	<b>88</b>
Introduction.....	88
1. The Combination Problem.....	89
1.1 The General Idea .....	90
1.2 The Summing of Subjects .....	97
2. Attempts to Solve or Avoid the Combination Problem.....	102
2.1 The Phenomenal Bonding Relation .....	103
2.2 Problems with Phenomenal Bonding .....	105
2.3 Combinatorial Infusion .....	109
2.4 Problems with Combinatorial Infusion .....	111
2.5 Cosmopsychism as an Alternative .....	116
2.6 Tye’s Representational Panpsychism and the Combination Problem.....	119
3. The Unity and Disunity of Consciousness.....	122
3.1 Unity of Consciousness.....	122
3.2 Disunified Consciousness .....	125
4. Discussion.....	132
4.1. Overview of the Proposals .....	133
4.2 Categorical Panpsychism and Combination.....	134
Conclusion .....	135
<b>Chapter IV: Monadic Panpsychism .....</b>	<b>137</b>
Introduction.....	137
1. Microphenomenal Structuralism.....	139
2. Monadic Panpsychism .....	142
2.1 Static Monadic Panpsychism .....	143

2.2 Dynamic Monadic Panpsychism.....	144
2.3 Global Monadic Panpsychism.....	145
3. Objections and Replies .....	147
3.1 Intuitions Against Monadic Panpsychism.....	148
3.2 Quality Combination: The Palette Problem .....	149
3.3 Structure Combination: The Structural Mismatch Problem.....	152
3.4 A Pilot in a Ship .....	155
3.5 Individuating the Dominant Microsubject in DMP.....	161
3.6 GMP and Symmetry.....	162
4. Discussion.....	165
4.1 Tye’s Representational Panpsychism vs. Monadic Panpsychism.....	165
4.2 Categorical Panpsychism and Monadic Panpsychism .....	166
Conclusion .....	167
<b>General Conclusion.....</b>	<b>170</b>
<b>References.....</b>	<b>174</b>

## General Introduction

Panpsychism, in its most basic definition, is the view that consciousness is a fundamental and ubiquitous feature of our world. Though it has a rich history in philosophical and religious thought in many cultures, only recently has it stepped into the spotlight of philosophical interest again. The term itself comes from the Ancient Greek words *pan* (πᾶν), meaning ‘everything’ or ‘whole’, and *psyche* (ψυχή), meaning ‘soul’ or ‘mind’. The term itself was first used by Frane Petrić (lat. Franciscus Patricius), a 16<sup>th</sup> century philosopher and scientist from Cres, Republic of Venice (now Croatia), in his magnum opus *Nova de Universis Philosophia (The New Universal Philosophy)*, in which he discussed the theories of *Pampsychia* (‘all-soul’) and *Pancosmia* (‘all-order’) (Petrić 1979, Purnell 2004). His panpsychist view is related to the concept of *anima mundi*, the world soul that is present throughout the whole cosmos (Petrić 1979), Purnell 2004.

However, much earlier than that, panpsychism and similar views have been present in human thought. Animism, one of the earliest of human spiritual beliefs, could be considered as a form of panpsychism or as its predecessor. Panpsychism-like beliefs are also present in some schools of Buddhism, such as the Flower Garland School (華嚴, *huáyán*), where the jewel-adorned net of the god Indra symbolises the mutual interdependence of everything, with each jewel reflecting every other jewel in the net (Norden and Jones 2019). The similarity to G. W. Leibniz’s (1902/2005) hierarchy of monads, which is also sometimes described as a panpsychist view, is striking: each monad is a mirror reflecting the entirety of reality. Some elements of panpsychist thinking can also be found in the Hinduist tradition of *Advaita Vedānta*, which sees the fundamental nature of the universe as consciousness (Dalal 2021). The first philosopher of the Ancient Greek tradition, Thales of Miletus, is also sometimes seen as a panpsychist, due to his belief that some non-living objects, such as magnets and amber, are

enminded self-movers, or his supposed belief that the universe is living and full of spirits (Goff, Seager, & Allen-Hermanson 2017).

In Western philosophy, Baruch Spinoza (1632–1677) and G. W. Leibniz (1646–1716) both held panpsychist views. However, the 19<sup>th</sup> century was particularly significant for panpsychism (Goff, Seager, & Allen-Hermanson 2017). Some prominent 19<sup>th</sup> century thinkers who espoused versions of panpsychism are: Gustav Fechner, Wilhelm Wundt, Willian James, Josiah Royce, and many others (Goff, Seager, & Allen-Hermanson 2017). Arthur Schopenhauer (1788–1860) was an important figure as well, due to his argument that will is the inner nature of all things (Goff, Seager, & Allen-Hermanson 2017). William James (1842–1910) is worthy of special mention considering that his argument against some forms of panpsychism is still discussed today, including in this work, extensively in Chapter III. Many more examples could be found since history is rich with panpsychist thinking, though I will leave it at this, for the sake of brevity.

But what makes panpsychism popular nowadays, in the 21st century? Panpsychists often claim that their theory is an attractive middle way between physicalism and dualism (Goff, Seager, & Allen-Hermanson 2017). After many years of physicalism dominating the debate in philosophy of mind, some philosophers became increasingly more interested in finding alternatives that can integrate consciousness into our general picture of the world. I believe that there might also be an emotional element to this, considering the often dry conclusions that physicalism offers: consciousness just *is* the brain or, more extremely, consciousness *does not exist*, at least not in a way we commonly think of it. So, for many, physicalism is counterintuitive and unable to provide a satisfying account of consciousness (Goff, Seager, & Allen-Hermanson 2017). In the case of dualism, the problem mostly pertains to how two radically different kinds of substances could interact. How can the immaterial soul affect my material body when I raise my hand? Dualism leaves us with a very disunified picture

of the world and offers little answers to this problem of interaction (Goff, Seager, & Allen-Hermanson 2017). Panpsychism, in contrast, ‘builds’ consciousness right into reality, leading to the (admittedly also counterintuitive) view that every fundamental entity of the cosmos is conscious, along with some aggregates of those entities, such as humans and sufficiently complex animals. There is no interaction problem since there are no two distinct substances: consciousness just is a part of the physical universe. The common intuition that we are ‘more than just our brains’ is also preserved. I hint again at the emotional element that might motivate an interest in panpsychism since I have noticed many interested non-academics accepting the theory for spiritual reasons. I will remain neutral on whether that is desirable, but it is clear that many see panpsychism as a more vivacious and complete picture of reality. Indeed, even within academia, some see panpsychism as more than just a theory of consciousness. Freya Mathews (2011), for example, states that panpsychism is her attempt to explain the *world* and argues for a holistic or cosmological version of panpsychism, with an emphasis on its significance for environmental philosophy.

In this work, however, I do treat panpsychism solely as a theory of consciousness. Though at times speculative, the main aim of my proposal is to present a coherent and robust version of panpsychism that avoids a plethora of issues and, despite being counterintuitive, offers a complete account of consciousness that integrates it into our general worldview. Specifically, my formulation of panpsychism will focus only on the fundamental level of reality, and I will try to construe the rich human consciousness using only the resources found at the smallest level. I will consider some historical sources, though the main intention will be to rely on cutting-edge research on panpsychism, such as on recent books and articles that push the theory forward. Throughout the work, other theories of consciousness will also be examined and compared to panpsychism, as well as a wide range of objections and replies.

In Chapter I, I will discuss physicalism and anti-physicalism. Physicalism is, simply put, the view that everything is physical. How the ‘physical’ should be defined, however, is less than straightforward, so the chapter will begin with a brief discussion on what physicalism is. After that, influential arguments against physicalism will be analysed, as well as objections and replies raised against them. These arguments share the general strategy of inferring an ontological conclusion against physicalism from epistemological premises. Then, I will examine a strong and general framework that the physicalist has available in order to block a wide range of anti-physicalist views: the phenomenal concept strategy, or the view that we have two different sets of concepts – phenomenal and physical – but that this does not necessarily mean that reality is non-physical. Finally, I will consider a unique and independent argument against physicalism, based on the intuition that consciousness is a sharp and determinate concept, which aims to preserve our intuitions about what consciousness is and accuse physicalism of failing to do so. The main aim for this chapter is to argue *against* the conclusiveness of common anti-physicalist arguments and *for* the strength of the unique and independent vagueness argument, which will limit and direct the further aims of this work.

In Chapter II, the focus shifts from arguments against physicalism to self-standing arguments for panpsychism. Here, I will first explore Russellian monism, which is nowadays arguably the most popular defence of panpsychism, as well as some particular objections to it. Russellian monists state that science only tells us what matter *does*, but not what matter *is*. In addition, I will also explore more recent modifications to Russellian monism and whether they succeed as better alternatives. The debate related to this will involve an extensive analysis of the ontology of properties, which pertains to different possible ways of explicating panpsychism. Another specific and unique pathway towards panpsychism, also based on considerations of vagueness with regard to consciousness, will also be addressed and compared to Russellian monism. The main aim of this chapter is thus defending panpsychism, mostly

through Russellian monism and its modifications, but also in a broader sense, through other arguments, and as a philosophical project that is worthy of serious consideration. These considerations will also further limit and define how I will present my original proposal in the forthcoming chapters.

In Chapter III, I will offer an extensive analysis of the combination problem, the question of how consciousness at the fundamental level leads to more complex forms of consciousness, which is likely the most serious and pressing challenge that panpsychism has to face. I will go over the main types of the combination problem, relating to how subjects and qualities of experience combine, as well as over a wide range of proposals for solving or avoiding the combination problem, while also connecting the discussion to the views mentioned in previous chapters. The aim of this chapter will also be to limit what I am allowed to argue for since I will accept that the combination problem *succeeds* in blocking (certain) forms of panpsychism. That is, I will accept that subjects, specifically, cannot combine to create new, more complex subjects. Thus, what I will have available for my original proposal in the next chapter will be whatever is left for constructing a form of panpsychism after accepting that the most consequential objection to it succeeds.

In Chapter IV, I will present an original version of panpsychism that respects all the limits defined in the previous three chapters. Partially inspired by the philosophy of G. W. Leibniz, my view will focus only on the fundamental layer of reality, using (arguably) minimal notions in order to present a full account of consciousness. This chapter will be slightly more speculative, though I believe that my proposal will be significant. If it succeeds in all its aims, it will at the very least show that current discussions on panpsychism are incomplete, and at most demonstrate how consciousness should be conceptualised. To strengthen some of the claims that I will make, I will rely on both historical and recent sources in this chapter, with the goal of demonstrating that there is a precedent for my view. Some alternatives will also be



considered, as well as contenders to my proposal. The main aim of this chapter is thus to respect the constraints given through the discussion so far, to address the combination problem, as well as to avoid certain other issues for panpsychism discussed throughout the work.

Therefore, the three limiting factors that will permeate the entire line of thinking in this work are the following: a) standard arguments against physicalism do not conclusively refute it, b) more independent routes towards panpsychism are needed instead, c) the combination problem is successful as an argument against some forms of panpsychism. Details will become clearer later on, though these factors should be kept in mind while reading this work. If successful, my proposal will be a novel way of framing panpsychism, different from the most commonly discussed formulations of the theory.

# Chapter I: Physicalism and Anti-Physicalism

## Introduction

Panpsychism, the theory that consciousness is a fundamental and ubiquitous feature of reality, has been gaining traction recently in the philosophy of mind. Some philosophers arrive at panpsychism after being convinced by common arguments against physicalism, which is broadly the view that everything that exists is physical, though definitions vary. In this chapter, I will call into question whether such arguments present a strong case against physicalism, as well as whether they offer a solid basis for establishing panpsychism, which will lead to examining a more promising alternative.

First, I will briefly discuss views on how to define ‘physical’, going over various proposals and explaining what options the physicalist has available. Second, I will explore influential anti-physicalist arguments, as well as the objections they have to face. Third, I will analyse objections to common arguments against physicalism, including the phenomenal concept strategy which has the aim of blocking anti-physicalists from inferring ontological conclusions from the apparent explanatory gap between the physical and the mental. Fourth, I will introduce the vagueness argument against physicalism, which relies on the intuition that the concept of consciousness is sharp and determinate and can therefore not be applied to any vague or indeterminate physical concept.

Finally, I will conclude that standard anti-physicalist arguments do not present a strong case against physicalism, nor do they lay a good groundwork for panpsychism, largely due to alternative explanations that the physicalist can give in order to prevent non-physicalist ontological conclusions. However, the vagueness argument is not undermined by the

objections raised against the previously discussed anti-physicalist views, and it directly suggests panpsychism as one of the options that can accommodate the robust intuition that consciousness either is or is not present, as a sharp concept with no possible borderline cases.

Both the negative conclusion that popular arguments against physicalism do not decisively refute it, as well as the positive conclusion that the vagueness argument is plausible and in support of panpsychism, will define how panpsychism will be approached throughout the rest of this work. As an implicit claim, the conclusion will entail that panpsychism is worthy of serious consideration in part *because* it is not necessarily tied to typical discussions on physicalism and anti-physicalism. I hope to show that there are different and independent routes towards panpsychism, each with its own set of difficult but new obstacles. That alone should pique the interest of researchers seeking new understandings of consciousness.

## **1. Physicalism**

Prior to exploring arguments against physicalism, I will examine how the view is defined, focusing first on a broad and theory-based conception of physicalism and the physical. Then, I will discuss Hempel's dilemma, which is a common and pressing objection that such conceptions face. This will be followed by a section on how the physicalist can respond to or avoid Hempel's dilemma. Finally, I will explain how this discussion pertains to the rest of what follows in this work, while arguing that the physicalist has the resources to offer a coherent conception of the physical.

## 1.1 Defining the Physical

In the most straightforward sense, physicalism is the thesis that everything is physical (Stoljar 2021). The physical, in turn, is defined as encompassing any property if and only if it is the kind of property described by physical theory (Stoljar 2021). This is a *theory-based* conception of the physical. For example, if physical theory (or just physics) describes the property of having mass, then having mass is a physical property (Stoljar 2021).

This signifies how physicalism differs from materialism, its historical precedent. Materialism was a metaphysical doctrine that tried to limit physics a priori, through conditions that matter must be “solid, inert, impenetrable and conserved, and to interact deterministically and only on contact” (Crane & Mellor 1990: 186). After physics has shown that matter does not meet these conditions, materialism transformed into physicalism, the claim that the empirical world contains “just what a true complete physical science would say it contains” (Crane & Mellor 1990: 186). Instead of making metaphysical claims to limit what the physical is a priori, physicalists make the linguistic thesis that every statement is synonymous with some physical statement (Stoljar 2021). As Philip Goff (2017: 14) puts it, pure physicalism is the view that it is possible, at least in principle, to describe the complete nature of fundamental reality through the vocabulary of the physical sciences. Thus, physicalism as it is known today is an a posteriori view, based on how physical theory describes the empirical world (Crane & Mellor 1990: 186).

While at first glance this broad definition of physicalism might seem intuitive and appealing, it faces various issues. Despite the current popularity of the view, the general question regarding theory-based conceptions of physicalism is whether there is a clear, credible, non-vacuous and non-circular definition of the physical (Crane & Mellor 1990: 185).

So, now, I will turn to a common objection aimed against theory-based conceptions of the physical.

### 1.2 Hempel's Dilemma

Carl Hempel (1969) introduced the notion that any definition of physicalism will end up being either trivial or false (Stoljar 2021). The main question is whether the physical is defined via reference to current physics, or future or ideal physics? The following are the horns of Hempel's dilemma (Stoljar 2001a: 394):

- 1) If physicalism is defined via reference to contemporary physics, then it is probably false because it is highly unlikely that contemporary physics is complete.
- 2) If physicalism is defined via reference to a future or ideal physics, then it is trivial because we cannot predict what future physics will encompass.

Both options present a theory-based conception of the physical. According to (1), current physics tells us what the physical is, but since contemporary physics is likely incomplete, any formulation of the physical based on it is also most likely false and incomplete (Stoljar 2021). According to (2), something is physical if it will be encompassed by a complete and ideal future physics. The problem is that we cannot predict what future physics will postulate – it might even contain mental items as part of its ontology and entail a form of panpsychism (Stoljar 2021). The conclusion of Hempel's dilemma is thus that there is either no notion of the physical, no clear notion, or no notion that can play the 'physical' role that philosophers want it to play (Stoljar 2001a: 394).

Tim Crane and D. H. Mellor (1990: 185) go a step further and claim that there is no clear and credible definition of physicalism and that no non-vacuous interpretation of it is true. Expanding on Hempel's dilemma, they argue that the 'physical' is "either taken in [a] very restrictive sense, in which case physicalism is clearly false; or it is taken in a very broad sense, in which case the doctrine is almost empty" (as reported by Crane 1993: 224). The challenge for the physicalist is thus to provide a defensible and substantial conception of physicalism (Crane 1993: 224).

### 1.3 Alternatives to Theory-Based Conceptions of the Physical

While Hempel's dilemma affects theory-based conceptions of the physical, it does not affect object-based conceptions, according to which a physical property is a property which *either* is the sort of property required by a complete account of the intrinsic nature of paradigmatic physical objects and their constituents *or* else is a property which logically (or a priori) supervenes on the sort of property required by a complete account of the intrinsic nature of paradigmatic physical objects and their constituents (Stoljar 2001a: 395–6). According to the object-based conception, if rocks are paradigmatic physical objects, then 'being a rock' is a physical property and, also, if mass is a property of paradigmatic physical objects, then mass itself is a physical property (Stoljar 2001a: 396).

Daniel Stoljar (2001a: 396–7) traces this distinction back to Kantian physicalism, which distinguishes between two classes of properties: a) the causal cum relational properties that physical theory describes, and b) the properties that duplicate physical objects will share. The former class corresponds to theory-based or *t-physical* properties, while the latter class corresponds to object-based or *o-physical* properties (Stoljar 2001a: 397). As Stoljar argues, if Hempel's dilemma shows something, it shows that a particular proposal about how to analyse

the concept of ‘physical’, in terms of reference to a particular stage of the development of physics, is mistaken – however, it does not necessarily show that there is no clear concept of ‘physical’ whatsoever (Stoljar 2001a: 395).

A further option that potentially avoids the dilemma centres on construing the physical in terms of what it is not. Proponents of the *via negativa* view offer the following definition (Stoljar 2021):

*F* is a physical property if and only if *F* is a non-mental property.

The aim is to equate ‘physical’ with ‘non-mental’ and consequently end up with a definition of physicalism with reference to a complete account of the non-mental (Spurrett & Papineau 1999). David Spurrett and David Papineau define ‘non-mental’ *causally*, via reference to physiology:

“It seems highly plausible that the non-mental is in fact complete [...] To deny this is to suppose that some non-mental effects are due to irreducibly mental causes. [...] But we take it that the empirical evidence, especially in the form of nineteenth- and twentieth-century physiological research, now weighs strongly against such irreducible mental causes.” (1999: 26–7)

Importantly, Spurrett and Papineau (1999) are not committed to the completeness of current physiological theory. Instead, they merely infer, based on physiology being unable to discover *sui generis* mental causes, that there are likely no such causes (Montero & Papineau 2005: 235). If there are no mental causes, then whatever is causal is physical. In that way, they attempt to refer to current physiology without suffering from the first horn of Hempel’s dilemma: physiology is incomplete, but it has not yet discovered mental causes, so it is probable that it will not do so in the future either (Montero & Papineau 2005: 235–6).

Finally, the physicalist can depend on a *global supervenience thesis*: the idea that there is no change or difference without a non-mental change or difference, so that two things “will never change or differ in any way without also changing or differing in some non-mental way” (Crane & Mellor 1990: 203). If that is the case, the physical “excludes the mental by being that on which everything else, including the mental, supervenes” (Crane & Mellor 1990: 203). This is the most minimal way of defining physicalism.

The goal of this section is not to firmly decide on one conception of physicalism and the physical, but to show that the physicalist has ways of circumventing Hempel’s dilemma, as well as to show how the objection leads to various differing definitions of the physical. For the purposes of this work, it is enough to establish that there are coherent conceptions<sup>1</sup> available to the physicalist, which, I believe, is an uncontroversial claim.

## **2. Knowledge-Based Arguments Against Physicalism**

I would first like to clarify why the title of this section refers to *knowledge-based* arguments against physicalism. While only Frank Jackson’s (1982) proposal is commonly known as the knowledge argument, I believe that Nagel’s (1974) proposal can also be seen as part of that category. As will be shown, Jackson focuses on knowledge of non-physical, qualitative properties of experience, while Nagel focuses on knowledge of what it is like to occupy a particular type of perspective. The title thus refers to knowledge-based arguments in a broader sense, not exclusively to Jackson’s view. Moreover, the goal of this section is to

---

<sup>1</sup> There is also Alyssa Ney’s (2008) attitudinal conception of physicalism, according to which physicalism is the attitude to base ontology purely on what physics describes. There are also structuralist definitions that define the physical via reference to mathematical, logical, nomic or causal vocabulary (see Chalmers 2020, Goff 2017). This is beyond the scope of this section, though it further shows that the physicalist has available options.



compare and connect the two arguments, before moving on to the (oftentimes shared) issues that they face.

## 2.1 Nagel's Bat Argument

In *What is it like to be a bat?*, Thomas Nagel (1974) introduced one of the most influential attacks on physicalism, which aims to show that physical theory will leave perspectives and facts dependent on perspectives out from an objective description of reality, thus failing to describe reality *as such*. Nagel chose the bat as an example because of the animal's radically different perceptual apparatus: bats use echolocation to perceive the world and rely on sonar to detect reflected sounds and get a sense of their environment. This is very much unlike human perception<sup>2</sup>: for example, bats are capable of finely distinguishing auditory information and so perceive distance, size, shape, motion, and texture, which we do through visual perception (Nagel 1974: 438).

Nagel's proposal goes as follows: even if we have an exhaustive third-person physical account of a bat's perception, behaviour, physiology, it does not seem that this will give us knowledge of what it is like to be a bat *for a bat* (Nagel 1974: 439). To understand this in greater depth, it is important to see how this connects to Nagel's conceptions of subjectivity and objectivity. He defined 'subjective' as that which is *perspective-dependent* and 'objective' as that which is *perspective-independent*. The notion of 'perspective' here is similar to the pre-theoretical 'point of view' as the locus of the 'I' which looks at the world, such as in the case

---

<sup>2</sup> See (Schwitzgebel & Gordon 2000) for an argument that humans do use echolocation as well, although most people's knowledge of that experience is very limited. I do not see this as undermining Nagel's position in any way considering that the bat's echolocation is just an example that can be replaced with other actual or hypothetical cases (e.g. an alien's natural sense of pulsar proximity or the chemical communication of various animals such as deer, mules, and big cats).

of someone looking at an apple in front of them, seeing its front but not its back, from the unique location that they occupy. For Nagel, the condition for subjectivity is having a perspective on the world, and there are facts which pertain to particular perspectives which would be missing from a purely objective description of reality. His perspective-independent conception of objectivity does not mean that a fact is completely disconnected from any particular perspective, essentially being perspective-*less*, as an ideal and completely impartial point of view, but rather that it is not tied to any *particular* perspective or perspective type. Simply put, the demand is that every fact should be accessible to every sufficiently complex type of perspective, not for there to be some godly, ultimate perspective which is completely objective and able to perceive reality in all its facets simultaneously, as some sort of an ideal and omniscient observer.

In *The View from Nowhere*, Nagel (1986: 27) argues that such a perspective-independent notion of objectivity is inappropriate for the task of understanding consciousness and giving us a complete understanding of reality because the more we try to find perspective-independent descriptions of reality, the more will we be unable to capture what he deems to be perspective-dependent, subjective facts. Nagel argues that it seems unlikely that we could understand the real nature of human experience *better* by leaving behind the human point of view in search of a more objective explanation – one that could be grasped by intelligent beings who could not even imagine what it is like to be human (Nagel 1986: 4).

To turn back to the bat, the idea is that there is *something it is like* to be a bat which is inaccessible to humans since we have a very different *type* of perspective. While he allows that subjects of a sufficiently similar type of perspective can form a fairly accurate conception of what it is like to be another subject, such as among humans, when it comes to such knowledge between two very different perspective types, such as when humans are trying to understand

bats, we cannot form a clear conception (Nagel 1974: 441–2). Some of the methods through which we might be able to extrapolate what it is like to be a bat from our own case are imagination and gradual transformation (Nagel 1974: 439). The former will not help because imagining that I am hanging upside down, flying around, having poor vision, or what echolocation might be like will not give me the answer about what it is like to have those experiences *for the bat* but only what it would be like for me to behave as a bat does (Nagel 1974: 439). That is, such a leap of imagination would always start from my own perspective and be confined to my perspective type.

In the case of the latter, a gradual transformation into a bat would not provide an answer either since I cannot imagine *now* what such a transformed stage of myself *in the future* would be like (Nagel 1974: 439). In that future stage, I would have a completely different neurophysiological makeup, so there would essentially be two (sufficiently) distinct subjects with regard to their perspective type, the old me and the future me. This would just confirm Nagel's argument since it would show that two very different types of perspectives cannot mutually understand what it is like to be each other. So, his conclusion is that there is a what-it-is-likeness pertaining to perspectives – a qualitative aspect of conscious experience – which is perspective-dependent, and which would end up missing in any fully objective description of reality. The crucial part is having the right kind of *knowledge*: what we are lacking is knowledge of experiences given through or as part of a certain perspective type.

To summarise, the argument is that there are facts which could never be represented or comprehended by human beings *in principle* since, whatever the status of such facts is, they appear to be facts that embody a particular point of view (Nagel 1974: 441). If the facts about what it is like to be an organism for the experiencing organism are accessible only from their perspective type, it is “a mystery how the true character of experiences could be revealed in

the physical operation of that organism” (Nagel 1974: 442). It is clear that Nagel does not directly argue that physicalism is false, but rather that we cannot see how it could be true since all descriptions of reality that physicalism can provide will necessarily be third-person objective descriptions, without accounting for first-person subjective descriptions involving perspective-dependent facts, leaving the physicalist with an incomplete picture of the world.

## 2.2 Jackson’s Knowledge Argument

Frank Jackson’s knowledge argument, which he described in *Epiphenomenal Qualia* (1982), is a thought experiment involving the scientist Mary who lives in a black-and-white room and knows all physical facts (or at least all physical facts related to vision and colour perception). When she leaves the black-and-white room and sees red for the first time, will she gain new knowledge from that experience? This is how Jackson outlines the thought experiment:

“Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room *via* a black and white television monitor. She specialises in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on. [...] What will happen when Mary is released from her black and white room or is given a colour television monitor? Will she *learn* anything or not? It seems just obvious that she will learn something about the world and our visual experience of it. But then it is inescapable that her previous knowledge was incomplete. But she had all the physical information. *Ergo* there is more to have than that, and Physicalism is false.” (Jackson 1982: 130)

Jackson's specific point is that no amount of physical knowledge will entail knowledge of what it is like to experience red. Because there is knowledge *other* than physical knowledge, namely knowledge of non-physical facts, physicalism must be false. After her release, Mary gets to know what red is like in her own experience or what it is like for red to *appear* in her experience. She gains knowledge of a non-physical fact not entailed by the complete physical knowledge she had while locked in her black-and-white room.

The qualitative, introspectively accessible what-it-is-likeness of having a particular conscious experience is what Jackson calls *qualia* (singular: 'quale') (Tye 2021a). Qualia are non-physical *properties*<sup>3</sup> which are epiphenomenal, meaning that they are not causally efficacious in any way: "Qualia cause nothing physical but are caused by something physical" (Jackson 1982: 134). Common examples for qualia in the literature include the throbbing pain of a headache, seeing the colour red, the taste of wine, the feeling of sandpaper on fingertips, and so on (Tye 2021a). For Jackson, knowledge that lies beyond the physical is not merely knowledge of certain perspective-dependent facts which are dependent on the subject taking up the relevant point of view, accessible only to particular perspective-types, but rather knowledge of *non-physical* or *qualitative properties* that feature in conscious experience.

For Nagel, 'what it is like' means 'what it is like to occupy a certain point of view on the world'. There is something it is like to be a bat because there is something it is like to occupy a bat-perspective. In the case of Jackson, referring to qualia means referring to 'raw feels', which he defines as "certain features of the bodily sensations especially, but also of certain perceptual experiences, which no amount of purely physical information includes" (Jackson 1982: 127). Therefore, in Jackson's framework, the 'what-it-is-likeness' of an

---

<sup>3</sup> It is important to note that Jackson does not specifically refer to qualia as non-physical *properties* but as non-physical *facts*, though it is most commonly understood that the facts in question are non-physical because they exemplify non-physical properties (Nida-Rümelin & O Conaill 2019).

experience is *qualitative* and constituted by the phenomenal character of conscious experiences or qualia, while in Nagel's framework, 'what-it-is-likeness' is *perspectival* in the sense that it is constituted by the subject having a unique perspective on the world. Jackson thus makes the more explicit ontological claim that there are non-physical properties, beyond the knowledge of purely physical facts.

### **3. Objections to Arguments Against Physicalism**

Now, I will discuss common objections to arguments against physicalism, starting with the ability hypothesis, which aims to refute Jackson's knowledge argument by blocking its ontological conclusion. Then, I will explore Martine Nida-Rümelin's (1998) Marianna scenario which has the goal of refuting both the ability hypothesis and Jackson's knowledge argument. Finally, I will turn my attention to the phenomenal concept strategy, which is a broad attempt of preventing any ontological conclusions from epistemological premises. The purpose of this section is to show that anti-physicalist arguments are inconclusive and thus to limit and define the direction of the rest of this work.

#### 3.1 The Ability Hypothesis

An influential objection raised against the knowledge argument is David Lewis's (1983, 1988) *ability hypothesis* – an alternative explanation of what it means to learn what an experience is like. Lewis argues that Mary does not acquire any new *propositional* knowledge (i.e. knowledge that something is the case or *factual* knowledge) about seeing red for the first time but only gains a set of abilities:

“Rather, knowing what it's like is the possession of abilities: abilities to recognize, abilities to imagine, abilities to predict one's behavior by means of imaginative experiments. (Someone who knows what it's like to taste Vegemite can easily and reliably predict whether he would eat a second helping of Vegemite ice cream.) Lessons cannot impart these abilities—who would have thought that they could? There is a state of knowing what it's like, sure enough. And Vegemite has a special power to produce that state. But phenomenal information and its special subject matter do not exist.” (Lewis 1983: 131)

So, knowing what an experience is like just *is* the possession of a certain set of abilities – it is knowing-*how* rather than knowing-*that* (Lewis 1988: 288). The difference between Mary being in the black-and-white room and Mary post-release is based on her gaining relevant cognitive abilities: if Mary had never experienced red, she would lack the ability to remember, recognise and imagine red (Tye 2000: 7). She might know, while still confined in her room, how to triangulate each colour experience exactly in a network of resemblances, but she will still not know what any node in the network is like (Lewis 1988: 266). According to the ability hypothesis, knowing what something is like is *identified* with abilities to be in certain propositional states, so Mary's failure to know what the nodes are like is merely the result of her lack of crucial abilities (Tye 2000: 6, 9). It does not follow that knowing what pain is like, for example, is knowing a *truth* but rather that it consists in the possession of certain cognitive abilities, such as the ability to recognise, imagine, compare pain, etc. (Tye 2000: 10).

Thus, when Mary leaves her room and sees red for the first time, she does acquire new knowledge, but that knowledge is nothing but knowing-*how* (Nida-Rümelin & O Conaill 2019). This does not exclude that there is some propositional knowledge that can be acquired through acquaintance with first-person experiences, though this does not undermine the ability

hypothesis since its proponents must only insist that propositional knowledge of that kind is accessible in other ways as well, rather than just through direct experience (Nida-Rümelin & O Conaill 2019).

To further strengthen the impact of the ability hypothesis, Lewis (1988: 270) argued that the *only* alternative to it is the *phenomenal information hypothesis*, the claim that coming to know what something is like is propositional in the sense of eliminating open epistemic possibilities. When we get physical information, physical possibilities are narrowed; when we have an experience and gain the phenomenal information, previously open possibilities are eliminated and that is what it means to learn what an experience is like (Lewis 1988: 270–1). While Lewis admitted that he could not refute the phenomenal information hypothesis, he did argue that the ability hypothesis should be preferred since it has the same explanatory power whilst also being compatible with physicalism (Lewis 1988: 277).

### 3.2 Mary vs. Marianna

Martine Nida-Rümelin (1998) introduced a thought experiment designed to demonstrate that the knowledge argument does not lead to an ontological conclusion that is incompatible with physicalism even if the ability hypothesis is unsuccessful, presenting another way of refuting Jackson's proposal. Marianna, just like Mary, lives in a black-and-white room and has never seen colour, not even in dreams or altered states of consciousness (Nida-Rümelin 1998: 52). However, before being released, she agrees to participate in a psychological experiment: she will not leave the room she spent her whole life in, but the interior decoration will be changed so that it is now full of artificial objects (e.g. furniture) of all colours (Nida-Rümelin 1998: 53). She is not allowed to see natural objects, such as ripe tomatoes or



photographs of landscapes, since she would know what their colour is, based on her knowledge of physical facts.

During the course of the experiment, she is presented with slides showing her the colours red, blue, green, and yellow (Nida-Rümelin 1998: 52). She especially likes the red slide and incorrectly identifies it as the colour of the sky, since she has been told that the sky is beautiful. While Marianna *nonphenomenally* believes the statement ‘the sky appears blue to normally sighted people’, she *phenomenally* believes that the sky appears red to normally sighted people since looking at the sky must be as beautiful as looking at the red slide (Nida-Rümelin 1998: 53-5). That is, she uses the word ‘blue’ to refer to what she sees when she looks at the red slide, believing that it is the colour of the sky. When Marianna is released and finally sees the sky, she understands her mistake and correctly states that the sky is blue, in the usual sense, realising that she conflated the two colours.

While she was looking at the slides, she gained the capacity to form new concepts about colour experiences, to ask questions, form new (if false) hypotheses, etc. After her release, she acquires knowledge about the correct *application* of phenomenal concepts. So, the Marianna argument distinguishes between these two steps in her epistemic process (Nida-Rümelin & O’Conaill 2019):

- a) acquaintance with kinds of colour experiences by having and remembering them,
- b) knowing what colour experiences other subjects have at a given time.

After her release, Marianna acquires a particular kind of phenomenal belief that the sky appears blue to normally sighted people and not red, which involves the application of the relevant phenomenal concept, correcting her former mistake.

This leads to two conclusions. First, Marianna's knowledge that the sky looks blue phenomenally does not consist in her knowing that it looks blue nonphenomenally *plus* certain cognitive abilities (Nida-Rümelin 1998: 70). Before Marianna realises that she is mistaken with regard to how the sky appears to normally sighted observers, she is still able to imagine blue at will, recognise it, and so on (Nida-Rümelin 1998: 70). That is, she can already think about the experience of blue, she has a phenomenal concept of blue, even if she incorrectly applies it. In order to correct her mistake, she does not need to acquire a further ability, but to acquire the corresponding phenomenal belief that the sky appears phenomenally blue (Nida-Rümelin 1998: 70). This refutes the ability hypothesis since it shows that abilities are insufficient to explain Marianna's epistemic progress.

Second, Nida-Rümelin's thought experiment counters Jackson's conclusion that there are non-physical properties that we can grasp propositionally. According to her, one can accept the epistemic state view of phenomenal belief and knowledge according to which "we can find some 'content-identical' item of nonphenomenal knowledge for every item of phenomenal knowledge", thus rejecting the idea that gaining phenomenal knowledge necessarily involves acquisition of *new* facts (Nida-Rümelin 1998: 69). If that is the case, the knowledge argument fails to demonstrate that there are non-physical facts.

### 3.3 The Phenomenal Concept Strategy

Phenomenal concepts are used both by anti-physicalists, who think that they refer to non-physical phenomenal properties, and physicalists, who argue that phenomenal concepts are only conceptually but not ontologically distinctive, referring to nothing over and above the physical properties, which we can also refer to using nonphenomenal physical concepts (Papineau 2007a: 111). The phenomenal concept strategy corresponds to the latter: while we

might have two different kinds of concepts, they both refer to a monistic physical reality (Chalmers 2007: 170). So, the phenomenal concept strategist agrees that there is an explanatory gap between physical facts as referred to using physical concepts and experiential states as conceived through phenomenal concepts, though they argue that our possession of phenomenal concepts can be explained in wholly physical terms (Chalmers 2007: 167). The general aim of the strategy is to block anti-physicalist *ontological* conclusions.

In the case of Jackson, he has to face the phenomenal concept strategy because his argument involves the move from an epistemological claim to postulating additional ontological entities – non-physical properties or qualia. There is a clear ontological conclusion to Jackson’s argument that the phenomenal concept strategy aims to block. That is, it is possible to construe Jackson’s argument in purely epistemological terms: according to proponents of the phenomenal concept strategy, when Mary leaves the confines of her black-and-white room, she only acquires the relevant phenomenal *concept* of red – merely a different way to refer to the same physical reality that she already knew through nonphenomenal concepts. Nida-Rümelin and O’Conaill (2019) make this distinction as well: the stronger version of Jackson’s argument entails the ontological claim that there are non-physical properties, while the weaker version merely entails that there is non-physical knowledge, but not that there are non-physical facts, which is compatible with the phenomenal concept strategy<sup>4</sup> (and physicalism). As Luca Malatesti explains, the deflated form of Jackson’s argument amounts to the following scenario:

---

<sup>4</sup> The phenomenal concept strategy developed out of the new knowledge/old fact view (Fürst 2011: 66), according to which the phenomenal character of experiences, such as the redness we see when we look at a rose, is a *physical* property of experiences; and to gain knowledge of ‘what it is like’ to have an experience with a particular phenomenal character requires the subject to acquire the relevant phenomenal concept of that phenomenal character (Nida-Rümelin & O’Conaill 2019). The acquisition of that phenomenal concept can be fully described in broadly physical terms, even though the subject can only acquire that phenomenal concept if they had an experience of the relevant phenomenal kind (Nida-Rümelin & O’Conaill 2019). So, post-release Mary gains knowledge about phenomenal characters under phenomenal concepts but what makes that knowledge true are the physical facts that Mary *already knew* in her black-and-white room, though under another conceptualization (Nida-Rümelin & O’Conaill 2019).

“[W]hen Mary sees a coloured object, she acquires phenomenal concepts that co-refer with her scientific concepts that concern qualia. Thus, this reply to the knowledge argument requires that Mary’s new beliefs involve phenomenal concepts that are *a priori* detached from physical concepts. Moreover, these concepts to be possessed require having a relevant type of experience. Finally, her phenomenal concepts refer to physical properties.” (2011: 394)

Similarly, if Nagel’s bat argument involves the ontological claim that there are phenomenal *facts* that we can only know through experience (i.e. by taking up the relevant perspective) then this is more-or-less equivalent to Jackson’s knowledge argument. While the epistemological gap in Jackson’s case obtains between total physical knowledge and knowledge of non-physical qualitative properties, in Nagel’s proposal the gap obtains between total physical knowledge and knowledge of subjective facts. If this interpretation is correct, then Nagel also has to face the phenomenal concept strategy: there are no subjective facts, but only subjective concepts that we use differently from objective concepts to refer to the same ontologically monistic physical reality. If Nagel does not make an ontological claim of this kind, then it is unclear what his challenge to physicalism even is since there is no necessary expectation that physical theory should encompass every single type of subjective representation or perspective<sup>5</sup>.

---

<sup>5</sup> According to the phenomenal concept strategy, phenomenal concepts do not necessarily refer to non-physical facts, which is similar to the argument that experiences *are* representations: when one introspects an experience, one forms a second-order representation of the first-order representation of a physical fact (Lycan 1996). So, when Mary leaves the black-and-white room and sees red for the first time, she is able to correctly think to herself ‘this is what it is like to experience red’ since, before seeing red, she was unable to represent the first-order state of seeing red “by the introspective application of self-referential indexical concepts” (Mandik 2001: 184). In the case of Nagel, Pete Mandik (2001: 182) thinks that he is essentially making the same claim as Jackson, so the same strategy can be applied. Perspective-dependent facts are simply second-order representations of physical facts, and these second-order representations need not figure in a complete physicalist description of the world.

## 4. Zombies and Possible Worlds

It seems that knowledge-based arguments do not manage to conclusively refute physicalism. Physicalist counterarguments discussed so far offer both specific and general ways of avoiding anti-physicalist ontological conclusions. Because of that, I will now turn to the conceivability argument as presented by David Chalmers (1996, 2010), which offers a different approach to the anti-physicalist. Following an outline of Chalmers's proposal, I will examine some of the issues it faces, as well as his master argument against the phenomenal concept strategy. My aim is to show that the conceivability argument also fails to disprove physicalism and that it does not avoid the phenomenal concept strategy either.

### 4.1 Chalmers's Conceivability Argument

Relying on a jump from what we can conceive to what is possible, Chalmers invites us to think of beings who are completely like us – atom-for-atom copies – yet there is nothing it is like to be them, they do not experience the feelings of pleasure or pain, but they look, behave, and talk just like us. Those beings are aptly named *philosophical zombies*. Going even further, one could imagine an exact copy of our universe, where all the same microphysical facts obtain, yet phenomenal consciousness does not arise at any point – there is nothing it is like to be *anything* in that world. If we use  $P$  to denote the conjunction of all such microphysical truths about the universe and  $Q$  to denote any arbitrary phenomenal truth, such as that someone is conscious or instantiates some phenomenal property, then the most straightforward version of the conceivability argument goes as follows (Chalmers 2010: 142):

1.  $P \ \& \ \sim Q$  is conceivable.
2. If  $P \ \& \ \sim Q$  is conceivable,  $P \ \& \ \sim Q$  is metaphysically possible.
3. If  $P \ \& \ \sim Q$  is metaphysically possible, physicalism is false.

- 
4. Physicalism is false.

The first premise is the epistemic thesis that philosophical zombies or zombie worlds are conceivable (Chalmers 2010: 142). If we can imagine a sophisticated android that acts just like a human but feels nothing inside, then it is not a giant leap to imagine a human copy that exists in the same manner. The second premise is Chalmers's assumption that whatever is conceivable is metaphysically possible. It moves from the epistemic thesis into a modal thesis, about what is possible (Chalmers 2010: 142). This is the most hotly debated premise of the conceivability argument and Chalmers (2010) has written extensively on conceivability and possibility to defend it. The third premise follows from the second: if an exact physical copy of me can exist, or an exact copy of this universe but without consciousness, then microphysical truths do not necessitate the existence of phenomenal consciousness. This is the step from the modal thesis to a metaphysical thesis, about the nature of reality (Chalmers 2010: 142). The third premise is largely uncontroversial since it is widely accepted that physicalism at the very least entails a modal thesis, so if the exact same physical facts obtain in a copy of our world but that world is phenomenally different, then physical facts do not fix the phenomenal facts and physicalism is false (Chalmers 2010: 142–3).

The essence of the argument is that if a physically exact copy of our universe without phenomenal consciousness is possible in the primary sense, then it is also possible in the secondary sense, and if it is possible in the secondary sense, then physicalism is false (Chalmers 2010). This includes the jump from conceivability to possibility to the metaphysics of reality

(Chalmers 2010). Let us take the concept ‘water’ as an example. The primary intension of ‘water’ is ‘watery stuff’ since it is conceivable and logically possible that the concept of ‘water’ refers to something which is watery but *not* H<sub>2</sub>O in the actual world. The secondary intension of ‘water’ is ‘H<sub>2</sub>O’ and it picks out H<sub>2</sub>O in all possible worlds, making it *metaphysically necessary* that water is H<sub>2</sub>O (Chalmers 1996: 132).

To clarify, the idea behind this is that our phenomenal and physical concepts give us direct epistemic access to the essential nature of the properties they pick out, so they are in a sense *transparent* (Schroeter 2021). If we can conceive of zombies, they are actually possible since zombie-related concepts reveal actual possibility due to their transparency. That is, the conceivability of zombies is guaranteed by the primary intension, while their metaphysical possibility is guaranteed by the secondary intension and, if both phenomenal and physical concepts involved in the idea of a physically exact copy of our world without phenomenal consciousness are semantically stable, then the primary and secondary intensions will be true with regard to exactly the *same* world (Schroeter 2021). The point is that phenomenal concepts do not have a primary intension different from their secondary intension, so if one accepts the water example, one would be forced to apply the same principle to the mind-brain case, which could only mean that the secondary intensions of e.g., pain and c-fibres, are different. If the secondary intensions of pain and c-fibres are different, then those distinct concepts do not refer necessarily to the same ontologically monistic physical reality. So, if phenomenal concepts are accepted as such, the jump from conceivability to possibility goes through. That would mean that the physical and phenomenal concepts involved in comparing the actual world to the zombie world are revelatory with regard to the properties they pick out, which supports Chalmers’s conclusion that there are non-physical properties necessary to explain phenomenal consciousness, beyond the conjunction of all microphysical truths about the universe.

Still, physicalists argue that the conceivability argument only demonstrates the conceivability of zombies using the primary intensions of the concepts involved but not using the more relevant secondary intensions (Chalmers 1996: 132). If that is the case, then physical and phenomenal concepts “may pick out the same properties *a posteriori* despite the *a priori* distinction” (Chalmers 1996: 132). In response, Chalmers argues that even the primary intension alone is potentially sufficient to justify the conceivability argument:

“The property of being watery stuff is a perfectly reasonable property, even though it is not the same as the property of being H<sub>2</sub>O. If we can show that there are possible worlds that are physically identical to ours but in which the property introduced by the primary intension is lacking, then dualism will follow.” (1996: 132)

By analogy, if it is conceivable that there is a world without ‘watery stuff’, we would still establish that there is a world just like ours in which there is no H<sub>2</sub>O, so it is unclear why it would be any different with respect to consciousness, using only the primary intension (Chalmers 1996: 132–3). The notion of ‘watery stuff’ is sufficiently ‘accurate’, so to speak, to distinguish the actual and the no-H<sub>2</sub>O world.

Furthermore, Chalmers (2010: 150–1) argues that the physicalist either has to accept that zombies are metaphysically possible or accept Russellian monism, the view that “consciousness is closely tied to the intrinsic properties that serve as the categorical bases of microphysical dispositions” (Chalmers 2010: 151). If zombies are metaphysically possible, physicalism is false, as per the original argument. If zombies are not metaphysically possible, then it must be the case that the structural properties of physics in the actual world, combined with the intrinsic properties of physics, *do* necessitate the existence of consciousness, which amounts to Russellian monism (Chalmers 2010: 151). If Russellian monism is true, when we



conceive of zombies, we fix the structural properties of physical systems in the actual world, but we do not fix their intrinsic properties – and if those properties qualify as physical properties, then the zombies we conceive of are not full physical duplicates (Chalmers 1996: 152). Zombies so conceived (purely structurally) are then metaphysically impossible since the conception does not capture the full picture. However, Chalmers (2013) presented a conceivability argument against forms of Russellian monism as well. He argued that if  $PP$  &  $\sim Q$  is conceivable and possible, Russellian panpsychism<sup>6</sup> is false;  $PP$  being the conjunction of all microphysical and microphenomenal truths and  $Q$  being an arbitrary macrophenomenal truth. The goal for both the physicalist and the Russellian monist facing the conceivability argument should be establishing a necessitation relation between  $P$  or  $PP$  and  $Q$  or denying that conceivability indicates possibility. (Since I am not convinced that Chalmers’s proposal goes through, I will focus on other arguments for panpsychism in Chapter II.)

#### 4.2 Conceivability and Possibility

Ultimately, Chalmers’s conceivability argument relies on the jump from the epistemological notion of conceivability to the metaphysical notion of possibility. He has to defend the idea that a priori ideal coherence is an accurate guide to genuine metaphysical possibility, so that every epistemically possible scenario describes a genuine metaphysically possible world (Schroeter 2021). While most philosophers accept that zombies are conceivable but not necessarily metaphysically possible, some reject that we can even coherently conceive of zombies. For instance, Daniel Stoljar (2006) argues that zombies are not even genuinely conceivable but only seem conceivable because of our limited knowledge about physical facts. If we had complete knowledge of physical facts that pertain to phenomenal consciousness, then

---

<sup>6</sup> Or panprotopsyichism if it is a conjunction of all microphysical and all *protophenomenal* truths.

we would know how those physical facts result in conscious experience (Stoljar 2006). In response, David Papineau (2007b) argues that it is unclear how physical, non-experiential facts could give us knowledge of subjective and perspectival facts and render zombies inconceivable.

Others argue against the idea that conceivability is a guide for possibility. For example, Stephen Yablo (1993) distinguishes between ‘possible’ and ‘presented as possible’, arguing that conceiving involves the *appearance* of possibility, but that it is not an actual guide to what is metaphysically possible. Keith Frankish (2007) argues that the argumentative method itself is flawed and that it leads to contradictions, presenting a conceivability argument in support of *physicalism*: if anti-zombies – beings who are physical duplicates of us, where their physical properties are sufficient and responsible for the production of consciousness – are conceivable then they are possible (Frankish 2007). He tries to show that the conceivability to possibility nexus works both ways and that it does not necessarily imply any strong refutation of physicalism. Since the two conceivability arguments for physicalism and anti-physicalism cannot both be true, Frankish (2007) takes it as a strong indication that the conceivability argument is seriously flawed. It is now clear that Chalmers’s proposal is not as forceful as it initially appears.

The phenomenal concept strategist can argue that zombies are indeed conceivable, which is largely uncontroversial, but that the properties we conceive of under a phenomenal concept are also physical, which would block the move from epistemology to ontology. If the zombie world is an exact physical copy of our world, it would necessarily include phenomenal properties. As Brian Loar (1999: 467) states, the physicalist should request a justification of the assumption that conceptually distinct concepts *must* express metaphysically distinct properties. Loar (1999: 468) also argues that phenomenal concepts express the physical

properties they pick out and that this explains why a zombie world is conceivable, without there being a possible world in which the relevant physical properties are distinct from consciousness. If it can be shown that in the actual world phenomenal concepts refer to physical properties, then any exact copy of the actual world will include the phenomenal concepts that pick out physical properties.

### 4.3 Chalmers's Master Argument

Chalmers (2007) mounts a strong counter to the phenomenal concept strategy through his *master argument*, which is presented as a dilemma: the phenomenal concept strategy either fails to explain the explanatory gap *or* it fails to prevent the move from an epistemological to an ontological gap (Chalmers 2007: 173–4). If *C* denotes the thesis that humans possess phenomenal concepts, then type-B physicalists – those that accept that there is an epistemological gap between the physical and the phenomenal but no ontological gap – require: (a) for *C* to explain our epistemic situation with regard to consciousness, and (b) that *C* is explicable in physical terms (Papineau 2007a: 136). Chalmers argues that there is no version of the thesis *C* that satisfies both (a) and (b): either *C* is physically explicable or it helps explain our epistemic situation, but it cannot do both (Papineau 2007a: 136).

To further elaborate on the dilemma, Chalmers (2007) considers two scenarios, one in which  $P \ \& \ \sim C$  is conceivable and one in which it is not, with *P* being the complete physical truth about the universe and *C* the thesis as described above. If the physicalist says that it *is* conceivable then phenomenal concepts are conceptually independent of all physical claims, which means that the gap between the mind and the brain now reappears as a gap between phenomenal concepts and the brain, so *C* fails to be physically explicable (Papineau 2007a: 136). If the physicalist says that  $P \ \& \ \sim C$  is *not* conceivable then phenomenal concepts are not

conceptually independent of *P*, which means that zombies would be conceived of as having phenomenal concepts and not as lacking consciousness (which is how we usually conceive of them), so *C* fails to explain our epistemic situation (Papineau 2007a: 136–7).

David Papineau accepts both horns of the dilemma, claiming that both ways of thinking about phenomenal concepts allow them to be simultaneously physically explicable and explanatory of our epistemic situation: “we can conceive of phenomenal concepts in a way that makes them conceptually independent of the physical facts *and* conceive of them in a way that doesn't make them conceptually independent” (Papineau 2007a: 137). That is, we can think about first-order phenomenal concepts either *phenomenally*, using second-order phenomenal concepts about relevant phenomenal states, or *nonphenomenally*, using physical or functional terms (Papineau 2007a: 137). The first way of thinking about first-order phenomenal concepts means that they are a priori distinct from physical or functional concepts, which in turn means that *P* &  $\sim C$  is conceivable, while the second way of thinking about first-order phenomenal concepts makes *P* &  $\sim C$  inconceivable (Papineau 2007a: 137). These two ways of thinking still refer to the same thing: first-order phenomenal concepts, which will have the same nature and role no matter how they are referred to (Papineau 2007a: 137–8). Thus, the way they are referred to “ought to make no difference to whether they are physically explicable and explanatory of our epistemic situation” (Papineau 2007a: 138). With regard to the conceivability argument, this means that we can think of zombies “as sharing our physical/functional properties but as lacking our *phenomenal properties phenomenally conceived*” (Papineau 2007a: 141; emphasis added).

The first scenario of Chalmers's dilemma is equivalent to the conceivability argument, so it faces the same objections, primarily concerning the jump from conceivability to possibility. The second scenario can be addressed by the physicalist in the manner outlined by

Papineau (2007a). So, Chalmers's master argument does not manage to conclusively discredit the phenomenal concept strategy.

## 5. The Vagueness Argument

Vagueness arguments against physicalism are especially interesting with regard to the purposes of this work. As Jonathan A. Simon explains, the challenge to physicalism that arises from the vagueness argument is distinctive because “it is independent of the more problematic steps in arguments such as Jackson's Knowledge argument or Chalmers's Conceivability argument, and accordingly it is immune to many of the leading objections to those arguments” (2017: 2107–8). In addition, if the argument is sound, it might nudge us in the direction of panpsychism. So, after discussing objections and replies, I will argue that the vagueness argument is a better alternative to knowledge-based and conceivability arguments, both for confronting physicalism and motivating panpsychism.

### 5.1 Vagueness and Consciousness

Michael V. Antony discussed the vagueness argument in an influential paper, stating that “[i]f our concept *conscious state* is sharp rather than vague, then common versions of the identity theory, functionalism, and dualism are false” (2006a: 515; emphasis added). Those theories tie conscious states to neurophysiological states. The former concept is *sharp*: conscious experience either is present, or it is not, and “we can make no clear sense of the idea of a borderline phenomenally conscious state” (Tye 1996: 681). The latter concept is *vague*: if conscious states are tied to neurons, for example, then it is possible to imagine a borderline case of the ‘neuron’ concept, perhaps by removing atoms (or other sufficiently small parts)

from neuronal components, arriving at structures that are neither clearly neurons nor clearly *not* neurons (Antony 2006a: 522). As Antony argues, for any neurophysiological state N, it is “a near certainty that, regardless of what N is, a borderline N will eventually be reached” (2006a: 522). That is, there is no determinate point at which the brain switches from being in N to not being in N. In contrast, any given conscious state C either is or is not present. Antony’s (2006a: 527) argument can then be outlined as follows:

(I1) If  $N = C$ , then 'N' has borderline cases if and only if 'C' has borderline cases.

(I2) There are borderline cases of 'N' but no borderline cases of 'C'.

(I3)  $N \neq C$ .

This formulation is specifically aimed against the mind-brain identity theory, though he argues that it can be applied to functionalism and some forms of dualism as well. In the case of the mind-brain identity theory, the argument is straightforward: conscious states C have no borderline cases, while neurophysiological states N have borderline cases, so C cannot be identical to N (Antony 2006a: 527).

In the case of functionalism, Antony (2006a: 527) first claims that the concepts for inputs and outputs for any functional definition of a conscious state, whether the inputs are environmental or sensory and whether the outputs are bodily motions or behaviours, will be *vague*. Second, many functionalists require for causal relations to be realised only *ceteris paribus* or in normal or ideal conditions, which leaves much room for vagueness (Antony 2006a: 527). Third, as per Lewis (1972), if “a mental state is the type it is if *most* of the type’s proprietary relations are realized”, vagueness is introduced through ‘most’, which is “notoriously vague” (Antony 2006a: 527). Fourth and final, Antony (2006a: 528) states that a system realising a functionalist theory of consciousness does so by instantiating

neurophysiological properties, so the functional state F is tied to the neurophysiological state N. If we remove atoms from the brain to create a borderline case of N, then we will also arrive at a borderline case of F, where it is unclear whether the system realises F (Antony 2006a: 528). If that is the case, N can just be replaced by F in the argument outlined above.

If the argument succeeds, Antony (2006a: 530–3) argues that there are three unaffected options: eliminativism about conscious states, non-interactionist dualism, and theories that place consciousness at the level of fundamental physics, such as *panpsychism*. Eliminativism entails the claim that the concept ‘conscious state’ might be replaced by a vague future concept, so they are not committed to C in the argument (Antony 2006a: 532). However, eliminativism is a fairly unpopular view nowadays in debates on the philosophy of mind, so it will not be discussed further. The same goes for non-interactionist dualism, which does not tie C to N. The final option is especially significant for my aims since it indicates support for panpsychism:

“In seeking physical properties whose concepts are sharp, the obvious place to look is fundamental physics. Notice, however, that if the nature of consciousness resides at that level, the likelihood that panpsychism is true would appear to increase dramatically.” (Antony 2006a: 531)

We now have an argument for anti-physicalism that is independent of the discussion on Nagel, Jackson, and Chalmers, that relies on the very intuitive notion that the concept of ‘conscious state’ is sharp, and which suggests that one possible upshot is a theory that binds consciousness to fundamental physics, such as panpsychism. This perfectly leads to Chapter II, where I will introduce independent arguments for panpsychism, including Russellian monism, as well as Michael Tye’s (2021b) recent book *Vagueness and the Evolution of Consciousness* in which he explicitly argues for panpsychism on the basis of the vagueness argument.

## 5.2 Phenomenal Concepts are Vague

The intuitive force behind the vagueness argument is strong since the pre-theoretical notion of a conscious state seems to be binary: either consciousness is present, or it is not. However, David Papineau (2002: 203, 228) argues that this intuition is a remnant of dualistic thinking, where there is an *inner light* within us representing our consciousness as distinct from any physical properties, leading one to believe that there must always be a definite fact whether that light is switched on or off. Since Papineau (2002: 203) does not accept dualism, he sees this as just another case of a dualist intuition of distinctness distorting our thinking about phenomenal consciousness. Thus, he claims that both intuitively determinable phenomenal concepts, such as seeing something red or hearing a middle C, as well as the phenomenal concept of consciousness-as-such or what Antony (2006a, 2006b) simply calls ‘conscious state’, are vague (Papineau 2002: 178–9).

This vagueness lies not in our epistemological limitations but in the concepts themselves, and Papineau (2002: 179) argues that they are too vague to draw sharp lines once they are extended beyond their everyday range of application. While it is less controversial that the phenomenal concept of redness is vague (or in general that the *contents* of conscious states can be vague), the real problem lies with *consciousness-as-such*. Assuming that there is such a concept, Papineau (2002: 186–7) holds that it is possible to identify a physical property as the referent of the phenomenal concept of consciousness-as-such. He claims that there are several competing candidates for the physical nature of phenomenal consciousness, although there is no fact of the matter “as to which of these candidates our phenomenal concept of consciousness-as-such really latches on to” (Papineau 2002: 204).

One example he discusses are *higher-order theories of consciousness*, specifically higher-order judgments or “first-person phenomenal judgements made using phenomenal



concepts” (Papineau 2002: 205). According to Papineau, the phenomenal concept of consciousness-as-such is vague in more than one dimension. First, it is indeterminate whether the concept “refers to dispositional Higher-Order judgeability or to any of the other correlated properties which are similarly present whenever humans report themselves conscious and absent whenever they deny this” (Papineau 2002: 225). Second, if there is consciousness without higher-order thought, then it is difficult to exclude the possibility of hidden conscious states in humans, in which case “the phenomenal concept of consciousness-as-such seems to become indeterminate between all the many properties which are present whenever humans report themselves conscious, even if not absent whenever humans deny this” (Papineau 2002: 226).

On my understanding, to undermine the intuition that conscious states are sharp, Papineau does not necessarily need to provide a positive account of which specific physical property the concept refers to. Though that would certainly make for a more forceful claim, by demonstrating and giving examples of potential candidates, such as higher-order judgeability, Papineau already provides an alternative to pre-theoretical understanding. Moreover, as he asserts, the underlying problem behind the whole issue is the intuition that phenomenal properties are distinct from physical properties, which is why one finds it hard to accept that consciousness is a vague concept (Papineau 2002: 228). If there is nothing extra to the physical, no ‘inner light’, then there is no reason to think that there is always a precise fact of the matter about whether consciousness is present or not (Papineau 2002: 228). Rather, the point is to free ourselves of dualistic thinking (Papineau 2002: 228).

### 5.3 Replies to Papineau's Objection

In response, Antony (2006b: 477) defends the vagueness argument from Papineau's objection. He explains that a key premise in Papineau's argument is that the semantic properties of phenomenal concepts are determined by a naturalistic theory of content<sup>7</sup>, which is precisely what underpins the vagueness of phenomenal concepts. On Papineau's view, the vagueness of phenomenal concepts is "in a sense *external* to the psychological nature of such concepts [...] in that the vagueness, sharpness, or (in)determinacy of one's concept is entirely independent of the concept's psychological nature" (as reported by Antony 2006b: 478–9). That is a "highly idiosyncratic view", according to Antony (2006b: 479). In fact, Antony (2006b: 479) places the burden of proof on theories of content since they conflict with our pre-theoretical intuitions, and that is usually taken as a reason to re-evaluate the theory of content.

Furthermore, there is also the worry that semantic and teleological theories of content generate too many semantic values, and this applies to Papineau's objection as well since it must be shown for each physical candidate on Papineau's list, i.e. any relevant physical property, that it is correlated with the relevant phenomenal concept, as well as that "*only* the correlated material candidates so figure; and also that our sharp or otherwise determinate *non*-phenomenal concepts are not assigned a comparably large set of candidates" (Antony 2006b: 479). A sceptical attitude is justified, according to Antony (2006b: 479), since causal and teleological theories of content are known for generating too many semantic values. Thus, Antony (2006b) concludes that our current phenomenal concepts are not vague or indeterminate in any significant way, nor do they have "a sufficiently rich conceptual role,

---

<sup>7</sup> Theories of content, such as causal or teleological theories, aim to explain why conscious states have the contents that they have.

since they refer directly and are associated with no further conceptual descriptions” (2006b: 480).

In addition, Antony (2006b: 480) states that Papineau’s physical candidates only appear on the list because he sees them as contenders for the truth about consciousness, but this should not be relevant. That is, if our current concepts of consciousness are *not* vague or indeterminate, then nothing about the true nature of consciousness can change that (Antony 2006b: 480). If this is true, Papineau’s argument is a non-sequitur since his contenders are irrelevant with regard to the nature of phenomenal concepts (Antony 2006b: 480). The concept of consciousness-as-such is currently understood as a sharp concept, but Papineau can counter this by saying that a correct *future* concept will be vague or indeterminate (Antony 2006b: 481). For example, in *Philosophical Naturalism*, Papineau states the following:

“[I]f consciousness is like life [...] any physicalist account of consciousness is likely to make consciousness depend similarly on the possession of some kind of structural complexity [...] [y]et any such complexity is likely to come in degrees, with no clear cut-off point beyond which you definitely qualify as conscious, and before which you don’t.” (1993: 124)

However, as Antony (2006b: 481) counters, this line of thought is not open for Papineau’s phenomenal concepts since they might be hard-wired, genetically determined, otherwise immutable, so that they cannot change in response to theoretical pressures. Alternatively, even if they could change, those future concepts would no longer be *phenomenal* “since phenomenal concepts do not represent their referents as material” (Antony 2006b: 481). So, Antony (2006b: 481) concludes that such future concepts could not be both vague or indeterminate *and* phenomenal at the same time.

Furthermore, non-phenomenal concepts are not necessarily vague or indeterminate. Michael Tye (1996: 682) proposes the Crick-Koch hypothesis as an example: having phenomenal consciousness could be the same as having neurons oscillate at 40 MHz, so consciousness would be an all-or-nothing affair on this view. Papineau does not exclude such theories, nor does he provide a positive theory about the nature of phenomenal consciousness, so it is hard to see how he establishes its vagueness (Tye 1996: 682–3). Similarly, Antony (2006b: 481–2) identifies this as a particular form of materialism where only sufficiently complex material properties entail borderline cases, which would leave space open for the claim that a future non-phenomenal concept of consciousness will be vague. Still, this conclusion is as strong as the reasons to accept this form of materialism (Antony 2006b: 482). After all, what is the view that consciousness will be explained in fully physical terms, as well as that our current *folk* concepts pertaining to consciousness will be replaced by more accurate, scientific concepts in the future? It is *eliminativism*, accompanied by all the advantages and disadvantages that it faces, and Tye agrees when he states that Papineau “is an eliminativist with respect to [phenomenal consciousness]” (1996: 683).

Ultimately, I find Papineau’s (2002) argument that consciousness is vague unconvincing, and I side with proponents of the vagueness argument. Considering the strength of the intuition that consciousness-as-such is a sharp concept, as well as the fact that the vagueness argument is independent of previously discussed objections to physicalism while lending support to panpsychism, I believe that I am justified in taking it as a steppingstone for further discussion.

## Conclusion

I have discussed several ways of defining physicalism, Hempel's dilemma and the issues it raises for theory-based conceptions of the physical, standard arguments against physicalism as presented by Nagel, Jackson, and Chalmers, various physicalist counterarguments including the phenomenal concept strategy, Chalmers's master argument against it, as well as the vagueness argument which sees consciousness as a sharp and determinate concept that cannot be tied to vague and indeterminate physical concepts.

One reason for why panpsychism presents itself as an attractive topic for research is precisely because it does not necessarily rely on what might be considered the 'traditional' debate centred on physicalism and anti-physicalism. As was discussed, the vagueness argument suggests panpsychism as one of the options that can accommodate the strong intuition that consciousness is sharp and determinate because it ties consciousness to fundamental physical entities. This is especially intriguing considering that the phenomenal concept strategy is plausibly one of the strongest tools that physicalists have against anti-physicalists, if not *the* strongest. It is clearly a serious obstacle to the proposals of Nagel, Jackson, and Chalmers, in one form or another (though it might be unclear whether Nagel's approach leads to an anti-physicalist ontological conclusion in the first place). Assuming that the phenomenal concept strategy is successful in preventing their conclusions, then all that is left for the anti-physicalist, at least within the scope of this work, is the more robust vagueness argument and a more open pathway towards panpsychism.

So, both the negative conclusion that Nagel, Jackson, and Chalmers do not present a knockdown case against physicalism, and the positive conclusion that the vagueness argument is solid, independent, and compatible with panpsychism, will serve as the basis for how panpsychism will be discussed and defended in Chapter II, as well as for how it will be

addressed throughout this work, particularly in Chapter IV, where I will present my own formulation of the theory. In Chapter II, other independent arguments for panpsychism, stemming from Russellian monism, will also be discussed and compared to the vagueness argument to further strengthen the view. All this leads me to conclude that panpsychism is worthy of serious consideration, as a proposition with a potentially unique starting point.

## Chapter II: Arguments for Panpsychism

### Introduction

In Chapter I, I have looked at the traditional arguments against physicalism and examined common objections raised against them, concluding that they do not offer a clear notion of the non-physical – one that is truly ontological without the danger of collapsing into a purely epistemological notion – as well as other objections and strategies which weaken the anti-physicalist’s case. I have also argued that these arguments do not provide us with a good pathway towards panpsychism, with the theory usually treated as an appealing alternative to both physicalism and dualism (Goff & Coleman 2020) instead of being independently convincing. One option that did seem more promising was the vagueness argument, independent of standard objections to anti-physicalists and based on the strong intuition that consciousness is a sharp and determinate concept. Since the vagueness argument points towards panpsychism as a potential option that can accommodate this strong intuition, it will be explored here, as an independent route towards panpsychism.

However, first, I will discuss the standard Russellian monist argument, which is arguably one of the most popular and commonly discussed strategies for defending panpsychism, tracing its heritage to the ideas of Bertrand Russell (1927/1992). Second, I will examine specific objections to the Russellian monist framework, as well as replies that the Russellian monist can offer. Third, I will turn to Hedda Hassel Mørch’s (2018) proposal that tries to show that dispositionalism, which is usually presented as a counter to Russellian monism, can also be used to support panpsychism, giving the panpsychist more options to defend their claims. Fourth, I will outline Mørch’s (2018) proposal for unifying the standard Russellian monist argument for panpsychism and her dispositional argument, as well as

Dondoni's (2022) modification of her views which aims to avoid certain issues that Mørch's proposal faces and strengthen the overall case for panpsychism. All three options – the original argument, the dispositional argument, and the unified proposal – will be presented as potential defences of panpsychism. Fifth, I will compare Tye's (2021b) line of reasoning in support of panpsychism, based on the vagueness argument as discussed in Chapter I, to the original Russellian monist argument. Finally, in a slightly more speculative section, I will clarify my position on what was discussed throughout the chapter and discuss a further potential option in support of panpsychism. Ultimately, the goal of this chapter is to present argument for panpsychism independent from the common frameworks of Nagel, Jackson, and Chalmers, as they were presented in Chapter I, and show that the theory can be supported in many different, robust ways.

## **1. Russellian Monism**

In this section, I will start by presenting a detailed outline of Russellian monist commitments, drawing on both early and more recent sources. Then, I will discuss several objections raised against Russellian monism, which call into question its position on categorical and dispositional properties, its claims of integrating consciousness into our general picture of the world, and its overall consistency as an independent view. After that, I will examine Russellian monist replies to these objections, with the aim of demonstrating that they fail to conclusively refute the theory. The purpose of this section is also to present Russellian monism as an independent case for panpsychism, unrelated to the common anti-physicalist arguments and the objections that they face, as I discussed them in Chapter I.



## 1.1 The Standard Russellian Argument

The main motivation behind Russellian monism is the idea that physical theory only reveals the behaviour of matter without telling us anything about what matter is intrinsically. Goff (2017: 17) calls this the *positive component* of Russellian monism: physics only describes the relational, dispositional, extrinsic, or structural properties of matter, without saying anything about the underlying non-relational, categorical, intrinsic, or non-structural properties of matter – its ‘deep nature’. For example, physics tells us how an electron behaves, but not how it is in and of itself (Goff, Seager & Allen-Hermanson 2017). Earlier formulations of this line of thought stem from Bertrand Russell (1927/1992), who argued that observational science reveals the mathematical structure of the reality but not what reality is *intrinsically*, and Arthur Eddington (1928: 259), who claimed that physical theory cannot reveal the nature of the atom because it only attempts to describe it through ‘pointer readings on instrument dials’. This is the first step of the Russellian argument.

The second step involves the notion that this categorical nature either is consciousness or is somehow related to consciousness. The reasoning behind this is that consciousness, or something closely related to it, is the best candidate to play the categorical role since it is the only positive conception of a ‘deep nature’ we have available. Consciousness, as Seager (2006: 137) argues, is something we have ‘ready to hand’ to play the categorical role, and positing something else is unjustified and based solely on the demand that it has to be non-mental. In *Mortal Questions*, Nagel (1979) similarly argued that either consciousness is a fundamental and ubiquitous property of reality, or it emerges from non-conscious properties, which would lead us back to the same starting position and problems with regard to explaining consciousness. Goff calls this the *negative component* of Russellian monism:

“[T]he deep nature of matter transparently explains consciousness, in the sense that there is an a priori entailment from facts about the deep nature of matter to

facts about consciousness. If you were able to access the deep nature of my brain, you would in principle be able to deduce the nature of my consciousness. The mystery of consciousness, then, results from our scientific ignorance concerning the deep nature of matter.” (2017: 17–18)

This was also expressed in the work of Eddington (1928), who appealed to parsimony in claiming that it is more reasonable to tie the ‘pointer-readings’ of science (i.e. physical facts as described by physical theory) to a background of consciousness. As he argues, tying it to something else, which is inconsistent with thought, would be “silly” since it would leave us wondering “where the thought comes from” (Eddington 1928: 259).

Because of these considerations, Russellian monists believe that this is an elegant and parsimonious way of integrating consciousness into the physical world (Goff, Seager & Allen-Hermanson 2017). The following is a more detailed formulation of the Russellian monist argument, as discussed by Russell (1927/1992), Eddington (1928), Seager (2006), and Goff (2017):

1. Physical theories are cast in a *causal-structural language*. They reveal how matter is interrelated and how it behaves under certain conditions. The fundamental properties that we attribute to elementary particles, such as mass and electric charge, are described entirely in *relational* or *extrinsic* terms. For example, a mass of  $m$  is just a property such that something with it will obey the relation  $m = F/a$ , where  $F$  is a force and  $a$  is acceleration (Seager 2006).
2. Thus, physical theories say nothing about what matter is *intrinsically*, or *categorically*, in and of itself. Qualitative concepts are completely absent from modern physics (Goff 2017).

3. There is the intuition that reality cannot be wholly relational: there must be something intrinsic or categorical that underpins these relations, or else “all the things in the world will merely be each other’s washing” (Russell 1927/1992: 325).
4. We know of at least one such candidate: *consciousness*. As Seager (2006) argues, introspective awareness of consciousness reveals it as an inner, categorical, deep nature of matter.
5. The panpsychist thus argues that consciousness or protoconsciousness *must play the categorical role*. Postulating anything else would mean introducing an unknown entity into our ontology, so considerations of parsimony suggest against that move. As Goff (2017) argues, either we accept that the categorical nature of matter is phenomenal, or we accept that it is ‘we know not what’<sup>8</sup>.
6. Physical theories cannot provide *a complete account* of concrete reality, which follows from the first three premises (Goff 2017).
7. The upshot is either *panpsychism*, the view that all basic matter instantiates consciousness, or *panprotopsychism*, the view that non-phenomenal fundamental properties can account for phenomenal consciousness, from premises (3) to (5).

---

<sup>8</sup> Goff’s (2017) simplicity argument, echoing the philosophy of Eddington who has argued that it is unparsimonious to suppose that the nature of matter is incongruent with mentality and then wonder where mentality comes from (Eddington 1928: 259–60), can be introduced as auxiliary to premise (5): “in the absence of any reason to suppose otherwise, the most simple, elegant, parsimonious hypothesis is that the matter outside of brains is continuous with the matter of brains in also having a consciousness-involving nature”. Therefore, the onus is on the opponent of panpsychism to provide a candidate for the categorical nature role with enough reasons to prefer it to the panpsychist proposal, which is *prima facie* much simpler (Goff 2017).

Regarding (7), if the categorical nature just *is* consciousness, then we end up with panpsychism, the view that consciousness *qua* consciousness is fundamental and ubiquitous. If the intrinsic nature is proto-conscious, the result is panprotopsychoism, the view that non-phenomenal but explanatorily relevant properties – or protophenomenal properties – are fundamental and ubiquitous (Goff, Seager & Allen-Hermanson 2017). In this work, I will focus on panpsychism, so panprotopsychoism and protophenomenal properties will be scarcely discussed.<sup>9</sup>

Commonly, it is understood that the appeal of Russellian monism lies in its apparent success in avoiding the issues facing both dualism and physicalism. In terms of the former, it avoids the interaction problem or the problem of explaining how a completely immaterial and non-causal soul or mental substance can affect the physical, causally efficacious body. Such forms of dualism also violate the principle of the causal closure of the physical, which is the idea that all physical effects have only physical causes, and that consciousness cannot be a further additional cause. When panpsychists claim that consciousness is causally efficacious, they attempt to do so without violating the common account of causality, while substance dualism directly introduces *sui generis* mental causes. In terms of the latter, panpsychism is unphased by the common arguments raised against physicalism, which I discussed in Chapter I (though it still has to respond to the issues raised by the vagueness argument, which will be discussed later). The Russellian monist case for panpsychism (or panprotopsychoism) is also markedly independent from issues that standard anti-physicalist arguments suffer from, which is an advantage of the view. That is, it is possible to argue for panpsychism without relying on

---

<sup>9</sup> Briefly, the reason for this is that panprotopsychoism, in my opinion, faces the same explanatory gap that physicalists face since it purports to show that consciousness comes about from physical ultimates, such as quarks, that are somehow tied to consciousness yet are non-mental. They are *protophenomenal* and not *phenomenal*, so the view just seems less clear and more problematic when compared to the more straightforward and unified panpsychist proposal.

the proposals of Nagel, Jackson, and Chalmers, and ways for how to do that will be explored in the following discussion.

## 1.2 Objections to Russellian Monism

### *1.2.1 Dispositionalism and Dispositional Essences*

The first objection concerns the need for positing categorical properties. This is connected to the wider debate about the ontology of properties, specifically to the distinction between dispositionalism, categoricalism, and hybrid views of properties. The dispositionalist argues that the world can be wholly characterised in terms of dispositions – how something behaves under certain conditions. The most straightforward way of analysing dispositions in terms of counterfactual conditionals is the *simple conditional analysis* (Choi & Fara 2018):

An object is disposed to *M* when *C* if and only if it would *M* if it were the case that *C*.

A common example is mass: it is because of mass that particles have the disposition to exert a force of attraction on each other. The condition *C* is the two particles standing in a certain relation, while the manifestation *M* is the exertion of the force of attraction. The broader idea is that stimulus conditions and manifestations can be identified for any disposition, so that any disposition would exhibit its characteristic manifestation under the relevant stimulus condition (Choi & Fara 2018). There are various counterexamples to the simple conditional analysis, but for the purposes of this section, I believe it is sufficient. In the simplest possible terms, it is intuitive that the definition of a disposition rests on how the object possessing that disposition behaves under certain conditions, which is also how Russellian monists frame the issue when making the claim that science only ever describe what things *do*.

The dispositionalist can object to the Russellian monist by asking for a justification of the claim that dispositions have to be grounded in something categorical. The strongest forms of dispositionalism, such as dispositional monism or pan-dispositionalism, state that either all suitably qualified properties are essentially dispositional, for the former, or that all properties in the broadest sense of the term have dispositional essences, for the latter (Choi & Fara 2018). How does the Russellian monist refute these views, apart from the demand that dispositions must be based in something categorical? In addition, if the dispositionalist can give a positive account of a dispositional essence, then it is unclear why consciousness should be regarded as the only suitable candidate to play the role of the ‘deep nature’ of matter. For example, Brian Ellis (2013) and George Molnar (2003) argue that dispositional essences are *dimensions* pertaining to the spatial and temporal features of the property, which ground how the property is disposed to behave or how it is *oriented*, where it is *located*, and how it is *identified*. If this proposal is valid, it is unclear<sup>10</sup> why the Russellian monist insists that dispositions need a categorical basis, instead of being fully explained through dispositional essences.

Regarding this objection, an influential critique of dispositionalism comes from David Armstrong (1997), who argued that dispositions involve a reference to something that does not exist. Specifically, Armstrong draws an analogy between how a disposition still contains a reference to its manifestation and the intentionality of the mental, which is the characteristic of mental states to be about or to represent something; or simply the ‘aboutness’ of mental states (Jacob 2019). According to Armstrong (1997), this presents a problem for the dispositionalist physicalist since they argue that intentionality is different from dispositional orientedness and fully explicable via physical facts. However, since they admit such powers for dispositional

---

<sup>10</sup> More extremely, the dispositionalist can accuse the Russellian monist of begging the question against dispositionalism, though the Russellian monist can in turn rely on common arguments against dispositionalism, as will be explored.

properties, an intentionality-like relation is present in everything there is (Armstrong 1997: 79). Armstrong (1997: 79) further asks: how can an object's having of a property involve within itself a relation of any sort to the manifestation – a state which very often does not exist? He identifies this as *Meinongian metaphysics*, introduced by Alexius Meinong (1904), which includes non-existent objects into its ontology. Since this metaphysics is uncommon and usually seen as undesirable and unparsimonious, Armstrong presents this as an objection to dispositionalism.

Furthermore, he also argues that under dispositionalism, we arrive at a very counterintuitive picture of the world, where everything is a potency and “act is the mere shifting around of potencies” (Armstrong 1997: 80). In a purely dispositionalist ontology of properties, he says: “particulars would seem to be always re-packing their bags as they change their properties, yet never taking a journey from potency to act. For 'act', on this view, is no more than a different potency” (Armstrong 1997: 80). This could be seen as a form of a vicious circle in which the only existing things are endless chains of manifestations, without them being grounded in anything concrete, at least in the case of pure dispositionalism. On my understanding, this is a rather abstract picture of the world in which the behaviour of matter is all there is. There is certainly intuitive force against it.

### *1.2.2 Russellian Monist Properties and Causality*

The second objection attacks the idea that Russellian monism succeeds in integrating consciousness into the physical world. For example, Robert Howell (2015) claims that, on Russellian monism, a single type of physical disposition might have been grounded by a categorical nature or quiddity different from the one that actually grounds it, which means that this grounding relation is metaphysically contingent. Moreover, as Howell (2015: 32) argues,

a Russellian monist property has a phenomenal categorical ground and some causal dispositions, which means that there are two distinct aspects of such properties: one that grounds phenomenal resemblance relations, and one that grounds resemblances between causal profiles. If all physical events have sufficient causes due to the aspects that ground causal profile resemblances, then the phenomenal aspect of Russellian monist properties makes “no unique causal contribution to the physical world” (Howell 2015: 32). To clarify, the aspect grounding phenomenal resemblance relations is just the consciousness-involving categorical nature that Russellian monists postulate, while the aspect grounding causal profile resemblances just is a dispositional property (or a set of dispositional properties). Simply put, the argument is that the categorical nature is causally inefficacious, with all causality resting on the dispositional aspect.

Furthermore, even if the categorical natures somehow contribute to physical causation, consciousness does not necessarily inherit any physical efficacy from them, which is in contrast to the claim of Russellian monists (Robinson 2018). As William Robinson (2018: 102) argues, Russellian monism violates the *causal intuition* that sensations of certain kinds have effects on our behaviour in virtue of being sensations of relevant kinds. Russellian monists claim that physical structural relations are grounded in categorical natures, but these natures are not like the sensations that cause our behaviour, so it is unclear whether our sensations gain a causal role from the physical ultimates or ‘inscrutables’ (Robinson 2018: 107). That is, Robinson (2018: 106) states that there is already a complete causal explanation of behaviour available, in terms of physical elements, their structural relations, and laws of nature, so the addition of consciousness-related categorical properties to our ontology by the Russellian monist seems to be redundant and it does not have an explanatory contribution. Among other things, this is because the Russellian monist accepts the premise that the “contribution that an inscrutable



makes to the physical (structural) relations of a physical entity is independent of the relations in which that entity stands to other physical entities” (Robinson 2018: 109).

In response to this objection, Nino Kadić (2017) has indirectly argued against the premise that, in Russellian monism, the relationship between categorical and dispositional properties is contingent. Alter and Coleman (2020: 236–7) also explored this option, presenting the option for the Russellian monist to argue that dispositional and categorical properties are not modally separable. The Russellian monist can claim that, just like proponents of many other theories, such as some categoricists (notably Armstrong 1978, 1983), they can posit brute laws or necessitation relations that establish a relation between categorical and dispositional properties, as described in Russellian monism. An alternative proposal by Kadić (2017: 49–50) is that panpsychists can accept the identity theory of powers, according to which ‘categorical’ and ‘dispositional’ are merely ways of describing the same property, though this is beyond the scope of this section (for now; similar views will be explored later).

In contrast, some (Hawthorne 2002, Alter and Coleman 2020) have argued that causal relations do not generally hold with metaphysical necessity, so there is no need that the relationship between qualities and dispositions should hold with metaphysical necessity either. If the Russellian monist posits contingent phenomenal-dispositional grounding laws, they would still be compatible with the causal closure of the physical, giving the view an advantage over e.g., interactionist dualism, in which causal closure is rejected (Alter and Coleman 2020: 236). So, the view would still have a unique set of advantages and offer a novel approach to consciousness. It seems that in both cases, whether the relation between the two aspects of a Russellian monist property holds with necessity or contingency, the second objection does not conclusively refute the overall framework.

### *1.2.3 The Collapse of Russellian Monism*

The third and final objection was raised by Amy Kind (2015) who claims that Russellian monism is not a better alternative to either physicalism or dualism, but it is actually in danger of collapsing into one or the other, depending on whether consciousness is a fundamental component of the universe or not. Kind (2015) does not necessarily claim that Russellian monism is false but expresses pessimism about whether it can transcend the divide between dualism and physicalism. The dualist will insist that consciousness is inscrutable – a fundamental aspect of reality, while the physicalist will insist that inscrutables are non-phenomenal (Kind 2015). Depending on whether the Russellian monist endorses panpsychism or panprotopsychism, the view will collapse into dualism or idealism in the case of the former, and physicalism in the case of the latter. So, according to Kind (2015), while Russellian monism might sound as an alternative to the two traditional views, it is merely a reconceptualization or reinterpretation of the same age-old problems. It is unclear then why Russellian monism should be considered as progressing the debate with regard to the mind-body problem.

Alter and Coleman (2020) have argued that Kind's (2015) argument misconstrues Russellian monism since the theory does not purport to settle if consciousness is fundamental or not, but just to provide a framework to discuss that question – a framework different from the ones provided by physicalism or dualism. With regard to traditional dualism and traditional materialism, Alter and Coleman claim that the following is how Russellian monism can make a different contribution and what sets the view apart:

“By construing (proto)phenomenal properties as categorical grounds of physical dispositional properties, Russellian monism provides a framework for developing a view that has neither of those shortcomings: a view that

adequately integrates consciousness into nature without denying or distorting consciousness's distinctive features.” (2020: 239)

I also think that the danger that Russellian monism faces is not actually between dualism and physicalism, but between *idealism* and physicalism. That is, if all *concrete* reality is mental, it could be said that reality simply *is* mental. If all dispositions are grounded in an underlying categorical nature which is purely phenomenal, then reality at the fundamental level is wholly mental, which could be construed as a version of idealism. (Though, it is questionable whether panpsychism collapsing into a form of idealism is much of an issue, considering that the view is already considered counterintuitive.) So, whether or not Russellian monism collapses into one of the mentioned frameworks, it still constitutes a new way of thinking about consciousness. Specifically, it differs from traditional views by integrating consciousness into the natural, causal order “without disregarding or distorting consciousness's distinctive features”, in virtue of how it applies the dispositional/categorical distinction to the mind-body problem (Alter and Coleman 2020: 239).

## **2. A Unified Approach**

In previous sections, I hope to have demonstrated that Russellian monism has resources to reply to common objections. Still, dispositionalism is a very influential and popular view among philosophers, and the Russellian monist should offer a strong refutation of dispositionalism and defend their objection to it. However, if it can be shown that dispositionalism can also serve as a basis for panpsychism or that it directly *leads* towards panpsychism, the strength of the case for the theory will be significantly bolstered, and dispositionalism, as an objection to panpsychism, will lose much of its force. So, first, I will outline Mørch's argument from the experience of causation and her corresponding

dispositional argument for panpsychism. Second, I will describe the issues that Mørch's argument faces, as well as her attempt of resolving those issues by reconciling her argument with the original Russellian monist view. Finally, I will turn to Luca Dondoni's (2022) attempt at strengthening the case for panpsychism through a modification of Mørch's reconciliation attempt. The aim of this section is to further explicate the resources that panpsychists have available.

### 2.1 Mørch's Dispositional Argument

Hedda Hassel Mørch (2018) offers a strong argument for panpsychism based on dispositionalism. She argues that the only fundamentally dispositional properties that we can positively conceive of are phenomenal properties, in particular those associated with agency, intention, and/or motivation (Mørch 2018: 1073). Her argument from the experience of causation rests on: a) *non-reductionism*, or the claim that all physical things have causal powers, b) the already mentioned idea that the only causal powers whose nature we can positively conceive of are *mental* powers, and c) *non-skeptical realism*, which is the notion that the nature of the causal powers of physical things is knowable or positively conceivable (Mørch 2020: 276). In very simple terms, when I observe what I am doing, such as when raising my hand, I notice that the cause for the hand-raising is mental – my intention caused it. Why should I postulate that other causes are different<sup>11</sup>?

This view, if correct, leads to the notion that dispositionalism *entails* panpsychism or, at the very least, that dispositionalism can be used to present an argument for panpsychism. If so,

---

<sup>11</sup> This very basic line of reasoning, as I see it, is tied to the attitude of giving primacy to phenomenal consciousness when building an ontology. It seems more parsimonious to assume that all causes are of the same fundamental kind, collapsing the intentionality of the mental and the directedness of the dispositional into one unified category.

the force of dispositionalism as an objection or alternative to panpsychism will be weakened. Mørch starts from the original argument for Russellian panpsychism and then builds an analogous argument for panpsychism from dispositionalism. If the original argument is at least valid, then the dispositional one should be as well. Her proposal goes as follows (Mørch 2018: 1080):

The original arguments share the following two premises:

*Categoricalism*: All (or most) physical properties are realised by categorical properties.

*Mental Categoricity*: The only categorical properties we know or can positively conceive of are phenomenal properties.

Analogous premises can be constructed in terms of dispositional properties:

*Dispositionalism*: All (or most) physical properties are realised by fundamentally dispositional properties.

*Mental Dispositionality*: The only fundamentally dispositional properties we know or can positively conceive of are phenomenal properties.

The first part, about categoricalism, is equivalent to the Russellian argument: the world cannot be wholly dispositional, it must be grounded in something categorical, and the only categorical nature we know of is consciousness. The second part, about dispositionalism, is the claim that even if dispositionalism is true, the only dispositions we know of are of a phenomenal nature<sup>12</sup>.

---

<sup>12</sup> Interestingly, Karl Pfeifer (2016) entertains the possibility that the analogy between how dispositions are oriented towards or contain a reference to their manifestations and mental intentionality is no mere coincidence. He argues that most typical descriptions of intentionality fail to distinguish between mental states and physical, dispositional states, focusing especially on the notion of directedness. In a manner like Armstrong (1997), Pfeifer

The upshot is that both categoricalism and dispositionalism can be used to make an equally effective argument for panpsychism, so the dispositionalist objection to panpsychism does not go through. This alone is significant since it means that there are two very different and opposed pathways to panpsychism, though, as we will see later, Mørch tries to reconcile her dispositional argument and the original Russellian monist proposal.

Historically, this idea was discussed by many thinkers, most notably by Gottfried Wilhelm Leibniz:

“The clearest idea of active power comes to us from the mind. So active power occurs only in things which are analogous to minds” (Leibniz 1704/1981: 171, as reported by Mørch 2018)

William James proposed and later endorsed a similar argument:

“[T]he concrete perceptual flux, taken just as it comes, offers in our own activity-situations perfectly comprehensible instances of causal agency ... If we took these experiences as the type of what actual causation is, we should have to ascribe to cases of causation outside of our life, to physical cases also, an inwardly experiential nature. In other words, we should have to espouse a so called ‘panpsychic’ philosophy” (James 1911: 218, as reported by Mørch 2018)

---

argues that dispositions are directed or oriented towards something which does not exist, which he compares to intentionality being directed at unfulfilled states, using the example of desiring for something. Implicitly, he demonstrates the similarities between the two notions, which might lead to us doubting whether we need two and whether it would be more parsimonious to just have intentionality. Based on this, he argues for a form of panpsychism or, rather, panintentionalism, which he further develops into a pantheistic conception of God. An argument analogous to Mørch could be made: the only orientedness or directedness we know of and can positively conceive of is intentionality (e.g., desire), so it is more parsimonious to consider *all* orientedness or directedness as intentionality. This would both provide an argument for panpsychism and a more parsimonious alternative to postulating both intentionality and dispositional directedness; or an argument against reducing intentionality to directedness, since we would then lack an explanation of consciousness, as panpsychists argue.

Mørch (2018: 1087) argues that her notion of mental dispositionality is the only premise of her argument required to refute the challenge from dispositionalism. All other premises form part of the original Russellian monist case for panpsychism. So, while the standard Russellian monist argument has difficulties in refuting the challenge raised by dispositionalism, it can be countered by her line of thinking that leads to the conclusion that dispositionalism entails panpsychism. That is, the panpsychist can simply claim that dispositionalism also leads to panpsychism, so the dispositionalist case against Russellian monism is also weakened. After all, if their position leads to panpsychism as well, how exactly is it a refutation of panpsychism in general, including Russellian panpsychism?

## 2.2 Issues with Mørch's view

The dispositional version of the argument might seem to contradict the intention of the original categorical argument, and because of this it might not appeal to panpsychists as a way of strengthening their case. The first issue is that the idea that phenomenal properties are the only fundamentally dispositional properties we know of goes against the original idea that all physical properties are dispositional (Mørch 2018: 1080). This apparent contradiction disappears if we understand the original claim as the notion that “physical properties are dispositional in the non-fundamental sense of being reducible to relations between circumstances and manifestations”; as the notion that physics reveals only purely *relational* or *structural* properties but no intrinsic or non-relational properties “of either the categorical or fundamentally dispositional kind” (Mørch 2018: 1080). That is, the original claim does not necessarily specify what is fundamentally dispositional, so this objection can be avoided.

The second problem is the apparent contradiction between the view that phenomenal properties are dispositional and the view that phenomenal properties are categorical. The

solution that Mørch (2018) proposes is the identity view of powers (Martin & Heil 1999, Heil 2003, Strawson 2008), the idea that all properties are necessarily both dispositional and categorical. That is, those terms are merely different ways of conceptualising the same property. Accepting this view avoids the contradiction, though it renders Mørch's argument weaker since it now depends on the specific commitment to the identity theory of powers. (This will be addressed in greater detail in the next section.)

However, even if the identity view is unconvincing, the panpsychists who endorsed the original categorical argument “would have been wrong to say that phenomenal properties are categorical” but they would still “have been right to say that phenomenal properties are *intrinsic* (given that both categorical and fundamentally dispositional properties are intrinsic)” (Mørch 2018: 1080). Their error would then lie only in conflating the notions of categoricity and intrinsicity, which is “arguably not a very radical mistake” (Mørch 2018: 1080). So, she argues that proponents of the original argument could endorse her argument without contradiction, if they accept that the original argument contains ambiguity or a mistake with regard to conflating ‘categorical’ and ‘intrinsic’ (Mørch 2018: 1080). Still, I have presented the original Russellian monist argument using ‘categorical’ instead of ‘intrinsic’ precisely because I think that the notion of an intrinsic nature is less clear and less defensible than the notion of a categorical property, so whether this non-radical mistake really is a mistake, as Mørch says, is questionable.

Her conclusion is that some phenomenal properties are or can be conceived of as fundamentally dispositional – namely *intentional*, *volitional* and *motivational* properties (Mørch 2018: 1074), as well as that phenomenal properties are the *only* fundamentally dispositional properties we know of or have a positive conception of since “no physical or other non-phenomenal properties qualify” (Mørch 2018: 1081). Based on these considerations,



Mørch presents a strong case for panpsychism which dismantles common and pressing dispositionalist objections. Of course, picking just one of the two arguments would mean that the panpsychist would have to deal with the objections from the other side, as well as that they would have to build a case for the advantages of whatever option they pick. For Mørch, it is imperative to show that the two versions of the argument do not contradict each other, despite the already significant advantage of having two distinct routes towards panpsychism. So, some attempts to reconcile her dispositional argument with the original Russellian argument will now be discussed.

### 2.3 Reconciliation – Why Not Both?

Mørch offers an alternative that should appeal both to proponents of the standard Russellian argument and her dispositionalist argument for panpsychism by relying on the identity theory of powers, as discussed by C. B. Martin and John Heil (1999). It might seem that the Russellian monist argument and Mørch's argument are incompatible, *if* Russellian monism is understood as a form of categoricism, and *if* categorical and dispositional properties are opposites, as David Armstrong (1997) conceives of them since he claims that categorical properties are completely non-relational and non-dispositional (Dondoni 2022: 52). Standard Russellian monists would deny the view that mental properties are dispositional in their nature, which is what Mørch claims, while pure dispositionalists would simply reject the need for categorical properties, which is a part of the Russellian argument under this construal. Similarly, dispositionalists would reject that there is a need for a categorical aspect or part of a property, which is the claim that proponents of the identity theory of powers and of the compound view make (Dondoni 2022: 54).

In order to accept both arguments and to make the case for panpsychism stronger, Mørch (2018: 1079) subscribes to the view that all properties are necessarily both dispositional and categorical, as opposed to purely one or the other. This means that ‘categorical’ and ‘dispositional’ are just two ways of referring to the *same* property – the categorical and dispositional aspect, as well as the property itself, are *all* identical (Dondoni 2022: 54). So, if the panpsychist accepts the identity theory of powers, they can accept both that the only causal powers we know of or can positively conceive of are mental *and* that all physical objects have categorical properties (Dondoni 2022: 50). The argument goes as follows (Dondoni 2022: 50):

I. *Non-reductionism*: All physical things have causal powers.

1. *Categoricalism*: All physical things have categorical properties.

II. *Mental causation*: The only causal powers whose nature we can know, or positively conceive of, are mental.

2. *Mental categoricity*: The only categorical properties whose nature we can know, or positively conceive of, are mental properties.

III. / 3. *Non-sceptical realism*: The nature of the causal powers and of the categorical properties of physical things is knowable, or positively conceivable.

Therefore,

IV. / 4. *Panpsychism*: All physical things have mental properties.

The aim of this argument is a strong and unitary defence of panpsychism. The theory is often understood as a less problematic alternative to dualism and physicalism, not as a positive

proposal which is desirable in its own right (Maung 2019). The attempt at reconciliation thus widens and strengthens

“the resources that the panpsychist (who is already convinced by the traditional argument) has to support her view. Thus understood, the upshot of *Reconciliation* is essentially to give to the traditional panpsychist the possibility to subscribe to a new, strong argument to support her view, without forcing her to give up on the resources that she already has (namely, the argument from categorical properties) – thus producing a comprehensive case in *favour* of her view.” (Dondoni 2022: 51)

#### 2.4 Dondoni’s Solution

However, there is the danger that the identity theory of powers just collapses into pure dispositionalism, which would block the effort to reconcile the traditional categorical Russellian argument and Mørch's dispositional argument. According to Henry Taylor (2018), both pure dispositionalists and identity theorists about powers agree that properties are qualities. For example, Heil (2012: 59) characterises qualities as *the way things are*, while Martin (1996) characterises being qualitative as being *really there in the object*. Similarly, pure dispositionalists, such as Molnar (2003: 99), argue that properties are *really there in the object*. That is, both theories treat properties as *qualities* of the object.

Notably, both identity theorists about powers and pure dispositionalists reject David Armstrong's (1997) notion of categorical properties as complete non-relational opposites to dispositional properties. Pure dispositionalists reject this notion because they deny that there are non-dispositional properties, while the categorical aspect in the identity theory of powers – if it is truly identical to the dispositional aspect – must also be powerful. The object’s quality

of ‘being like this’ simply *is* its power ‘to do that’; the diamond’s hardness simply *is* its power to scratch glass, which simply *is* the diamond having a tetrahedral arrangement of carbon atoms (Jaworski 2016). Lastly, both pure dispositionalists and identity theorists believe that the whole nature of a property is *powerful* – “there is no aspect or part of a property which is not powerful” and there is “no part or aspect of a property that is independent of its dispositional nature” (Dondoni 2022: 53). Based on these considerations, Taylor argues that pure dispositionalism and the identity theory of powers are quite simply *the same view*.

In order to avoid Taylor’s objection, Dondoni (2022: 55) proposes that we replace the identity theory of powers with the *compound view*, which is the claim that a single property has separate purely dispositional and purely categorical parts *essentially*, but neither part alone completely exhausts the property. So, the main difference between the identity theory of powers and the compound view is that, under the compound view, categoricity and dispositionality are *not* identical and there is a part of the property that is truly non-dispositional, which would allow the panpsychist to hold both the standard Russellian categorical argument and Mørch’s dispositional argument *without contradiction* (Dondoni 2022: 56).

This is a very attractive proposal considering that Russellian monism can be interpreted as a version of the compound view in its original form, which I think is correct. This is what Howell (2015) does, without acknowledging it, when he presents his exclusion argument against Russellian monism. If this is the case, then the reasoning behind both the standard Russellian argument, as per Howell, and Mørch’s dispositional argument, as per Dondoni, neatly leads to the compound view. However, this opens the theory up to the exclusion argument, which renders the categorical part of the property epiphenomenal (Howell 2015: 32):

1. There are two distinct and separable aspects of Russellian monist properties (i.e. the compound view interpretation of the theory), laws that ground phenomenal resemblance relations and laws that ground resemblances between causal profiles,
2. All physical events have sufficient causes in virtue of those aspects that ground resemblances between the causal profiles of Russellian monist properties.

Therefore, those aspects that ground phenomenal resemblance relations make *no unique causal contribution to the physical world*, which leads to the idea that phenomenal natures are epiphenomenal. Note that this is the same objection that was discussed in Section 1.2.2, now applied Dondoni's (2022) unified approach. Nevertheless, whether the form of epiphenomenalism that ensues from Dondoni's argument is problematic can be doubted:

“In sum, the phenomenal part of the property is responsible for grounding the property, while the dispositional part of the property is responsible for activating the causal capacity of the property — thus both parts are essential to the property, and both parts play *equally relevant* roles (even though *different* roles) towards the contribution of the property to the structure of physical reality. Again, this does *not* mean that the phenomenal parts of properties are not epiphenomenal, for they are indeed (as they are causally inert); it means that this specific kind of epiphenomenalism is not a threatening one — phenomenal parts still contribute to the structure of physical reality (by grounding the properties they are parts of).” (Dondoni 2022: 62)

Furthermore, as was mentioned previously, a different strategy might be to postulate a necessitation relation between the categorical and the dispositional parts of the property (Kadić 2017) in order to avoid Howell's objection that two distinct categorical properties might play the same grounding role for dispositions, as well as to further alleviate the danger of

epiphenomenalism by connecting the two parts, as per the compound view, via necessity. This is an option that Dondoni (2022: 59–60) discusses yet rejects.

One other option is to deny the inference “from two different categorical properties playing the same grounding role” to the conclusion that they are epiphenomenal or not causally efficacious by accepting a Humean picture of reality, where causal regularities are understood as contingent (Dondoni 2022: 60). However, as Dondoni (2022: 60) notes, this would distance the compound theorist from both pure dispositionalism and dispositional essentialism since they generally reject Humean causality, as well as from Mørch's argument since she explicitly rejects a Humean account of the laws of nature (Mørch 2018, 2020). Still, Dondoni seems to favour this option and, overall, it is clear that the compound theorist has resources to either accept or challenge the accusation of epiphenomenalism.

### **3. Vagueness and Panpsychism**

Now, after I discussed Russellian monism, the dispositional alternative, and reconciliation attempts, I will turn to Michal Tye's (2021b) recent proposal, in which he directly argues for what he calls *representational panpsychism*, on the basis of the vagueness argument. He contrasts his view with Russellian monism, so I believe it is important to analyse it as yet another option the panpsychist has available to defend their theory. If Tye can present a version of panpsychism that avoids the problems faced by Russellian monism, the issues raised by the vagueness argument, as well as the issues faced by Nagel, Jackson, and Chalmers, as discussed in Chapter I, then it would certainly be one of the best options on offer. However, it is unclear whether his proposal successfully does so. The aim of this section is thus to explore Tye's view and how it connects to Russellian monism, as well as the unique challenges that it faces. This will set the basis for Chapter IV, in which I present a proposal that avoids both

concerns about vagueness and those facing Russellian monism, as well as those facing Tye's representational panpsychism.

### 3.1 Tye's New Argument

In his new book *Vagueness and the Evolution of Consciousness*, Tye entertains the idea that consciousness seems both vague and sharp at the same time. To explain this, he claims that while conscious *states* might be vague, consciousness itself – or consciousness\*, as he calls it – must be sharp (Tye 2021b: 2). Vague concepts are indeterminate, so it might be vague whether the colour I am seeing in my conscious experience is *this* shade of red or *that* shade of red, while sharp concepts are determinate or 'binary' – there either is conscious experience or there is not, with no borderline cases. Since Tye (2021b) posits that consciousness\* is a fundamental feature of micro-reality, tying it to determinably sharp physical ultimates, he avoids the objections raised by the vagueness argument, as discussed in Chapter I<sup>13</sup>. If this is correct, it would preserve both the intuition that consciousness is a determinate on-or-off phenomenon, as well as the conflicting intuition that our experiences come in degrees. So, while our representational content might be indeterminate, a determinate and sharp aspect of consciousness is guaranteed.

To further clarify the idea, he states that consciousness essentially has content, though what distinguishes it from other representational entities is not content but rather consciousness\*, which is not a representational property (Tye 2021b: 80). Human beings have both consciousness\* and consciousness – the former being the irreducible and fundamental,

---

<sup>13</sup> Another reason for why Tye introduces consciousness\* are zombies. Consciousness\* can explain why we are not zombies – beings who lack consciousness\* at the fundamental level – since “microphysical duplication does not logically guarantee duplication at the level of quiddities” (Tye 2021b: 89). So, zombies can act like us but they lack consciousness\* and are thus not conscious (Tye 2021b: 89). However, since I claimed that the conceivability argument is inconclusive in Chapter I, I will not focus on this aspect of Tye's proposal.

non-representational property of consciousness-as-such, the latter being the representational aspect of our experiences (Tye 2021b: 79). Fundamental entities, like quarks, only have consciousness\* but not consciousness, so they have no representational states at all (Tye 2021b: 79). Consciousness\* is also not an emergent property, but an intrinsic physical property of fundamental entities, and it is not a phenomenal property of token experiences either since then it would have a representational character (Tye 2021b: 80).

### 3.2 Clarifying Consciousness\*

So, consciousness\* is fundamental, irreducible, sharp, determinate, as well as non-representational and non-phenomenal. At this point, one might wonder whether this concept corresponds to any regular, pre-theoretical view of consciousness that we have. It might be tempting to think of it as *protoconsciousness*, a property which is not consciousness itself but is somehow responsible for the coming about of consciousness in complex systems. Indeed, Tye seems to suggest something along these lines in the following:

“[W]hat underwrites the transfer of consciousness\* from fundamental entities to certain complexes is their being organized functionally in the right way. It is because the fundamental parts are conscious\* and arranged in this way that the complex states so formed are conscious\*; further, in being conscious\* and playing the right role, these states are conscious.” (Tye 2021b: 88–9)

However, I believe that there is an alternative. After all, Tye named his view representational *panpsychism* rather than panprotopsychism. The better way to think about consciousness\*, in my opinion, is in terms of it being the condition for experience as such; the field of consciousness; minimal subjectivity; the conduit for experiences; an abstracted viewpoint. Notably, these options are all sharp. They do not have to be something representational, as Tye



says, but they can be *that* in which experiences happen. Admittedly, this is fairly abstract, though it better corresponds to Tye's aims. Similar views have been discussed in contemporary literature on phenomenology<sup>14</sup>. They would allow for consciousness to be both something sharp, that either is or is not present, and for it to be something that has vague representational content. In contrast, if consciousness\* is actually protoconsciousness, then Tye would inherit the problems of panprotopsychism, most significantly the question of how something non-conscious can lead to consciousness as we know it, which is the same explanatory gap that panpsychists accuse physicalists of having. The simpler solution is that quarks 'provide' us with consciousness\*, while the complex, such as the brain, provides us with consciousness as representational content. This is a crucial idea, which will be relevant to my proposal in Chapter IV.

### 3.3 Tye on Russellian Monism

Tye (2021b) compares his view to Russellian monism, arguing against both the *reductive* and *primitive* version of the view. The former is an identity thesis on which macrophenomenal states are identical with having micro-parts with the relevant set of phenomenal quiddities arranged to form a certain brain state (Tye 2021b: 24). This is equivalent to panpsychism. The latter is a grounding view on which macrophenomenal states are metaphysically grounded in having micro-parts with the relevant set of *proto*-phenomenal quiddities arranged to form a certain brain state (Tye 2021b: 24–5). This is equivalent to panprotopsychism. He claims that both options do not fare better than a posteriori physicalism,

---

<sup>14</sup> For example, see Dan Zahavi (2010, 2014) and Evan Thompson (2015), among others.

which identifies mental states with brain states, though the connection cannot be seen a priori (Tye 2021b: 28–9)<sup>15</sup>.

In the case of reductive Russellian monism, Tye (2021b: 25) argues that the physical types with which microphenomenal states are identified with are vague because neurophysiological types themselves are vague, and the same applies to relevant functional types, so if that is the case, it seems that Russellian monism does not have any advantage over physicalism since it suffers from the same type of objection. That is, it cannot accommodate the intuition that consciousness is a sharp and determinate concept. Furthermore, the quiddities or categorical properties postulated by the Russellian monist are generally very different from macrophenomenal states, “for it would be absurd to hold that quarks feel pains and itches and experience colors” (Tye 2021b: 26). How can the Russellian monist then claim that the presence of any given macrophenomenal state can be deduced a priori from the presence of a set of quiddities and their relations (Tye 2021b: 26)? While this might seem like the combination problem – the question of how complex consciousness comes about from simple consciousness at the fundamental level – Tye (2021b: 26) is asking rather why quiddities and their relations give rise to *that* macrophenomenal state rather than *this* macrophenomenal state. As he argues, the explanatory gap that is found here is very similar to the one that arises for standard physicalist views (Tye 2021b: 26). According to Tye (2021b: 28), Russellian monism effectively accepts that pain is one and the same as a certain brain state, while adding that the brain state is built out of ‘hidden combinations of hidden quiddities’, so it is unclear whether it

---

<sup>15</sup> Additionally, as a general objection to Russellian monism, he raises a form of a conceivability argument against the idea that phenomenal quiddities necessitate phenomenal consciousness (Tye 2021b: 31). Seeing that one of the motivations for Russellian monism is to preserve the jump from conceivability to possibility, this undermines the view (see also Chalmers 2013). Since I find conceivability arguments unconvincing, as I argued in Chapter I, I will not expand on this, though I mention it to better clarify Tye’s motivations.

avoids the issues raised by the vagueness argument and whether it has any advantages over a posteriori physicalism.

In the case of primitivist Russellian monism, Tye (2021b: 29) argues that it is ‘radically unclear’ how a brain state, which is complex, vague, and partly neurological or functional, could metaphysically ground the sharp type of consciousness. There is nothing that necessitates the sharp concept of consciousness arising from the set of quiddities and their relations (Tye 2021b: 29). Since metaphysical grounding is supposed to have an explanatory dimension, Tye (2021b: 29) claims that this is yet another explanatory gap that the primitivist Russellian monist must face, just like physicalists.

I do not find Tye’s arguments against the two kinds of Russellian monism convincing. Most fundamentally, I do not see why the Russellian panpsychist could not simply claim that the categorical property is what Tye calls consciousness\* and the dispositional property is connected to representational states and arrive at basically the same view that Tye endorses. Admittedly, this is not an obvious interpretation of Russellian monist views, so Tye could counter this. However, as we will see in Chapter III, just like Russellian monists, Tye must face the combination problem, including its subject-summing version, which is considered as the most serious and pressing objection to any (combinatory) form of panpsychism. Importantly, I agree with Tye’s distinction between consciousness\* and conscious/representational states, and I will apply a similar distinction when presenting my original version of panpsychism in Chapter IV. My goal will be to avoid both the standard issues that panpsychism faces, such as the combination problem, and the issues stemming from the vagueness argument.

## 4. Discussion

This is a speculative section in which I first offer my thoughts on Mørch's (2018) and Dondoni's (2022) reconciliation proposals. Then, I consider what a purely categorical formulation of panpsychism would be like, which is an option that is not commonly discussed in literature. The goal of this section is to identify the routes open towards a convincing argument for panpsychism, as well as to set up my line of thinking in Chapter III, where I will address the most pressing issue for panpsychism and, more importantly, in Chapter IV, where I will outline my own theory.

### 4.1 Thoughts on Reconciliation

While I find the reconciliation arguments convincing, I do not necessarily want to argue for reconciliation as the best option for the panpsychist. Rather, I would like to make a dialectical argument: whatever option one picks in terms of the ontology of properties, there will still be space for panpsychism to make its case. If one is a Russellian monist, the standard argument can be defended. If one is a dispositionalist, then Mørch's argument stands, and we can make a case for panpsychism. If one accepts Mørch's attempt at reconciliation between the categorical and the dispositional argument, then there is a way to clarify that view, as suggested by Dondoni (2022), and arrive at a unitary and convincing defence of panpsychism. Also, even if Mørch's reconciliation attempt collapses into dispositionalism, one can still argue for panpsychist from a purely dispositionalist standpoint, without seeking reconciliation. So, no matter which popular ontology of properties one chooses, one could make an argument for panpsychism. That being said, out of all these options, I found Dondoni's (2022) proposal to be the most complete and robust.

However, a problem that might undermine all of the reasoning behind the standard Russellian argument is the fact that dispositional essentialists introduce an alternative to consciousness as the only positively conceivable notion of an ‘inner nature’. Dispositional essences might be sufficient as a ‘basis’ for reality, without the need for fundamental consciousness, so the dispositionalist objection to Russellian monism is still forceful. Interestingly, in the case of dispositional panpsychists who follow Mørch's line of thinking, there is no danger because there is no obvious positive conception of causation which is *not* experienced as a mental phenomenon – at least provisionally. However, with regard to the standard Russellian argument, there seems to be an alternative to consciousness and it is thus no longer the only candidate that can play the grounding role. As mentioned previously, the *dimensions* discussed by Ellis (2013) and Molnar (2003), which orient, locate, and identify a property, are an example of such an alternative.

The problem is that those alternatives will not help us in explaining consciousness or filling out our ontology in order to make it more coherent. There is still a disconnect between theoretical entities like dimensions and spatiotemporal orientedness, and the first-hand experience of being conscious and understanding that as our intrinsic, inner, deep, or categorical nature. This is especially so when it comes to mental causation because it seems to directly be the basis of our actions, as per Mørch's argument. Because of that, I still think that the overall consciousness-oriented approach is preferable because it provides a more integrated picture of reality. Ultimately, the broadest justification for this project comes from the need to try an approach which puts consciousness first, as the epistemically and metaphysically most immediate object of our knowledge, as well as the phenomenon we know of with most certainty, making claims based on consciousness seem more promising than trying to replace it with a purely theoretical notion, like a dispositional essence. Quite simply, we *do not need* such a notion – we have consciousness ready at hand to play that role.

## 4.2 A Categorical Formulation of Panpsychism?

There is one remaining option for the panpsychist, which is a *purely* categorical formulation of panpsychism. I have not seen this extensively discussed in literature on the philosophy of mind. What would happen if, instead of assuming that dispositions must be grounded in something categorical, the Russellian monist simply denied that there are dispositions and accepted pure categoricalism? Under this view, there are only categorical properties and the theoretical roles that they play are fixed by the laws of nature. How consciousness at the fundamental level of reality forms more complex forms of consciousness would thus be fully determined by relevant laws of nature. Importantly, these theoretical roles are *contingent*. In another possible world, they could be different, so there is no modal necessity obtaining that fixes them in the same manner across all possible worlds. This is because categoricalism usually relies either on a regularity view, where regularities are necessary but not sufficient conditions of being laws of nature (Bird 2007: 68), or the nomic necessitation view as discussed by David Armstrong (1983: 85), where a necessitation relation can hold between universals. So, categoricalists separate the property and the theoretical role that it plays, while dispositionalists define a property in terms of the theoretical role that it plays and, as such, that property would necessarily play the same role in all possible worlds – because the property *is* the role. Whether the categoricalist sees laws of nature as regularities or as involving necessitation relations, the theoretical roles that the laws determine are contingent in both cases.

Why is this important? First, it gives the panpsychist another option dissimilar both to Mørch's view and the standard Russellian argument, if Russellian monism is understood not as standard categoricalism, but as endorsing a hybrid view in which there are both categorical and dispositional properties (as discussed by Mørch 2018, Howell 2015, and Dondoni 2022). Also, it potentially avoids the vagueness argument if it is compatible with the claim that consciousness\* is tied to the fundamental level of reality as a categorical property, leaving

representational consciousness and how it relates to consciousness\* to the laws of nature, which does not seem to be critically different from Tye's representational panpsychism. Moreover, in Chapter III, the categoricalist formulation of panpsychism will be examined as a possible solution to the combination problem. In Chapter IV, categorical panpsychism will also be addressed as a possible way of strengthening my proposal. Indeed, all ontologies of properties that have been discussed so far will be relevant to how I will formulate my original theory later on. For now, it is enough to simply acknowledge this as a possible option for the panpsychist.

## **Conclusion**

The point of this chapter was to show how one can make the case for panpsychism based on several different ontologies of properties. There is the original Russellian argument, which claims that physical theory only ever describes the behaviour of matter and proposes consciousness as the most robust candidate to play the grounding categorical nature role. Then, there is Mørch's (2018) dispositional version of the view, aimed at deflating the strength of the dispositionalist counterproposal to panpsychism. What strengthens the case for the theory even further are the two reconciliation proposals: Mørch's reconciliation attempt based on the identity theory of powers, aimed at appealing to both Russellian monists and dispositionalists, and Dondoni's (2022) re-imagination of her view in terms of the compound view, with the goal of solving some of the issues that her proposal suffers from. I have also discussed Tye's (2021b) representational panpsychism, based on the vagueness argument, and how it relates to Russellian monism. His distinction between consciousness\* and conscious states will be useful in Chapter IV since I make a somewhat similar proposal. Then, I have examined what a purely categorical formulation of panpsychism would look like and what its unique advantages would

be. Finally, I stated that, instead of firmly accepting any particular approach, I tried to show that all of these views offer a way to make an argument for panpsychism convincingly.

What I did not address is the most pressing issue for any form of panpsychism, the combination problem. So, in Chapter III, I will discuss the combination problem at great length, specifically its most pressing version pertaining to the combination of subjects. The quality combination problem, or the question of how qualities combine to produce the rich complexity of human mental life, will also be addressed. I will also look at some of the most promising attempts at solving or avoiding the problem, including Tye's (2021b) proposed solution and my suggestion for a purely categoricist formulation of panpsychism, and see whether they are successful. All of this will define how I present my novel formulation of panpsychism in Chapter IV, which avoids both concerns about vagueness and those facing Russellian monism.



## Chapter III: The Combination Problem for Panpsychism

### Introduction

In the previous chapter, I have examined various arguments for panpsychism and the objections that they are faced with. I have concluded that standard arguments against physicalism are not good bases for establishing panpsychism, arguing instead for the primacy of independent arguments for the theory. However, all forms of panpsychism relevant to the purposes of this work have to provide some sort of an answer to the combination problem, the question of how consciousness can combine, either in the form of a solution or an evasion attempt.

So, in this chapter, I will first extensively discuss what the combination problem is and how it is usually presented, what the different kinds of the problem are, with a special focus on its distinctively problematic subject-summing formulation, which asks how small subjects can combine to form larger subjects, such as humans. Second, I will analyse various possible answers to the combination problem from the literature, as well as objections and replies. I will start from Goff's phenomenal bonding solution, which argues that microsubjects can stand in relations that necessitate the coming about of a further subject. After that, I will outline Seager's combinatorial infusion proposal, according to which microsubjects can fuse together to form a new subject. The next proposal examined will be Tye's representational panpsychism and its relation to the combination problem. Cosmopsychism, an alternative theory that is unburdened by the standard forms of the combination problem, will also be addressed. Third, I turn to broader strategies that the panpsychist might endorse in order to provide an answer to the combination problem. The first one will centre on a discussion regarding the unity of consciousness, while the second one will centre on putative cases of disunified consciousness. Finally, I will offer an overview of the discussed proposals and explain what my position is with regard to them, which ones I found more convincing than others, and how that will inform

my original proposal in Chapter IV. Briefly, I will also entertain the option of a purely categorical formulation of panpsychism, which I discussed in Chapter II, in relation to the combination problem, seeing whether it can offer a way out for the panpsychist.

Ultimately, I will conclude that, despite the various proposals defending panpsychism against the combination problem, the strength of its challenge is preserved. Specifically, I will side with those who argue that it is positively inconceivable how subjects might combine, which will serve as a limit for what I am allowed to argue for in the next chapter. While I will not, in most cases, directly argue that the panpsychist cannot solve the combination problem, I believe that it is safer (and more interesting) to assume that that is the case and see what sort of panpsychism can arise from such constraints.

## **1. The Combination Problem**

In this section, I will first outline the general form of the combination problem, considering its commitments and explicating how it presents an objection against panpsychism. I will then describe the various kinds of combination problems that the panpsychist must face, related to subjects, qualities, and structure. Then, I will expand on the subject-summing version of the combination problem and explain why it is the most pressing issue for panpsychists. The aim of this section is thus to analyse the combination problem in great depth, to better understand attempts to solve or avoid it, which will be discussed in the following sections. In addition, it is also important to understand the challenge it presents to panpsychism since that will limit and define how I avoid the combination problem with my original proposal in the next chapter.

## 1.1 The General Idea

The main motivation for postulating basic or fundamental consciousness is the goal of making it more intelligible how consciousness emerges at higher levels. The combination problem for panpsychism<sup>16</sup> is the question of how these experiences, subjects or “feelings” at the lowest level of reality come together to form more complex higher-level experiences, subjects or “feelings”. Most famously, it was formulated by William James in *The Principles of Psychology* (1890). James criticised what he called the “mind-dust” theory, according to which mental states are understood as combinations of elemental mental states (Chalmers 2017). It is also important to note that James’s broader argument was aimed against the possibility of intrinsically organised wholes of any sort (Shani 2010). In fact, James approvingly quoted Royce in the *Principles of Psychology*: “No summing of parts can make a unity of a mass of discrete constituents, unless this unity exists for some other subject, not for the mass itself” (Royce 1881: 375–76). Itay Shani (2010: 419) identified three main theses within the broader paragraph quoted by James, all of which play an important role in his argument against mental combination:

1. All complex natural systems are aggregates
2. Such aggregates do not constitute unified wholes
3. Aggregates may act as unities by exerting combined influence on an external entity, such as when appearing as unities to an observer.

Importantly, in the third thesis, aggregates appearing as unities are still intrinsically non-unified: “For the spectator it is one; in and for itself it is an aggregate“ (Royce 1881: 375–76, as reported by James 1890). Thus, James endorsed a form of mereological nihilism, where the

---

<sup>16</sup> Strictly speaking, the combination problem is only an issue for combinatory forms of panpsychism, where elementary units of experience merge together, in an aggregative manner, to yield higher forms of experience (Seager 1995, Shani 2010).

existence of composite objects is denied, and from which he drew the conclusion that no mental combinations exist either (Goff 2016). He wrote:

“Where the elemental units are supposed to be feelings, the case is in no wise altered. Take a hundred of them, shuffle them and pack them as close together as you can (whatever that may mean); still each remains the same feeling it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean. There would be a hundred-and-first feeling there, if, when a group or series of such feelings were set up, a consciousness belonging to the group as such should emerge. And this 101st feeling would be a totally new fact; the 100 original feelings might, by a curious physical law, be a signal for its creation, when they came together; but they would have no substantial identity with it, nor it with them, and one could never deduce the one from the others, or (in any intelligible sense) say that they evolved it.” (James 1890: 160)

The idea here is that “private minds do not agglomerate into a higher compound mind” (James 1890: 160). It should be noted that ‘private’ might introduce additional constraints, though I believe that James did not use the word in the sense in which philosophers of mind discuss the privacy of the mental nowadays. Since the argument can be rephrased without reference to privacy, I will ignore it. The focus should be on the main idea that minds do not agglomerate since it is very difficult – perhaps unintelligible – to imagine how a group of phenomenal simples could unite in such a way that they produce new and distinct higher-level phenomenal entities. Considering that we have no problems with the idea of bricks forming a house, the combination problem pertains specifically to the mental realm (Goff 2017).

Now to further specify James’s argument, we need to elucidate two different ways of reading it. On the one hand, it could be taken as a more abstract argument against the idea of

qualitative elements or “feelings” combining into a unified macroexperience, without reference to the experience being enjoyed by a subject (Coleman 2012, 2014). That is to say, the experience is then understood as a “mere qualitative element” of one’s overall experience, where the notion is abstracted from the existence of the subject (Coleman 2014: 27). This is merely a way of indexing the particular qualitative elements or token experiences without referring to the subject.

On the other hand, the word “feelings” could be interpreted as denoting subjects considering that James uses “a feeling” and “a consciousness” interchangeably (Coleman 2014: 27). The notion of *subject* relevant here is that of a *phenomenal* subject, a centre of experience, something for which there is an answer to the question of “what it is like to be it?”, rather than the notion which defines subjects in terms of agency, function, or behaviour. This distinction between feelings as qualitative elements and feelings as subjects is introduced to extrapolate in more detail how James’s argument can be understood, as well as to indicate which interpretation is pertinent to the debate at hand. While the first reading might be relevant to certain forms of the combination problem focusing on the compounding of qualities, it is the second reading of James’s passage that most philosophers nowadays consider to be more challenging for panpsychism. To further clarify this, it is important to explain what other kinds of the combination problem exist in order to precisely see why the subject-summing variant is the most difficult one. David Chalmers (2017) introduced a thorough and useful taxonomy of the various kinds, with three main problems and several specific sub-problems. The three main ones are the subject combination problem, the quality combination problem, and the structure combination problem<sup>17</sup> (Chalmers 2017: 182). All of them present a different challenge to panpsychism, though they are commonly intertwined, as are the views presented to solve them.

---

<sup>17</sup> There are some other versions of the problem, such as the unity, boundary, awareness and grain problems, but they are plausibly subsumable under the three main variants (Chalmers 2017).

As was noted, the subject combination problem is the general question of how microsubjects combine to yield macrosubjects. The specific and most difficult form of it is the *subject-summing problem*, the reasoning that it seems possible in principle for any set of subjects to exist without a further subject existing, which means that no amalgamation of microsubjects *necessitates* the existence of a macrosubject (Chalmers 2017). Put simply, it always seems possible for *any* group of subjects to exist without a further subject existing. This means that no combination of microsubjects can be said to necessitate the existence of a macrosubject (Chalmers 2017). There are also two ways of responding to the subject-summing problem. For some, the problem is simply difficult to solve but not impossible. Eventually, the hope is that panpsychist will either provide a solution or otherwise circumnavigate the problem (Goff, Seager & Allen-Hermanson 2017). For others, subject-summing is taken to be incoherent or impossible. For instance, as will be explored in more detail later, Coleman (2014) has argued that it is incoherent because of the fact that each subject's viewpoint excludes the viewpoints of all other subjects. Coleman's viewpoint argument will be addressed in more detail in the next section.

The quality combination problem is the issue of how microqualities combine to yield macroqualities. A macroquality is something like phenomenal redness (what is it like to see red), for example (Chalmers 2017: 183). The specific difficult version of the quality combination problem is the *palette problem*: there is presumably a limited set of microqualities, so how do they produce the wide spectrum of macroqualities (Chalmers 2017: 183)? This becomes especially salient when we imagine the supposedly few microqualities producing all of the macroqualities of sight, smell, taste, etc., but also the experiential macroqualities of emotions, moods, and maybe even thoughts in general (for the latter, see Bayne and Montague 2011). The idea that there is a limited set of microqualities, however, is merely an assumption which is strongly suggested only by Russellian panpsychism, where microqualities must

correspond to the known small number of fundamental microphysical properties (Chalmers 2017: 183). This problem disappears if there is no reason to think that there is a limited set of either microqualities or microphysical properties.

The structure combination problem entails the palette problem: how does a presumably limited microexperiential structure produce a rich and broad macroexperiential structure (Chalmers 2017: 183)? To illustrate what ‘structure’ means here, consider how our visual experience corresponds to the experienced spatial environment (Goff, Seager and Allen-Hermanson 2017). For example, when I see a chair, the structure of my visual experience of the chair will correspond to where the chair is spatially located. Contrarily, the structure of this experience does not correspond to the structure of our brain: “Macrophysical structure [...] seems entirely different from the macrophenomenal structure we experience” (Chalmers 2017: 183). The specific form this problem takes is the *structural mismatch problem*: the macrophysical structure of human brains appears completely different from the macrophenomenal structure of conscious experience, so how do the microphenomenal elements whose structure presumably corresponds to microphysical structure combine to yield a macrophenomenal instead of a macrophysical structure (Chalmers 2017: 183)?

The panpsychist must address all of these problems in an ultimate solution, but the subject-summing problem is commonly taken as the most difficult one to solve. The reason for this is that one can more easily imagine a solution to the other problems. On this, I agree with Coleman (2014), who argues that there is no particular challenge in the idea of qualitative instances combining into qualitative wholes, where the result is “the intelligible product of the qualities of the components plus their arrangement” (2014: 28), while the combining of subjects strikes him as a “demonstrably incoherent notion” (2014: 29). One of the mechanisms for that could be simple addition or some form of merging, fusion... etc. In the case of subjects,

the additional challenge lies in explaining how a *further* subject emerges, one which has an utterly new perspective.

To demonstrate the difficulty of this problem, Goff (2009, 2017), following James's original passage, invokes a form of the conceivability argument: for any group of subjects and any conscious states, there could be a set of subjects instantiating a set of conscious states without there being a further subject that is not identical with any member of the original set. Put in simpler words, this means that one can conceive of a "micro-experiential zombie", who is: a) completely physically identical to an actual human being, b) its micro-level parts have conscious experience, but c) none of their macro-level parts has conscious experience (Goff 2017). The upshot is that facts about microsubjects do not necessitate any facts about macrosubjects, nothing that would metaphysically ensure the coming about of a new subject (Goff 2017). In other words, it is possible for the facts about microsubjects to obtain without facts about macrosubjects obtaining, so that the existence of the former cannot account for the existence of the latter (Goff 2017). Paired with Coleman's (2014) conclusion that subject-summing is demonstrably incoherent, I believe that the gravity of this version of the combination problem has been shown. Since the same form of the conceivability argument is generally used to reject physicalism, then it is a great worry for the proponent of panpsychism if the same argument applies to them: they have no edge over physicalism in terms of explaining phenomenal consciousness, which was the main motivation for the theory in the first place. As I have argued in Chapter I, I do not generally think that conceivability arguments are convincing. However, in this case, Goff's formulation merely attempts to demonstrate the difficulty of imagining subjects combining into new subjects. The better way to state it, in my opinion, is merely that a set of microsubjects does not necessitate the existence of a further, more complex subject, in the same way that a set of physical ultimates does not seem to necessitate it, so an explanation for how that occurs must be given in both cases. So, his



argument can be rephrased as a more general objection, independent of the conceivability-possibility nexus.

Part of the explanation that the panpsychist provides relies on them introducing consciousness at the lowest level of reality, with the aim of making it more intelligible how consciousness emerges at higher levels. If consciousness is present throughout, they argue, then it is easier to see how it might emerge, compared to the physicalist where the emergence is from non-phenomenal simples to complex phenomenal wholes. Whether this coming about of consciousness should be called emergence under panpsychism is debatable but made clearer by distinguishing between weak and strong or radical emergence. In cases of weak emergence, complex properties found at the higher level are wholly explainable in terms of the behaviour of the lowest level. A commonly used example is liquidity, which at first seems like a property not found in the molecules that comprise water, for instance, yet is actually merely a particular molecular structure where molecules are not bound but slide past each other (Strawson 2006: 13). Nothing new comes into being when that molecular structure obtains; the weak emergence of liquidity is traceable, so to say, to the goings-on at the lowest level (Strawson 2006: 13).

In the case of strong or radical emergence, the higher-level property is ontologically distinct from the lowest level, and it cannot be traced to the facts obtaining at the lowest level. Something ontologically new comes into being in the case of strong or radical emergence. The panpsychist introduces consciousness at the lowest level to avoid radical emergence and instead replace it with a gradual, weak form of emergence. But, as Goff (2017) attempts to show, the panpsychist loses this advantage if it is unclear whether the relationship between phenomenal simples and macroexperiences is in any way more transparent than the relationship between purely physical simples and macroexperiences (Morris 2017). That is, it is unclear what explanation the panpsychist can then offer for how consciousness comes about at higher levels, and what merit it would have.

Goff's argument that the presence of microsubjects or phenomenal simples does not necessitate the coming about of macrosubjects or complex phenomenal consciousness can also be applied to elucidate the previously mentioned intuition that it is difficult to *imagine* – or even that it is *unintelligible* – how subjects could combine. Of course, it is difficult to imagine a lot of things that might be intelligible, such as chiliagons, so that hardly constitutes a serious argument against combination. The key issue here is intelligibility. The panpsychist claims that the non-reductive physicalist's combination of accepting realism about phenomenal consciousness paired with accepting non-phenomenal microphysical simples results in an unintelligible form of brute emergence, where something ontologically new at the macrophenomenal level emerges from a non-phenomenal micro-level. As Galen Strawson puts it, such emergence involves “magic passage across the experiential/non-experiential divide, something that, ex hypothesi, not even God can understand” (2006: 24). This attack based on intelligibility is rather ferocious. Also, Goff (2017) argues, the same form of the gap is present in both the physicalist and the panpsychist case when scrutinised: both cases involve an as-of-yet unexplained “magic passage”, neither of which is more transparent at first sight (Morris 2017). So, it is unclear that the panpsychist has an advantage here.

Considering the above, it seems convincing that the subject-summing problem is indeed the most difficult obstacle to a coherent panpsychist worldview. However, not enough has been said, so in the next section, I will explore in more detail the precise difficulties pertaining to subject summing.

## 1.2 The Summing of Subjects

The key issue with subject-summing is the idea that for any group of subjects and for any group of conscious experiences, it is conceivable that each subject instantiates some token

conscious experience, but it is not the case that there is another, further subject necessitated by the preceding group of subjects that is not identical with any of them (Goff 2017: 174). Based on this, Goff presents an argument similar to Chalmers's (2013) conceivability argument, as discussed in Chapter I. He argues that we can conceive of an atom-for-atom copy of a human being, whose micro-level parts all have conscious experience, but whose macro-level parts lack conscious experience altogether (Goff 2017: 175). However, as I mentioned before, this same objection can be rephrased without reference to conceivability and possibility: it is merely the claim that there is no obvious necessitation relation between phenomenal simples and macrophenomenal states. So, this serves to demonstrate that consciousness at the micro-level is not a necessary and sufficient condition to produce macro-level consciousness. For how would a group of subjects necessitate a further subject if the further subject is not identifiable with the preceding subjects? If S1 is a subject having a particular conscious experience and S2 another subject having a particular conscious experience, it does not follow from that that there is another subject, S3, having both of these experiences unless we say that  $S1=S2=S3$ . In addition, with regard to the unity of consciousness, if there is something it is like for S1 to have a particular experience, such as seeing red, and something it is like for S2 to have a different particular experience, such as seeing green, it does not follow that there is something it is like for a subject to experience both. Of course, one subject can see red and green at the same time, by looking at a red and green spot on a wall, for instance, but that is not the same as saying that a subject can integrate the experiences of two other subjects in order to experience red and green at the same time as a *new* subject.

To clarify this, let us look at Sam Coleman's phenomenal point of view argument. According to Coleman, the combination of microsubjects into macrosubjects is a "demonstrably incoherent notion" (2014: 29). It is not just that it is unintelligible how subjects sum, but rather that it is *demonstrable* that they *cannot* sum. So, unlike Goff (2017), who argues

that it is difficult to understand how that could be the case, Coleman (2014) makes the more explicit, stronger claim that there is no way for subject summing to happen. He argues that this is the case because the *point of view* or the *qualitative field* of one subject excludes the points of view of all other subjects (Coleman 2014). A conjunction of the experiences that a subject is having and of those that they are *not* having constitutes that subject's unique point of view<sup>18</sup> so that if those experiences were combined with the experiences of another subject, the new subject would have to instantiate a conjunction of these two sets of experiences (Coleman 2014). This is contradictory unless the two original minds have identical sets of experiences, which is *prima facie* impossible because there would still be at least two distinct *tokens* of experiencing (Coleman 2014: 33). While it may be possible for two subjects to have identical experiences, it is impossible for these two instantiations of the relevant point of view to entail a third instantiation of the relevant point of view. The third subject, in whatever way it comes about, would be a further viewpoint resulting from two distinct token points of view, which, as Coleman argues, truly does seem incoherent, especially if their contents are mutually exclusive. For example, take two subjects, S1 and S2, experiencing red and green, respectively. The point of view of S1 is red with the exclusion of green (and all other experiences), while the point of view of S2 is green with the exclusion of red (and all other experiences) (Coleman 2014: 33). It is impossible for a further subject, S3, which is the result of the combination of S1 and S2, to experience red (excluding green and all else) and green (excluding red and all else) *at the same time*, for that would be contradictory (Coleman 2014: 33). While we might imagine the phenomenal contents of these experiences combining (e.g. into brown) and contributing to a subject's point of view, that still tells us nothing about the idea of a subject having two distinct points of view synchronously, which is what true combination requires (Coleman 2014: 33).

---

<sup>18</sup> I believe that the notion of a 'point of view' used by Coleman here is different from the pre-theoretical notion of a perspective from which we observe the world. His definition is rather based on sets of experienced (and, conversely, unexperienced) qualities, so that my 'point of view' is just the set of my thoughts, feelings, memories, etc.

Coleman concludes: “While the qualitative contents of consciousness may perhaps combine, consciousnesses themselves—subjects—cannot: this is precluded by the metaphysical logic of points of view“ (2014: 34).

To make it clearer what is meant by *point of view*, recall the notion used by Nagel (1974) from Chapter I. Nagel considers the specific point of view of bats since it is so radically different from the human point of view, considering that it involves cognitive capacities, such as echolocation, that we do not have. This is because he sees perspectives as *types* that can fundamentally differ, as in the case of the human compared to the bat. Furthermore, he argues that it is difficult to understand what one could mean by the objective character of some experience, abstracted from a particular point of view from which a subject apprehends the experience (Nagel 1974: 443). Thus, he claims that experiences, being necessarily tied to points of view, do not have an objective nature that can be apprehended from many different viewpoints (Nagel 1974: 443). So, while we might gain knowledge about bats’ experience *from our point of view*, we can never know it from its own point of view. This connects to Coleman’s idea that points of view or qualitative fields are sets of experiences that exclude other sets of experiences (or viewpoints), with our set excluding the bat’s set and vice-versa. In Nagel, points of view are *not* fundamentally private but are instead thought of as *types* (1974: 441). In fact, Nagel remarks that it is possible for us to take up a point of view other than our own so that we can know or describe the quality of another’s experience (1974: 441–42). They are subjective in the sense of only being adoptable by sufficiently similar subjects, who share the same perspective type, which is why he uses the example of a bat as a radically different subject from us (Nagel 1974: 442). Whether or not they are private also does not seem to be an issue to Coleman, who focuses on the mutual exclusivity of points of view. The difference between them is that Nagel is more focused on the epistemological aspect of viewpoints: we can never *know* what it is like to be a bat; we can never gain objective knowledge about subjective

viewpoints or the qualitative notion of “what’s-it-likeness”, regardless of whether he makes an ontological conclusion that there are perspective-dependent *facts* beyond physical facts, as I discussed in Chapter I. Coleman, on the other hand, is more focused on the metaphysical aspect: viewpoints are unique sets of experiences that necessarily exclude all other viewpoints – they are all qualitatively unique and had by particular subjects. This is crucial to Coleman’s argument. An essential feature of any consciousness that contributes to a further consciousness mereologically, and that has to be shared with that further consciousness, is the exclusion of all the experiential phenomena that it excludes. This is why Coleman (2014) argues that only viewpoints with exactly the same set of qualities or the same ‘zone of exclusion’, so to speak, could perhaps combine. However, the two approaches discussed seem to be compatible: Nagel addresses the exclusivity of viewpoints as *types*, across species, while Coleman addresses the exclusivity of viewpoints as *tokens*, even within the same species or type.

In addition to these definitions of viewpoints, Coleman also presents a specific argument against the form of combination desired by the panpsychist. He starts off by drawing an analogy with the combination of hydrogen and oxygen atoms into molecules of water: in that case, all three atoms forming the water molecule continue to persist, only in a modified way, after they combine (i.e. the oxygen atom “takes” an electron from each hydrogen atom) (Coleman 2014: 30). Thus, he defines combination as “the formation of a whole from components where the components continue to exist in the whole, but are intrinsically altered by combining with one another” (Coleman 2014: 30). A stark contrast to this type of combination would be what Coleman calls *mere aggregation*, where the entities in question are not necessarily conditioned or united in any particular way (2014: 30). The resulting higher-level properties of water, which are distinct from the lower-level properties of the atoms in isolation, are *structural properties*, in the sense that they are not ontologically autonomous from or inexplicable with reference to the lower level, meaning that one can intelligibly trace

back from the higher-level properties to the lower-level properties (Coleman 2014: 31). In the case of microsubjects combining into macrosubjects, if the combination is to be intelligible, it would have to obtain in the same way as it does in the case of water molecules. That is, microsubjects would have to “add up to” macrosubjects in the same way that atoms add up to form water molecules (Coleman 2014: 32). The points of view or qualitative fields of those subjects would have to combine so that the microsubjects having those points of view persist in the whole in the same way that atoms still persist while forming molecules (Coleman 2014: 32). This is impossible because, by hypothesis, the combination proposed by the panpsychist results in a single subjective point of view (Coleman 2014: 32)<sup>19</sup>.

## **2. Attempts to Solve or Avoid the Combination Problem**

After discussing the various forms of the combination problem, with a special emphasis on its subject-summing version, I now turn to attempts to solve or avoid it. I will discuss Goff's phenomenal bonding relation, Seager's combinatorial infusion, Tye's representational panpsychism and how it tries to respond to the combination problem, as well as cosmopsychism as an alternative theory that is apparently unburdened by the issue. The aim of this section is to call into question the strength of these proposals, without always strictly arguing against them. That is, I merely want to show that neither proposal firmly resolves the issue raised by the combination problem, which still retains its force against panpsychism.

---

<sup>19</sup> However, does the panpsychist have to accept that the microsubjects disappear when the macrosubject is formed? While Coleman's argument seems convincing, and I agree with it, it seems, in principle, possible to construct a form of panpsychism in which the microsubjects contribute in virtue of the phenomenal qualities they possess, their structural features, etc., instead of in virtue of their subjectivities. My proposal, that I will outline in Chapter IV, shares similarities with this (for now) provisional idea.

## 2.1 The Phenomenal Bonding Relation

One recent attempt at solving the subject-summing form of the combination problem is Philip Goff's idea of phenomenal bonding. He argues that the combination problem excludes the possibility of a group of fundamental subjects necessitating a further subject merely in virtue of them *existing* but that it does not exclude the possibility of them being related in some specific way that necessitates the existence of a further subject (Goff 2016: 292). So, even if subjects are *metaphysically isolated*, that does not mean that they cannot be related in a way that produces further subjects.

The notion of *metaphysical isolation of subjects* employed here by Goff needs to be clarified. He starts with the following epistemic principle (Goff 2016: 287):

Conceptual Isolation of Subjects (CIS) – For any group of subjects, instantiating certain conscious states, it is conceivable that just those subjects with those conscious states exist in the absence of any further subject.

From this, he draws a metaphysical conclusion (Goff 2016: 287):

Metaphysical Isolation of Subjects (MIS) – For any group of subjects, instantiating certain conscious states, it is possible that just those subjects with those states exist in the absence of any further subject.

This conclusion seems to be in line with what James's passage wanted to establish: the existence of a group of subjects does not necessitate the existence of a further subject. So far, this strand of thought seems coherent and in line with what we have previously discussed regarding the difficulties of subject-summing.

The proposed relation that then follows from these considerations is what Goff calls *phenomenal bonding*. However, the phenomenal bonding relation is merely a placeholder for something that is yet to be discovered. We do not have a transparent conception of such a



relation; else the solution to the combination problem would be obvious to all (Goff 2016: 292). That being said, even though we cannot form a transparent conception of phenomenal bonding, we can form *a* conception of it as “the relation such that when subjects stand in it they produce further subjects” and suppose that one exists (Goff 2016: 292).

The panpsychist could also identify a real physical relation as the phenomenal bonding relation, even though it is not originally conceived of as being phenomenal. If fundamental simples are phenomenal, the panpsychist can interpret the already obtaining physical relations between them as phenomenal too. That is, merely by obtaining between two phenomenal simples, the relation is potentially describable as phenomenal. By doing that, the panpsychist does not introduce any new relations but merely claims that some existing physical relation has served the role of phenomenal bonding all along. Of course, a lot has been written on the metaphysics of relations and on how to individuate them, on whether they are tokens or types, and so on, but the aim here is simply to show that there is a way for the panpsychist to make this move, regardless of how unconvincing it might seem. Additionally, according to Goff (2016: 293–94), the theoretical attractions of panpsychism would allow for us to presuppose the existence of such a relation in order to save the theory. There are also independent reasons for thinking that there are relations in the world for which we lack a transparent conception: for instance, intuitively, there must be a deeper nature<sup>20</sup> to spatiotemporal relations which underpins the structural understanding of them that we get through physics (Goff 2016: 294–95).

The relation that Goff specifically proposes as a candidate for phenomenal bonding is the spatial relation (or perhaps *spatial relations*; it is unclear why Goff uses singular here). However, this only works if we accept a phenomenal form of mereological universalism – the

---

<sup>20</sup> This is the crux of contemporary Russellian panpsychism.

view that all subjects are spatially related and always combine into further subjects – rather than a phenomenal form of mereological nihilism – the view that subjects never combine (Goff 2016: 299). Given that mereological nihilism rules out combinations of any sort, it is a non-starter since the subjects that we are pre-theoretically committed to appear as composite objects (Goff 2016: 299). If we accept universalism, then the identification of the spatial relation with phenomenal bonding could be defined as such that when phenomenal entities are spatially related, they determine a conscious subject (Goff 2016: 299–300). Provisionally, this is also compatible with the Russellian panpsychist’s demand for all structure to be grounded in a deeper nature: phenomenal bonding is the deep nature of the otherwise structurally conceived spatial relation (Goff 2016: 299–300). It is unclear why Goff favours universalism over other candidate theories that lie in between nihilism and universalism (see the following section). The reason he provides is that universalism allows us to identify spatial relations as phenomenal bonding relations (Goff 2016), but it does not seem to me that this is the only option in order to make such a move. Ultimately, Goff’s conception is fairly vague and leaves a lot of questions open, but it does function, at the very least, as a starting point in the search for an answer to the combination problem.

## 2.2 Problems with Phenomenal Bonding

With that being said, there are several problems with Goff’s proposal even if we take it as strictly propaedeutic. The first thing that strikes me as odd is the distinction Goff draws between subjects producing subjects in virtue of them *existing* versus in virtue of them *being related*. What does it mean to say that subjects produce other subjects merely by existing, without presupposing that they are related in some way? If we recall James’s passage quoted earlier, he does not fail to mention something akin to phenomenal bonding: “the 100 original feelings might, by a curious physical law, be a signal for its creation”, where that law could be

whatever relation we choose to identify with phenomenal bonding (1890: 160). He goes on to argue that such a production of further subjects would still be unintelligible: “they would have no substantial identity with it, nor it with them, and one could never deduce the one from the others, or (in any intelligible sense) say that they evolved it” (James 1890: 160). Thus, James did not forget to address a relation-based form of subject-summing and an opponent of panpsychism could similarly modify the combination problem to restrict such subject-summing. This makes me question whether the existence/relation distinction that Goff introduces does any theoretical work whatsoever.

In fact, Goff’s distinction seems of secondary importance compared to the sheer act of subject production. Whether subjects produce further subjects via existence or relations, the problematic part is that a *new* and *distinct* subject emerges, with its own unique perspective of viewpoint. Pre-theoretically, higher-level subjects such as human minds experience themselves as unified centres of experience, so how does the further subject, as a result of relations between other smaller subjects, acquire such unity and, with it, its identity? How does it acquire its new phenomenal character? The phenomenal bonding relation might give us a way of conceptualising how microsubjects could potentially combine but it does not bring us any closer to understanding how a *new* subject emerges as a result of those combinations. If the further subject is truly something ontologically novel, then the way it emerges according to phenomenal bonding must be brute. Considering that the whole point of contemporary panpsychism is to deny brute emergence (see Seager 2006, Strawson 2006), this is definitely not something that the panpsychist should endorse.

Another problem with phenomenal bonding is that it seems dangerously ad hoc. Presupposing that a phenomenal bonding relation exists either independently or via identification with a real physical relation in order to save panpsychism is permissible only if one has strong independent reasons for accepting panpsychism. Even then, the introduction of

a new relation should be motivated by something external, i.e. some additional motivating reason for accepting the relation that does not depend on panpsychist commitments, and the panpsychist should be expected to explain how the relation follows from previous premises. In response, the panpsychist could say that they need to have an overall theory that is internally consistent, so that they do not need independent justification for every part of the theory. However, their aim is to keep in place and not undermine the grounds for motivating the theory in the first place. If there are considerations of unintelligibility that are troubling the theory, as per the combination problem, it seems odd that they appeal to a relation that has unintelligibility problems ‘built in’, so to speak, because that vagueness with regard to what the phenomenal bonding relation is identified as would undermine the original motivations for panpsychism.

The panpsychist could instead claim a sort of mysterianism about the phenomenal bonding relation: precisely because it is a phenomenal rather than physical relation (or a physical relation conceived of as phenomenal in some respect), one cannot demand a definition of it in the same sense as one defines a physical entity; that is, we cannot in principle attain knowledge about what the phenomenal bonding relation is because: a) introspection does not seem to reveal what kind of relation that is, b) there is no way to study the relation in a way that is common to the hard sciences, so that we could form a mathematical, structural or nomic conception of it, and thus c) we cannot confirm or deny whether and which particular fundamental goings-on are phenomenal or not. For example, vis-à-vis Goff’s spatial-as-phenomenal relation, it would be impossible to confirm or deny that *that* relation, in particular, is consciousness-involving (rather than some other physical relation). Even if we could, as panpsychists, interpret physical relations as phenomenal because they obtain between phenomenal entities, that does not necessarily mean that *all* relations between those entities are or should be interpreted as such, and there is no way of confirming which one actually is

phenomenal. The guess with regard to which relation that *is* needs to be educated and explained, rather than seemingly arbitrary.

Apart from the spatial relation, there is something else that seems to be ad hoc in Goff's proposal: phenomenal universalism, the mereological claim that subjects always combine to make further subjects (Goff 2016: 299). Again, he accepts universalism instead of nihilism (without considering other alternatives) in order to salvage his combinatory form of panpsychism. This is particularly confusing considering the sheer implausibility of universalism. Any other theory that allows for some combinations into proper wholes would have worked better in Goff's argument than either of the two extremes. If anything, picking universalism makes the theory even less attractive to those who are yet to be swayed into accepting panpsychism. To many, I presume, the metaphysical baggage of universalism would be enough to drive them away from panpsychism forever.

Because he chose universalism, however, the naïve question comes to mind: if mental composition is truly unrestricted, why do human subjects remain isolated? The limit where mental composition stops clearly exists at some levels, such as with human subjects. Goff did not explain whether or why the limit exists and why it would pertain to some specific level of complexity if it does exist. It would have been interesting to see how theories other than universalism could accommodate panpsychism, or at least to get a defence of universalism qua universalism. Rather, it seems that Goff introduces universalism merely to save the theory instead of providing reasons why universalism is attractive, either in its own right or in conjunction with panpsychism.

### 2.3 Combinatorial Infusion

In this section, I turn to a different proposal aimed at solving the combination problem, the idea of *combinatorial infusion*, as discussed by William Seager (2010). According to Seager (2010), microsubjects could potentially stand in a certain relation that enables them to *merge* into a new aggregative mental state. On this view, the smaller constituents cease to exist and become “absorbed” by the resulting state, which potentially avoids Coleman’s previously mentioned viewpoint argument and the general notion that subject-summing is incoherent (Seager 2010). Specifically, Seager (2010: 13) introduces combinatorial infusion through three main characteristics:

1. The mental character of the combined or aggregative mental state stems from the mental characteristics of the constituents,
2. The combined or aggregative mental state is a novel state which in some way ‘absorbs’ or supersedes the mental states of the constituents, but
3. There is no radical or ontological emergence of the aggregative mental state; rather there is an intelligible relation which holds between the mental components and the resulting aggregative state.

The most important question for combinatorial infusion is whether all three characteristics can be consistently defended, with (3) being the most difficult one to elaborate, though for now, this will not be addressed. So, moving on, Seager (2010: 13) suggests that the correct way of understanding conscious states is through the notion of *large singles*, which are entities that are metaphysically simple, without any parts, but which are in some way extended (not necessarily spatially). To present an example, Seager draws an analogy to classical black holes: an entity which forms when a set of constituents is bound together via gravitational collapse, where the constituents could be any kind or form of matter (2010: 14). In the case of black

holes, the particular features of its constituents are destroyed and, in a sense, absorbed by the whole, meaning that classical black holes can be understood as large simples in the sense that they are partless yet extended (Seager 2010: 14). The way in which classical black holes come into existence should thus be understood as similar to how combinatorial infusion happens (Seager 2010: 15).

If Seager's proposal is viable, then combinatorial infusion can avoid some of the key issues of subject-summing, the unity of consciousness, and phenomenal bonding. Recall the problem of S1 and S2 entailing S3 without S3 being identical to them from before. Also, recall how we clarified and specified that problem through Coleman's viewpoint argument. Specifically, Coleman (2014) argues that microsubjects must persist after combination if that combination is to be true rather than merely aggregative, but that that is impossible since points of view cannot combine. Seager thus presents us with a form of combination where S1 and S2 need not persist for S3 to obtain. In combinatorial infusion, the original microsubjects, S1 and S2, cease to exist and are "absorbed" or replaced by the resulting "large simple" subject S3.

Furthermore, Seager's proposal seems to be a more viable option than phenomenal bonding. Rather than arguing that microsubjects stand in a certain relation that necessitates a further macrosubject, as Goff does, Seager argues that they instead stand in a relation that enables them to merge so as to form a further subject. Again, since Coleman's objection seems to apply to Goff but not to Seager, it seems *prima facie* plausible to favour combinatorial infusion over phenomenal bonding. Put simply, in Goff's case, microsubjects are related in some special way that necessitates a new macrosubject, while in Seager's case, microsubjects merge into a new macrosubject that is entirely the direct result of their interactions and properties (Seager 2010: 15).

Seager (2010: 11) also addresses James's original worry that subjects cannot combine by mere co-occurrence and that the resulting whole will not have any distinctive role, with its features being wholly determined and exhausted by the combined action of its parts. If the abovementioned example of classical black holes truly does provide us with a robust example of how combinatorial infusion might work, then we already have a template for how combinatorial infusion might explain consciousness (Seager 2010: 15). Thus, Seager provides an intelligible model of how subject summing could happen. In addition, the features of the whole are the direct result of the lower level. In other words, brute or radical emergence is not present within combinatorial infusion even though a new subject is created (Seager 2010: 15).

#### 2.4 Problems with Combinatorial Infusion

Coleman (2014) raises an objection to combinatorial infusion similar to the one I previously discussed, with regard to viewpoints. He argues, contra Seager, that the fusion of subjects could only provide for the *emergence* of a macrosubject, which goes against the basic combinatory panpsychist commitment to anti-emergentism (Coleman 2014: 35). Furthermore, he argues that the lower-level properties must contribute to the new whole "*in virtue of their metaphysical nature, or, otherwise put, while remaining true to what they are*" (Coleman 2014: 35). Since subjects are discrete entities, they cannot contribute to the unified subjectivity of the new subject, which is the result of merging precisely because their original subjectivity is obliterated in the process (Coleman 2014: 36–37). That is, they cannot contribute by remaining 'true to what they are'. It is then unclear that combinatorial infusion retains the purported benefits of panpsychism relative to physicalism since it re-introduces a form of emergence and hence also a kind of unintelligibility.



He further argues that this results in a form of emergence: subjects cannot impart their subjectivity; rather, what is necessary in combinatorial infusion is the destruction of the property of their subjectivity, which is precisely the property needed in order to achieve a higher-level subject (Coleman 2014: 37). By ceasing to exist, Coleman interprets the microsubjects as brutally causing the further subject to come into being, which he considers a form of radical emergence (Coleman 2014: 37). Again, since the (constitutive) panpsychist holds a commitment to anti-emergentism, this goes against the spirit of their theory. The dilemma Coleman presents to Seager's view is thus the following: either the microsubjects are still present after combination, in which case his viewpoint argument applies, or they are not, in which case the combination is an instance of radical emergence.

However, Seager could argue that Coleman did not properly interpret his argument. Microsubjects do not merely cease to exist in combinatorial infusion. Much like in the case of classical black holes, the constituents are still present in the forming of consciousness, still existing, even though their individual features are overridden by the whole. Compare this idea to the previous discussion regarding the formation of water molecules: they both involve a form of combination where the constituents are heavily modified, but still present. Therefore, this is not the case of radical emergence since there is an intelligible relation obtaining between the mental constituents and the final aggregative whole (Seager 2010: 13). Additionally, Seager provides us with a potential physical correlate to how complex mental states are formed: quantum entanglement. When a set of particles is entangled, the resulting composite object has holistic properties that are not reducible to the properties of its constituents, but Seager contends that this is still not a case of radical emergence: the entangled state is still a predictable result of the basic laws of quantum mechanics, as predicted by Schrödinger (Seager 2010: 13). Both the black hole and the quantum entanglement examples can be intelligibly applied to consciousness. Indeed, these physical models might underpin how complex consciousness

comes about and how points of view truly can integrate by merging into an overarching, more complex, state.

In addition to this, Seager (1995) argued in an earlier paper that quantum mechanics provides us with a way of conceiving how particles can form wholes that are genuinely new rather than being mere aggregates while being committed neither to radical emergence nor to James's understanding of combination as exemplified in his classic combination problem for panpsychism. Seager (1995) invokes the double-slit experiment, where a beam of particles (such as photons) is directed towards a separated pair of slits. After passing through the slits, the particles hit a detector screen where an interference pattern is formed as a result of the interaction between the two paths that the particles can take to reach it (Seager 1995: 277). Intuitively, one might assume that the passing of the particles through the slits results in a mixture which contains half of the particles in state  $\Psi_1$  (they pass through the left slit) and the other half as being in state  $\Psi_2$  (they pass through the right slit), but that is not the case (Seager 1995: 284). Instead, they are in the superposition of the two states,  $\Psi_1 + \Psi_2$ , which is a combination of the states forming an entirely new state with properties different from those of the mixture (Seager 1995: 284). It seems thus that quantum wholes are not just mere sums of their parts, but the ability to enter into superpositions is certainly tied to the properties of the particles that enter into it.

This provides us with another way of conceiving how microsubjects could persist after combination. If they could enter what we might provisionally call *phenomenal superposition*, then we can think of them as forming a new subjective whole while persisting post-combination. Even though a new subject is the ultimate result, the properties of the preceding microsubjects, especially those that pertain to the ability to enter into superposition, continue contributing to the state of the whole. Again, without being committed to the truth or falsity of any particular empirical account, this still leaves us with a way of conceiving how subjects

could combine. However, if the *subjects* continue to persist and not just a set of their properties, how is this proposal not also susceptible to Coleman’s argument? A potential answer might be that a new subject emerges and “takes over”, as the whole, while the original microsubjects contribute in an idle way. The original microsubjects are modified in a sense that enables the formation of the subjective whole via superposition. This notion, ultimately, rests on recently discovered or merely potential relations in quantum mechanics that we do not yet fully understand, though they reveal new possibilities. It certainly does not significantly strengthen Seager’s overall case.

Alternatively, Seager could contend that microsubjects *do* cease to exist post-combination, but that this does not lead to any problematic emergentist view. If we imagine a set of three microsubjects, S1, S2 and S3, we might conceive of a form of asymmetrical fusion, where S1 and S2 confer their phenomenal contents to S3 while ceasing to exist. After the process is finished, S3 will be more complex than it was before, now having a ‘blend’ (or some other form of co-occurrence) of the mental contents previously had by S1 and S2. The original state of S3 does cease to exist as well, but not in the sense of being annihilated. Instead, it is simply modified in form, now being more ‘packed’ with phenomenal content. During this process, no instance of radical emergence appears. Rather, what was present in the original state (the set of S1, S2 and S3) is now altered, with S1 and S2 being annihilated, but with their phenomenal contents now being taken over by S3<sup>21</sup>.

There are several challenges that this understanding of fusion must face. The immediate response to this would be: why does this kind of view, where subjects cease to exist either completely or by undergoing change, not result in a problematic form of emergentism, but non-panpsychist combinatorial views do? For instance, why could not a physicalist,

---

<sup>21</sup> My proposal in Chapter IV will discuss a similar model, though it will not be burdened by fusion, absorption, and similar notions. It will also try to accommodate Coleman’s objections.

panprotopsychist or panqualityist insist that the fundamental properties they postulate are also sufficient for this non-radical kind of subject emergence? The panpsychist could argue that it is necessary to posit microsubjects since the phenomenal contents must be *had* by subjects and because those subjects must be able to then confer them to another subject, which grows in complexity. Considering that fundamental subjectivity is not present within physicalism, panprotopsychism and panqualityism, to take several examples, they cannot avoid a problematic form of emergence pertaining precisely to the coming forth of subjectivity. Where is the subject in their view to begin with? In contrast, within the purview of panpsychism, all problematic forms of emergence are eliminated under this model of combination (or fusion). Subjects change and cease to exist, but they do not emerge out of thin air. Postulating microsubjects is thus warranted since they are necessary to facilitate these kinds of processes: again, phenomenal contents cannot exist without being experienced by subjects and more complex kinds of subjects cannot come about without them absorbing or integrating the phenomenal contents of preceding subjects. This asymmetric model of phenomenal fusion thus amounts to a gradualist view of consciousness, where subjects can slowly grow in complexity by gaining more and more phenomenal contents.

However, another issue here is trying to understand whether this is a response to Coleman's original viewpoint argument or merely an abandonment of the idea that subjects persist post-fusion. I suggest that it lies somewhere in the middle. Namely, it is indeed a rejection of subjects persisting post-fusion, but not a rejection of the possibility of fusion per se. In other words, it offers a way of understanding fusion beyond the confines of Coleman's original conditions for combination. It thus also serves as a response to Coleman since it shows that his assessment was incomplete. That is, he did not manage to exclude all the relevant ways of understanding fusion. Thus, the combinatorial infusion panpsychist still has a way of making sense of how less complex subjects can produce more complex subjects. It just might not be in

a way that was previously anticipated by opponents such as Coleman, nor in the way that was originally intended by Seager. It is thus clear that Seager's (2010) proposal is not as robust as it originally appears, while Coleman's objection still succeeds, at the very least, in deflating combinatorial infusion. While there might be ways of justifying the idea of subjects fusing, the mere sight of danger with regard to re-introducing a form of emergence is enough to weaken the appeal of his argument.

### 2.5 Cosmopsychism as an Alternative

Cosmopsychism is the claim that the universe as a whole is the ontologically fundamental level of reality, not the lowest level, and that it instantiates consciousness as such (Goff 2017). It is based on Jonathan Schaffer's *priority monism*, the view that "the cosmos is the one and only basic actual concrete object, prior to any of its proper parts" (2010: 65). This is in contrast to *existence monism*, which is the view that only one concrete object exists – the cosmos (Schaffer 2010: 66). While existence monism entails priority monism, priority monism does not entail existence monism since it allows "for the existence of many derivative proper parts of the cosmos" (Schaffer 2010: 66). What priority cosmopsychism does is simply add consciousness into the picture, so that the cosmos as a whole is conscious or consciousness-involving, and that its derived parts can be conscious as well in virtue of the consciousness of the whole (Goff 2017: 234). Naturally, since cosmopsychism uses a top-down ontology, subjects do not need to *sum* – they are instead derived from the cosmos as a whole. This is how cosmopsychism avoids the subject-summing version of the combination problem. Another putative advantage of cosmopsychism is that it does not presuppose *fundamentalism*, the notion that entities at the lowest level of reality are fundamental (Nagasawa & Wager 2016). In fact, it is ontologically neutral on whether there is a fundamental, smallest level at all. That is, if one

could demonstrate that fundamentalism is false, one would also demonstrate that panpsychism is false, leaving cosmopsychism as a remaining option.

Priority cosmopsychism does face a different kind of issue: the *derivation problem*, or the question of how medium-sized consciousness, e.g. human, derives from the cosmic consciousness of the whole (Nagasawa & Wager 2016: 121)? One answer is that it is possible to imagine the consciousness of a human individual being divided into smaller, less fundamental segments. So, by analogy, it is also conceivable that the consciousness of the whole cosmos can be divided into smaller, less fundamental segments (Nagasawa & Wager 2016: 121). With regard to priority monism, Schaffer (2010) also offers several possible solutions to the derivation problem, through concepts such as *distributional properties*, *regionalised properties*, and *regionalised instantiation*. An example for a distributional property is ‘being polka-dotted’, so that the whole can be polka-dotted, which is a coherent and unproblematic claim that entails heterogeneity for the whole (Schaffer 2010: 59). An example for regionalised properties is ‘bearing the redness relation to here and the greenness relation to here’, so that a seemingly monadic property can be heterogeneous (Schaffer 2010: 60). An example for regionalised instantiation would be similar to the example for regionalised properties, but what is being regionalised are instantiations rather than the properties themselves, so the cosmos can be heterogeneous by instantiating-here red and instantiating-there green (Schaffer 2010: 60). It is less clear whether this can be consistently applied to consciousness, however, but the point is that the cosmopsychist has resources which seemingly render the derivation problem less serious than the traditional combination problem for panpsychism.

It is unclear whether priority cosmopsychism avoids the *combination* problem, even if we accept that the derivation problem is milder. The idea of a top-down ontology for consciousness seems to me incoherent, considering the natural processes of growth from

conception and evolution, which are evidently bottom-up processes. How would we reconceptualise the long and slow process of matter becoming more complex, until it reaches the threshold of consciousness, and then the threshold of self-awareness and abstract thought, as a *top-down* process? Similarly, on a shorter scale, how can we conceptualise the growth of an organism from conception until adulthood and death as a top-down process? At the very least, it is counterintuitive that it *is not* so.

The cosmopsychist could respond by admitting of bottom-up processes pertaining to consciousness in addition to the top-down process of derivation, but this is clearly unparsimonious and leads to them having to face the combination problem *and* the derivation problem. Alternatively, they could argue that derivation is unrelated to the physical processes of evolution and growth, or, more generally, that no bottom-up processes are necessary in order to explain consciousness. The process of derivation might be *diachronic*, in the sense that it happens at a specific point in time, as opposed to evolution and growth, which are *synchronic* processes, in the sense that they happen *over a period of time*.

However, E. J. Lowe (2012) has criticised Schaffer's conception of the cosmos, arguing that the cosmos can exist as a mereological sum of many parts, but that this is insufficient for the role of the ontologically prior, fundamental whole that is needed for priority monism. According to Lowe, such a sum is not prior to its parts but only exists *in virtue* of them. If that is true, then the cosmopsychist cannot use Schaffer's conception of the cosmos either, to play the role of the fundamental whole which instantiates consciousness. I find this convincing or at least strong enough to disregard cosmopsychism as a proposal that truly avoids the combination problem. Furthermore, even if it merely replaces the combination problem with the derivation problem, it still has to give a positive picture of how derivation happens. If a theory can avoid the combination problem altogether (including similar problems, such as

derivation), then that theory ought to be preferred. This will be discussed as part of my original proposal in Chapter IV.

## 2.6 Tye's Representational Panpsychism and the Combination Problem

According to Tye (2021b: 81), only generic or bare consciousness\* is at the base level, rather than multiple sorts of consciousnesses. To suppose otherwise would be unparsimonious and “impossible to reconcile with a general representational approach to the different types of conscious state” since there is “no representation at the most fundamental level” and thus nothing at that level which would ground representations for different types of conscious states (Tye 2021b: 81). If that is the case, then Tye seemingly avoids the combination problem. Moreover, as he argues, his proposal does not have to give an account of why *this* particular combination of ultimates produces human consciousness, while *that* particular combination of ultimates, such as in rocks, for example, does not produce consciousness, which is a problem he considers unanswerable (Tye 2021b: 81–2). Furthermore, Tye (2021b: 81) claims that it cannot be the case that spatial arrangements can transfer consciousness\* from parts to the whole. Instead, consciousness\* already exists at the level of fundamental parts (Tye 2021b: 84). The proposed solution to the combination problem that he describes is thus the following:

“A given complex entity C is conscious\* because its fundamental parts are conscious\* and they are arranged in the right way. The truth of this claim tracks the truth of the counterfactual that if the fundamental parts of C had not been conscious\* or they had not been suitably arranged, C would not have been conscious\*.” (Tye 2021b: 85)



To clarify, this is not a case of brute emergence, for Tye (2021b: 85) claims that brute emergence for consciousness\* with complexity is not an option. Instead, what is required for a transfer of consciousness\* is:

“an arrangement A of the fundamental parts that is not found in all complexes and that supports the counterfactual claim that if the fundamental conscious\* parts had not been arranged in way A, then the resultant complex state would not have been conscious\*” (Tye 2021b: 85)

So far, this seems somewhat unclear. Does Tye just assume that consciousness at the fundamental level requires the coming about of complex consciousness at higher levels, through counterfactuals? While I have that suspicion, Tye spells out the story in greater detail. Relying on the *global workspace theory*, the view that experiences are informational states that arise in an integrated central workspace, Tye claims that it offers a simple characterisation of complex conscious brain states (2021b: 86–7). From that, he goes on to claim the following:

“Consciousness\* is transferred, I want to suggest, from the fundamental parts to certain complex wholes as long as the fundamental parts are arranged so as to form states that play the sort of role assigned to conscious states in GWT. Given this arrangement, transfer occurs and the states that are formed in this way are conscious\*. Further, in being conscious\* and playing the relevant global workspace role, these states are thereby conscious.” (Tye 2021b: 88)

Recall what consciousness\* is in Tye’s proposal. It is a basic, non-representational form of consciousness, contrasted with conscious states, which are essentially representational. What he claims here is that physical ultimates transfer consciousness\* - the basic form of consciousness – to more complex wholes if and only if those physical ultimates are arranged in a way that plays the role for conscious states as described by the global workspace theory;

that is, that they are integrated in a centralised workspace. If that arrangement occurs, the whole is conscious\*, and if those states feature in a global workspace, the whole is conscious (without the asterisk; representationally). The transfer of consciousness\* from fundamental entities to complexes is thus based on fundamental entities being organised *functionally* in the proper way (Tye 2021b: 88).

While this is certainly a new and unique approach, I find it unclear whether it is convincing as a solution to the combination problem. While I have no problem with the idea that conscious states, conceived of as essentially representational, are separated from consciousness\*, and that those representational conscious states are constituted or defined functionally or relationally<sup>22</sup>, I worry whether the story of consciousness\* being transferred to the whole makes sense. Fundamental entities are conscious\*, but if they are conscious in *any* sense of the world, then they are minimal kinds of subjects. Since the complex attains consciousness\* from those fundamental subjects, it is unclear to me whether Tye can avoid the subject-summing version of the combination problem. Moreover, this proposal might also face similar objections as I discussed for combinatorial infusion, particularly Coleman's (2014): do the conscious\* microsubjects cease to exist or do they contribute to the consciousness of the whole?

One response that Tye (2021b: 91–2) provides is that consciousness does not necessarily require a person who is conscious and, consequently, there is no requirement that the consciousness\* found at the fundamental level belongs to a person (Tye 2021b: 92–3). He then goes a step further and extends this to psychological subjects: “there can be consciousness\* at the most fundamental level without any person, any rational being, any psychological subject” (Tye 2021b: 93). This is a highly unusual and counterintuitive view,

---

<sup>22</sup> In Chapter IV, I argue for something very similar.

however. I find Tye's proposal unconvincing if it demands for us to accept a highly idiosyncratic view of basic consciousness, or just consciousness in general, that is not had by any subject. This does not necessarily refute Tye's representational panpsychism, but in my view, it severely undermines its appeal.

### **3. The Unity and Disunity of Consciousness**

In this section, which is more speculative, I first discuss how the debate on the unity of consciousness might help the panpsychist with regard to the combination problem. Then, I examine whether apparent cases of disunified consciousness can serve as a potentially empirical case that consciousness can combine and de-combine, which might justify the panpsychist's commitments. Unlike the specific proposals addressed previously, these two approaches present a broader and more general strategy that the panpsychist can adopt. However, I will ultimately argue that their merit is limited and that it is unclear whether they deflect the subject-summing version of the combination problem.

#### 3.1 Unity of Consciousness

There are two main strands of thought on how consciousness is unified: the experiential parts view and the no experiential parts view (Brook and Raymont 2017). According to the former, a unified conscious experience is composed of other experiences (Brook and Raymont 2017). On the latter, a unified conscious experience has a complex object or content but no experiential parts (Brook and Raymont 2017). Rather, the new experience replaces previous experiences instead of containing them as parts (Brook and Raymont 2017). One particular version of the experiential parts view is *subsumption*. Under subsumption, two subsumed unified states produce a conjoint phenomenology so that there is something it is like for the

subject to be in both states at once (Bayne 2010: 32). A similar notion is *co-consciousness*, which involves the claim that there is a relation that conscious states have which enables them to be experienced together (James 1909). Dainton (2000), for example, differentiates two co-consciousness theses: the strong and the weak introspection thesis or I-thesis. The strong I-thesis is the view that co-consciousness is constituted by ‘introspectibility’, where experiences are co-conscious if and only if they are introspectable by a single introspective awareness (Dainton 2000: 35). The weak I-thesis claims that co-consciousness is not constituted by ‘introspectibility’ but that the two are correlated instead: “if a group of experiences are co-conscious they are all actual or possible objects of a single introspective awareness” (Dainton 2000: 35).

Regarding the no experiential parts view, one approach relies on postulating a unified conscious field where everything that is experienced is experienced together by entering into the same phenomenal content, where that phenomenal content is the content of a single non-composite experience (Tye 2003: 36). According to Michael Tye (2003), there are not many experiences, corresponding to each individual sensory modality, that combine to produce a new and unified experience, but rather just *one* experience. Another option is the *joint consciousness* view, where if “an experience that one is having provides consciousness of any item, then it provides consciousness of other items and of at least some of the items as a group” (Brook and Raymont 2009: 567–68). This view attempts to be neutral with regard to whether there are experiential parts (Brook and Raymont 2017).

The no experiential parts view could be compatible with forms of panpsychism that allow for the merging or fusion of subjects or experiences into new and unified distinct wholes, so that complex experiences are not currently composed of parts but are rather the end-result of their unification. If unified consciousness is produced by the aggregation of experiential parts, then there is hope yet of making sense of the idea of phenomenal characters of

microsubjects combining so that they produce a unified conscious experience and thus a new subject. Though the question remains: is the new subject a new centre of experience that now has a conjunctive phenomenal character? It is important to note that the experiential parts view does not regard experiential parts as each being possessed by distinct subjects, so it is not immediately apparent that the view will help us in addressing the subject-summing problem.

The case is that, in both the experiential parts and the no experiential parts views, the experience is *presented to a subject*. Indeed, experiences necessitate the existence of a subject, unless one can make sense of the idea of experiences existing without being experienced by anyone. It might then follow that a conjunctive phenomenal character necessitates the existence of a new subject who does not experience A, B and C in isolation but experiences A & B & C. The idea of a conjunctive phenomenal character existing *without* there being a new subject to which that character is presented is ultimately unintelligible and should be rejected without further consideration.

The problem with the more tenable option, where the conjunctive phenomenal character is had by a new subject, is that it does not avoid the subject-summing variant of the combination problem. At best, it might avoid the problem of how experiences combine, but that is unclear as well. The idea that complex experiences are mere aggregations of simpler experiences is unattractive. The simpler experiences must be unified in a coherent way, so that the resulting complex experience does not contain mutually exclusive, contradictory content, such as containing the experiences of brightness and darkness at the same time, for example. This option would not avoid Coleman's (2014) viewpoint argument. One must thus posit some sort of relation between the simpler experiences that assures the coherency of the complex experience (Brook and Raymond 2017). For instance, Susan Hurley (1994: 59) has argued that there must be a normative basis for the unity of consciousness: we cannot be aware of inconsistent experiences, such as seeing light that is red and green at once, but even if the

experiences are consistent, they could still be disunified in a single consciousness. Something about the unity of consciousness must guarantee that experiences unify in a coherent and consistent way. For instance, there might be a physical basis that is responsible for establishing coherent relations between experiences, a relation based on the interaction of brain and world, unity of behaviour, cognition, etc. (Brook and Raymond 2017). Still, while this approach can offer a potential solution to the experience-summing problem, it does not avoid the subject-summing problem. Even if it succeeds to explain how complex experiences result from simpler experiences, it tells us nothing about how a *subject* comes about, which is the central problem in question.

### 3.2 Disunified Consciousness

In this section, I will take a look at apparent cases of disunified consciousness to see if they constitute a counterexample to the idea that consciousness must necessarily be unified and thus allow for the possibility of several centres of consciousness combining. In the final paragraphs, I will connect the discussion to Coleman's viewpoint argument in order to explore whether these cases constitute empirical counterarguments to his challenge. There are three main ways in which consciousness can be disunified: by being split into distinct streams of consciousness, by being partially unified, and by being completely disunified or shattered. For each category, there are corresponding empirical cases that seem to suggest it.

The first and arguably the most important of these cases are commissurotomies, popularly known as brain bisection cases. A commissurotomy is an operation where the corpus callosum, a collection of about 200 million neurons connecting the left and right hemisphere of the brain and enabling their mutual communication, is cut (Brook and Raymond 2017). The operation was most commonly performed to treat severe epilepsy (Nagel 1971). Interestingly

enough, in day-to-day living, patients suffer from no difficulties and appear as experientially unified as they were before the operation (Brook and Raymont 2017). What is even more interesting is that under certain and very specific conditions, the patients behave as though they have two distinct centres of consciousness within them (Brook and Raymont 2017). Tachistoscopic stimulation, an example of such strict laboratory conditions, was discussed by Nagel, where signals are flashed on a screen “on one or the other side of the midpoint of the patient’s gaze, long enough to be perceived but not long enough to permit an eye movement which would bring the signal to the opposite half visual field and hence to the opposite side of the brain” (1971: 399). Nagel continues:

„The results are as follows. What is flashed to the right half of the visual field, or felt unseen by the right hand, can be reported verbally. What is flashed to the left half field or felt by the left hand cannot be reported, though if the word 'hat' is flashed on the left, the left hand will retrieve a hat from a group of concealed objects if the person is told to pick out what he has seen. At the same time he will insist verbally that he saw nothing“ (1971: 400).

One further example is an experiment where the word ‘taxable’ is flashed on a screen in a similar manner so that the left visual half sees only the letters ‘tax’, while the right sees only the letters ‘able’ (Brook and Raymont 2017). When asked what word is being shown, the patient will say ‘tax’, while their hand will write ‘able’. This is so because speech is usually controlled by the left hemisphere, while the left hand is usually controlled by the right hemisphere (Brook and Raymont 2017).

The second class of cases, concerning partially unified consciousness, is exemplified in the work done by Justine Sergent (1990). Sergent flashed a numeral in front of a patient’s left and right visual field and asked them to say whether the numbers are the same or different (as

reported by Hurley 1994). In this case, the patients were no better than chance, which is in line with what one would expect in commissurotomy cases since neither brain hemisphere has access to the content of the other (as reported by Hurley 1994). In the second experiment, they were asked to state which number is higher and press a lever on the side of the higher number. In this case, the patients were highly accurate and showed no difficulty in judgment (as reported by Hurley 1994). Similarly, in the third experiment, they could indicate with high accuracy whether the numbers were equal or not. Based on these considerations, Susan Hurley (1994: 53) raises the question of whether the patients can have a partially unified, partially disunified consciousness, where the disunity affects only some categories of information but not others?

The third class of cases, where consciousness is radically disunified or ‘shattered’, can be understood through the prism of severe schizophrenia. A patient suffering from it apparently loses the ability to have a unified experience of the world and of themselves (Brook and Raymont 2017). For example, the patient is unable to complete coherent sentences, to connect perceptions, beliefs, and motives into plans of action, including those necessary for basic survival, such as obtaining food or reacting to pain (Brook and Raymont 2017). The subject, in this case, can be said to have fragments of experience which are unintegrated and disunified to the extent of disabling the subject from leading a normal life or interacting with others (Brook and Raymont 2017). Emil Kraepelin, one of the earliest commentators of schizophrenia, considered a lack of the unity of consciousness to be one of its key features (as reported by Bayne 2010). However, even though it might seem intuitive to claim that consciousness is radically disunified in patients, that conclusion is not obviously supported by empirical data. For instance, Bayne (2010: 158) argues that there is no reason to conclude that patients have multiple conscious states that are not integrated by a single phenomenal field. Instead, he identifies schizophrenia as affecting the narrative rather than phenomenal unity of consciousness: patients have issues keeping their thoughts on track, but that does not



necessarily imply that they lack phenomenal unity of consciousness. Schizophrenia is a heterogeneous condition, best thought of as a set of disorders rather than as a singular disorder, so any simplified and singular characterisation faces the danger of being incomplete (Bayne 2010: 157).

There are many more conditions that seem to suggest disunified consciousness, such as dissociative identity disorder, anosognosia, simultanagnosia, dysexecutive syndrome, and even mirror twins, where one consciousness seems to span two subjects (Brook and Raymond 2017). All of them provide us with reasons to doubt whether consciousness necessarily needs to be unified. Especially pertinent are the commissurotomy cases because they seem to directly contradict the idea of a unified consciousness.

These cases are, of course, relevant to the panpsychist. If it is truly the case that consciousness can either wholly or partially split into two, such as in commissurotomy cases, or into many fragments of experience, such as in cases of severe schizophrenia, then it is possible for streams of consciousness to combine into a unitary one as well. For instance, if we could find a way to regenerate the corpus callosum in commissurotomy cases in the future, we could also expect their split streams of consciousness to reunify. In principle, this gives room to the panpsychist to claim that consciousness truly is a malleable property, something that can be split into multiple subjects and reunited potentially even at the fundamental level. Thus, if we confirm that commissurotomy patients, for example, are truly instantiating two separate streams of consciousness, then the panpsychist has a way of confirming the possibility of their theory – or of at least strengthening its prospects – despite facing pressure from the combination problem. In other words, the case of commissurotomy patients would be an empirical example where it is plausible to postulate two subjects where there previously was one, assuming that such an interpretation of split-brain cases is more plausible than Bayne's switch model, for instance. The combination problem would then transform into a weaker demand to explain how

consciousness splits and unifies, but it would no longer serve the purpose of denying that such malleability is impossible or incoherent. In fact, if only one of the abovementioned examples of apparent disunity of consciousness is plausible, it would serve as a foothold for the panpsychist.

However, there are several important arguments against interpreting these cases as true instances of disunified consciousness. Probably the most important opposition to such interpretations comes from Tim Bayne (2010), who has argued that consciousness remains unified in commissurotomy cases, but that a patient's single stream of consciousness switches between the two hemispheres, effectively taking turns in order to support consciousness. Bayne appropriately calls this the switch model. In the experiment with the word 'taxable', for instance, the patient can be said to be conscious of the word 'tax' at one time and of the word 'able' at another time, but they will not be conscious of both words at once (Bayne 2010: 210).

According to Elizabeth Schecter (2012), there is evidence that fluctuations do occur between hemispheres, which makes the switch model potentially explanatorily relevant to commissurotomy cases, but the switch model still fails in preventing the attribution of dual streams of consciousness to commissurotomy patients. It is implausible that conscious activity in one hemisphere strictly precludes activity in the other, which is supported by experiments where the hemispheres do indeed simultaneously participate in consciously-driven behaviour (Schecter 2012: 211). However, even if the hemispheres are never simultaneously conscious, they would still be associated with distinct conscious streams since they would be psychologically distinct from one another (Schecter 2012: 219). They would have different sets of conscious experiences which would occur at different times, and which would be produced and located within distinct mental systems (Schecter 2012: 219). The switch model thus is unsuccessful, according to Schecter (2012: 221).

It is important to note that the two-stream interpretation of commissurotomy, which is used here as the most salient example of apparent disunity, is controversial and that there is as of yet no consensus on how to properly understand commissurotomy cases. This brings us to the question: is a commitment to such a controversial interpretation of commissurotomy really necessary for the panpsychist to answer Coleman's previously discussed viewpoint argument? Is that really preferable to accepting Coleman's conclusion? In a sense, it is at the very least a possibility. For instance, one could claim that Bayne's switch model suggests that one subject has access to two points of view. This might just lead us back into the two streams model, but not necessarily for any wrong reasons. One could argue that every interpretation of commissurotomy that is committed to any kind of bifurcation in consciousness, be it narrative, phenomenal or information-based, necessitates that we postulate two different viewpoints in one patient and thus two separate streams of consciousness. It is unclear whether that necessitates that these two streams are had by two different subjects, which makes the issue even more puzzling, but it allows for the possibility of a single subject having access to multiple viewpoints which potentially brings Coleman's argument into question: if that is the case, one subject truly can have separate, mutually incompatible viewpoints. If the streams are had by two distinct subjects, then that allows the panpsychist to base their claims about the possibility of combination and, potentially, de-combination, on such cases.

To bring this into focus, let us briefly revise Coleman's argument: he argued that subjects cannot combine because: a) viewpoints are mutually exclusive, and b) combination requires for the original lower-level parts to survive the combination and contribute to the higher-level whole. If the commissurotomy example is interpreted as patients having two separate streams of consciousness, then two conclusions are possible: either the patients have two separate subjects within them with each having their own corresponding viewpoint, or there are three subjects now – S1 who has the first stream, S2 who has the second, and S3 who

is nonidentical to S1 and S2 but has access to the stream of each. Even though none of this is conclusive with regard to the defence of panpsychism, it helps us conceive of the idea of a subject (S3) having two streams of consciousness, where those streams are distinct and had by two subjects (S1 and S2), independently of whether we are committed to the truth or falsity of any particular empirical account. It also helps us conceive of how multiple viewpoints might be integrated into an overarching one: if it is plausible to suggest that a reconstruction of the corpus callosum would re-integrate the separated streams of consciousness, then that makes it at least conceivable how viewpoints can combine. As such, it serves as a potential counterexample to Coleman's argument since it demonstrates the conceivability of two streams (and thus two viewpoints) (re-)uniting.

Finally, let us briefly examine an interesting approach that was proposed by Thomas Nagel (1971), who seemingly dropped the concept of unity altogether instead of arguing against disunity cases. Specifically, he argued that the unity that we ordinarily experience might not be anything absolute (Nagel 1971: 410). Furthermore, he argued against the countability of subjects: "If I am right, and there is no whole number of individual minds that these patients can be said to have, then the attribution of conscious, significant mental activity does not require the existence of a single mental subject" (Nagel 1971: 409). Nagel concludes that the common-sense idea of a single person occupying a body might become obsolete one day, thus effectively arguing that cases of disunity are not special cases at all, but only more salient examples of what obtains for everyone (1971: 411). If Nagel is right in his prediction, then the combination problem for panpsychism becomes less severe: if there is no unity of consciousness at all, then the panpsychist is less pressed to provide a model of how subjects combine into a new, more complex subject. Perhaps they simply do not (or at least not in the strictly phenomenal sense demanded by the combination problem).

However, there is one major reason for rejecting cases of disunity altogether and siding with Coleman (2014). Unlike the rare cases that I discussed, combinatory panpsychism would require billions of microsubjects combining into one stream of consciousness, each of them being a distinct subject that either retains or loses its subjectivity in the process. If the microsubjects retain their subjectivity, then the danger of how their viewpoints can combine in a non-contradictory manner reappears. If they do lose their subjectivity, then the danger of radical emergence (of the new subject) reappears. It is unclear then whether panpsychist combination is similar in kind to the combination of streams of consciousness that might happen in cases of disunified consciousness. Coupled with the general arguments against disunified consciousness, which would be a shocking and likely discipline-changing phenomenon if true, I do not see this as a viable option for the panpsychist with regard to tackling the issues raised by the combination problem. As I will try to demonstrate in Chapter IV, there is a way of formulating panpsychism that does not require for subjects to combine or fuse at all. If such a version of the theory is viable, then putative cases of disunified consciousness would neither be a problem nor a potential resource for the panpsychist. That is, the theory would be neutral with regard to this debate.

#### **4. Discussion**

This is the final section in which I first offer an overview of the discussed proposals and define how I will go forward in this work. I will evaluate which proposals are more convincing than others and explain the constraint that accepting the force of the combination problem might give to the panpsychist. Then, I return to what I discussed in Chapter II and consider whether a purely categorical form of panpsychism would help with solving the combination problem. Cautiously, I will argue that it might be beneficial to the panpsychist,

but that it also depends on accepting the complete framework of categoricism and, consequently, inheriting all of the issues that categoricists face.

#### 4.1. Overview of the Proposals

In this section, I will clarify what my position is on the proposals for solving the combination problem that I discussed. The phenomenal bonding relation, as proposed by Philip Goff (2016), is best seen as a general framework or blueprint for what such a genuine relation might be. Ultimately, though, I do not find Goff's proposal convincing, considering his at times ad hoc commitments. For example, spatial relations being the phenomenal bonding relation, and, in turn, universalism being accepted solely in order to justify that. In contrast, the combinatorial infusion view proposed by William Seager (2010) seems more convincing since it at least provides a positive conception of how microsubjects could produce more complex subjects at higher levels, though I find the danger of reintroducing radical emergence, as well as the overall argument against combinatorial infusion by Coleman (2014), to be strong reasons for why the panpsychist should not rely on Seager's proposal. Similarly, I also do not find cosmopsychism particularly attractive because of its commitment to a very unusual picture of the world, which can be said for Tye's (2021b) representational panpsychism as well, if it requires for us to accept that consciousness is not necessarily had by persons or subjects. The two broader strategies that I discussed, pertaining to phenomenal unity and disunity, are certainly illuminating, though I find the results inconclusive and think that the panpsychist should look elsewhere for a solution to the combination problem.

Thus, I agree with either that the combination problem is successful in blocking several pertinent versions of panpsychism or, at the very least, that the combination problem should be *treated* as successful. It is definitely a strong objection to panpsychism and the constraint that

accepting the conclusion of the combination problem gives to the panpsychist might lead to some very interesting theories. In Chapter IV, I will argue, with this constraint in mind, for a form of panpsychism which avoids the combination problem, preserves the intuition that consciousness is a sharp and determinate concept, and differs in kind from combinatory and emergentist panpsychism, the two versions of the view that are most commonly discussed in the literature.

#### 4.2 Categorical Panpsychism and Combination

In Chapter II, I have discussed the potential for a purely categorical formulation of panpsychism. It consists of the commitment to the idea that only categorical properties exist and that all theoretical roles that those categorical properties play are determined either by the laws of nature or by a necessitation relation obtaining between universals. How might this relate to the combination problem? One option is for the categorical panpsychists to state that the rules behind combination, regardless of whether it is the combination of subjects or qualities, is simply a matter of laws of nature or necessitation relations. It seems like an easy way out: after all, if the standard categoricalist can resort to such commitments, then there is no reason for why the panpsychist could not do the same. Moreover, it seems that the combination problem does not even arise, at least not in its full force, for categorical panpsychism.

However, Coleman's (2014) argument that viewpoints necessarily exclude each other and that they cannot combine in any logical and non-contradictory manner might still be an obstacle. As a response, the categoricalist panpsychist could claim that laws of nature are brute and thus not in need of further explanation or, alternatively, that the production of more complex consciousness at higher levels does not require the summing of subjects. Still, for this to be a

viable form of panpsychism, the standard motivation for categoricism (instead of dispositionalism or hybrid theories of properties) must be endorsed. The view would also inherit all of the problems that categoricism has. The question is: is this too much to ask for a coherent theory of consciousness, that avoids the most pressing objection to panpsychism? The answer, of course, depends on the case for categoricism in general. Still, I find this option interesting enough and I will discuss it in relation to my original proposal in Chapter IV. Ultimately, if it can provide the panpsychist with yet another way of defending their theory, it is worthy of consideration.

## **Conclusion**

This chapter was an extensive overview of the combination problem and proposed attempted solutions. First, I have identified the subject-summing version of the combination problem as the most serious of its formulations, presenting the panpsychist with a seemingly insurmountable challenge. Second, I have criticised the various proposals that panpsychists give in order to avoid or solve the combination problem, siding, in most cases, either with Coleman (2014) or claiming that the discussed solutions are insufficient to deflate the strength of the objection. Third, I addressed both phenomenal unity and potential cases of phenomenal disunity in order to further explore what options the panpsychist has available, though ultimately denying that they are strong options. Finally, after summarising the discussion so far, I have speculatively discussed whether a categorical form of panpsychism might avoid the combination problem, arguing that its commitments might outweigh any benefits to the panpsychist that it might have.

This sets me up for what I will present in the next and final chapter, in combination with the constraints defined in previous chapters. In Chapter I, I have argued that relying on



common arguments against physicalism is not a good way to establish panpsychism, with the possible exception being the vagueness argument. However, as we have seen here, Tye's representational form of panpsychism, based on the vagueness argument, does not conclusively avoid the challenge presented by the combination problem, though it does have other advantages with regard to accommodating our intuition that consciousness is sharp. The limit from Chapter I is thus that I cannot rely on the standard arguments against physicalism when presenting my theory. In Chapter II, the constraint is that an independent argument for panpsychism is necessary in order for the theory to be convincing. I have shown that there are many ways for the panpsychist to make such an argument, though all options will ultimately face the combination problem. So, in Chapter III, I have discussed the combination problem. The constraint here is quite extreme: I ultimately accept that the combination problem is successful, or at least that it is better to treat it as successful.

Therefore, in Chapter IV, I will try to present a new version of panpsychism that respects these constraints, and which will be different from all commonly discussed formulations of the theory. Overall, the main goal of my proposal is to present a kind of panpsychism that is unburdened by the combination problem, that integrates the intuition that consciousness is sharp, and that presents a clear and coherent account of phenomenal consciousness.

## Chapter IV: Monadic Panpsychism

### Introduction

As we have seen in Chapter III, the combination problem, and particularly its subject-summing version, is a great obstacle for proponents of *combinatory* or *constitutive* panpsychism. Their proposal is that microsubjects combine in some way to form macrosubjects and they present various possible models of how that happens, such as *phenomenal bonding* (Goff 2016), where microsubjects can stand in a phenomenal bonding relation which necessitates the weak emergence of a further, more complex subject. In the case of combinatory panpsychism, macro-level facts about consciousness are *grounded* in micro-level facts about fundamental consciousness (Goff, Seager & Allen-Hermanson 2017).

In contrast, *emergentist* panpsychists think that macro-level consciousness, such as in humans or animals, is also fundamental (Goff, Seager & Allen-Hermanson 2017), either in the sense that it co-exists with micro-level facts about fundamental consciousness, which cause it, or alternatively in the sense that micro-level fundamental consciousness ceases to exist in the process of forming macro-level fundamental consciousness, such as in William Seager's (2016) and Hedda Hassel Mørch's (2014) *combinatorial infusion* view, where the microsubjects *fuse* into one big macrosubject.

As an alternative, I have also explored Michael Tye's (2021b) representational form of panpsychism, as well as its attempt to solve the combination problem by arguing that the relevant functional organization of phenomenal ultimates leads to the consciousness of the whole, if and only if those phenomenal ultimates feature in a global workspace. Since this depends on accepting Tye's proposal that consciousness does not necessarily need to be

experienced by persons or subjects, I argued that this view is unappealing. However, the vagueness argument, the motivation behind representational panpsychism, does indicate that consciousness is a sharp and determinate concept, which is an intuition that my form of panpsychism will try to preserve as well.

I have also discussed various ontologies of properties, such as Mørch's (2018) dispositional argument and reconciliation proposal, Dondoni's (2022) modification of her view on the basis of the compound view, as well as a speculative form of purely categorical panpsychism. Objections raised by dispositionalists have also been considered. The original theory that I will argue for here is neutral between the ontology of properties endorsed by Russellian monists, hybrid theories of properties such as the identity theory of powers and the compound view (which might be the same as the property account of Russellian monists), and possibly also categoricism, which I will compare to my theory.

What I will propose is *monadic panpsychism*, an alternative to both combinatory and emergentist panpsychism, as well as to Tye's proposal, inspired by the philosophy of G. W. Leibniz. Both of these options involve the coming about of a new subject, at a higher level of complexity, either through aggregation, fusion, or emergence. The view that I have in mind fully accepts the conclusion of James's combination problem – subjects do not combine – and does not rely on the coming about of a new subject. At the very least, I am prepared to treat the combination argument as succeeding, in order to explore what kind of a panpsychist theory can be built under these constraints. To present my view, I will first introduce the basic commitment that will serve as the basis for my theory. Then, I will describe the resulting forms of panpsychism that utilise it. After discussing some pressing objections and replies, I will list the advantages of my proposal in comparison to other potential theories of consciousness. Finally, I will conclude that my theory avoids some of the most serious issues that other forms of panpsychism face and, additionally, that it has distinct advantages which make it preferable.

## 1. Microphenomenal Structuralism

My ultimate goal is to present a version of panpsychism that does not have to face the subject-summing formulation of the combination problem. I believe this can be done whilst preserving the original motivation behind panpsychism: to integrate consciousness into reality without reducing it to anything else, but in a way that respects the common understanding of causal relations. To clarify, while panpsychists usually ascribe causal power to consciousness, they do it in a way that does not change physics: the electron would still be an electron and behave accordingly regardless of whether or not panpsychism is true. If a panpsychist theory can achieve this without being encumbered by the subject-summing problem, that alone would make it worthy of further consideration.

The commitment to the view that subjects cannot combine differentiates me from many panpsychists. I go the extra mile by accepting Sam Coleman's (2014) argument that it is not only difficult to see how subjects could combine, but that it is positively *incoherent*. I also accept Coleman's arguments against fusion accounts and emergentist panpsychism, as discussed in Chapter III. The challenge now lies in presenting a form of panpsychism that abides by these constraints. When one does not accept the existence of combined, complex macrosubjects, it seems natural to say that the consciousness of each fundamental, simple microsubject depends on its relation to other simples. This is the notion on which I will build my panpsychist proposal:

*Microphenomenal Structuralism.* The phenomenal character of any given microsubject is determined by its relations to other microsubjects featuring in the same relevant causal structure.

At this point, it is important to clarify what a microsubject is in its own right, isolated from all other microsubjects, and what it is when part of said structure. The basic, non-relational or

intrinsic consciousness that a microsubject has in isolation can be thought of in many ways: some rudimentary, non-specific form of consciousness, a point of view without experiences, with empty awareness, the minimal subject, a mere conduit for experiences, phenomenal space, etc. This shares similarities with Tye's (2021b) consciousness\* proposal, which is non-representational basic consciousness, tied to physical ultimates, distinct from representational conscious states. This is fully compatible with the aims of monadic panpsychism as well. A more specific proposal comes from Russellian panpsychism, where the claim is that every quark in the universe has the same quark-appropriate *type* of experience, that every electron has electron-type experience, and so on. In this way, Russellian panpsychists respect physics and causality: quarks are quarks, inside and out, and they cannot be anything else.

However, if I only accepted these basic, intrinsic type-properties, I would have a very limited palette to paint with. How can complex human consciousness be composed of the relatively small number of experiences corresponding to the types of physical ultimates? Worse yet, if microsubjects were only empty points of view without experiences, where would my rich human experience come from? This will be discussed in greater detail later, but for now it is sufficient to acknowledge the problem. To expand the palette, I argue that while microsubjects only have few intrinsic properties, they can have an indefinite range of relational properties. There is the quark-type experience in isolation, but also quark-type-experience-when-part-of-*this*-causal-structure or when part of *that* causal structure and so on. The phenomenal character of each such simple or microsubject thus depends on what position it occupies in the relevant causal structure. That is, the microsubject has its intrinsic quark-type experience, but in addition to that it experiences the relational phenomenal qualities determined by its position and role in the structure. If the microsubject is part of an appropriate causal structure, such as the human brain, it will experience its basic consciousness *and* the full, rich

human experience. The basic consciousness is intrinsic, while the human consciousness is relationally constituted.

What does it mean to say that phenomenal qualities are relational? There is precedent for this in literature on philosophy of mind. A version of what Sydney Shoemaker (1982, 2006) calls the *Frege-Schlick view*, based on (Frege 1956) and (Schlick 1959), postulates that *qualia* (i.e. the qualitative properties of conscious experience, phenomenal consciousness) are relational rather than intrinsic. Furthermore, David Hilbert and Mark Kalderon (2000), as well as Austen Clark (2000), construe qualitative character as relational and argue that “the qualitative character of color experiences is determined by their position in the subject’s color experience space, i.e., by their similarities and difference from other experiences in the repertoire of the subject” (as reported by Shoemaker 2006: 20). Rudolf Carnap (1928) also presented a view akin to phenomenal structuralism, where experiences can be completely described in virtue of relations of phenomenal similarity between them.

These proposals usually wanted to eliminate qualia or to make it compatible with physicalist theories of consciousness. Microphenomenal structuralism also views phenomenal qualities as relational, at the level of fundamental simples, though it admits of the basic non-relational consciousness of microsubjects-in-isolation as well. This basic consciousness is the point of view that I occupy, while the relational structure constitutes the rich and full human experience that I have. Thus, accepting this kind of basic consciousness is what makes the fundamental simples *micro-subjects*, and the theory a form of panpsychism. Since all of this happens at the level of microsubjects, the subject-summing version of the combination problem is avoided.

It is important to note that my proposal is compatible with the Russellian panpsychist commitment to a hybrid view of properties where there are both categorical and relational

properties. I just add one more thing to the relational realm – the phenomenal qualities of our experience – while keeping at the categorical or intrinsic level only the basic consciousness needed for experience *as such*. The phenomenal qualities that are relationally constituted ‘anchor’ themselves, so to speak, in microsubjects, the only difference being that this relational constitution of phenomenal properties has its endpoint at the fundamental level rather than at the level of a combined complex or emergent subject. Despite this, I will also explore how monadic panpsychism relates to other ontologies of properties in the following discussion as well.

## 2. Monadic Panpsychism

As stated, the main aim of microphenomenal structuralism is to avoid the subject-summing version of the combination problem. If the phenomenal character of each microsubject is determined by their mutual relations at the fundamental level of reality, then my proposal can account for full human consciousness without relying on the weak or strong emergence of additional subjects, which is a major advantage for the theory. Based on the discussion so far, I can now introduce the resulting form of panpsychism. Inspired by the philosophy of G. W. Leibniz, *monadic panpsychism* is the claim that microsubjects exist and interact in a hierarchical manner (Strickland 2014). Microsubjects thus resemble Leibniz’s *monads*, which are simple mind-like substances, organised in a hierarchical manner in which one monad is dominant:

“From this we see that each living body has a dominant entelechy, which in the animal is the soul; but the limbs of this living body are full of other living things – plants, animals – each of which also has its dominant entelechy or soul.” (as reported by Strickland 2014: 28)

So, only *one* monad serves as the ‘soul’ of the organism. Here, Leibniz also claimed that this goes on forever, with each subordinate monad having its own dominant soul, which is not something I endorse. To be clear, my proposal is by no means fully committed to Leibniz’s complex metaphysics, but it was the source of the initial thought on which my theory is based. The three versions of monadic panpsychism that I will discuss are static, dynamic, and global monadic panpsychism. All of them are committed to microphenomenal structuralism, where the phenomenal complexity that we associate with human consciousness is relationally constituted without postulating any additional and complex higher-level macrosubjects.

### 2.1 Static Monadic Panpsychism

I will start with the static version since I quickly want to dismiss it. While I am certain that a case could be made even for this option, the sheer strength of the alternatives, in comparison, makes it deeply unattractive. The basic idea is that the dominant microsubject keeps its dominant role permanently:

*Static monadic panpsychism (SMP).* All microsubjects featuring in the relevant structure<sup>23</sup> relationally determine the phenomenal character of only one particular microsubject in a way that gives it a full human experience. That particular microsubject plays the dominant role.

The obvious objection here is: what happens if the dominant microsubject stops being a part of the relevant structure, such as the brain, perhaps due to brain damage? This sounds similar to a view that Leibniz held in his youth, expressed in the doctrine of the *flos substantiae* (‘flower of substance’) in which a person’s soul was a part of a minimal piece of matter, no

---

<sup>23</sup> The relevant relational structure is whatever is minimally necessary to produce my current conscious experience at any given time – a cluster of neurons, regions of the brain, the whole brain, etc.



bigger than a mathematical point, which was located in the centre of the brain (Strickland 2014: 136).

After death, the body would be destroyed but the soul would persist since mathematical points are indivisible and thus indestructible (Strickland 2014: 136). The purpose of this doctrine was to ensure the soul's survival after death. It seems quite unbelievable to think that one microsubject – a consciousness-bearing physical ultimate – permanently contains the 'soul' of the organism, so that if only it was removed from the brain, the whole organism would perish, which is why I will not discuss static monadic panpsychism further.

## 2.2 Dynamic Monadic Panpsychism

To circumvent this issue, consider the dynamic version of the theory, where the dominant role can be played by any microsubject and switch dynamically, perhaps in a manner isomorphic to brain processes:

*Dynamic monadic panpsychism (DMP).* All microsubjects featuring in the relevant structure relationally determine the phenomenal character of only one microsubject in a way that gives it a full human experience. That microsubject plays the dominant role, but which particular microsubject that is can change at any given time.

The conscious experience of the dominant microsubject is relationally determined by its relation to other microsubjects. In the case of the dominant microsubject being removed from the brain, any other microsubject could replace it in its dominant role and become the new endpoint for the relational constitution of a full human experience. Leibniz expressed a similar view:

“But there is no need to suppose, as have some who have misunderstood my thought, that each soul has a mass or portion of matter of its own, or allotted to it forever, and that it consequently possesses other inferior living things which are forever destined to serve it. For all bodies are in a perpetual flux, like rivers, and parts are continually entering and leaving them.” (as reported by Strickland 2014: 28)

The dominant entelechy or soul always remains embodied, but its body is “subject to continual change such that no part of the body is permanently united to the entelechy” (Strickland 2014: 135). This follows from Leibniz’s view that “every created being is subject to change, and consequently the created monad also, and even that this change is continual in each one” (as reported by Strickland 2014: 16). Leibniz adopted this view after abandoning the *flos substantiae* doctrine.

### 2.3 Global Monadic Panpsychism

The third version of the theory moves away from the apparent similarities to Leibniz’s metaphysics. It is a more holistic and interconnected proposal that seeks to explain rich human consciousness primarily through interconnectedness. Specifically, it abandons the idea that only *one* microsubject plays the dominant role:

*Global monadic panpsychism (GMP)*. All microsubjects featuring in the relevant structure relationally determine the phenomenal character of all other microsubjects that make up the structure, so that the phenomenal character of every single one of them is the full human experience.

The result is a very large number of microsubjects who all experience human-level consciousness. Admittedly, this is the most difficult version of the theory to defend. When I am thinking, it is like saying ‘the cabinet is meeting to consider this proposal’ but everyone says the same thing, at the same time. The ‘me’ that I am referring to when I self-identify is like a layer of Ninos who are qualitatively the same. Because of this phenomenal sameness, I experience the many instantiations of ‘I’ as one.

It is not the case that the plurality *as such*, as a whole over and above the individual microsubjects, experiences the richness of my consciousness since this would lead back to the subject-summing version of the combination problem. Rather, it is the case of many individual microsubjects experiencing the same thing, but that multitude is not apparent within my consciousness since there is no phenomenal difference between the microsubjects – they are qualitatively identical. In GMP, all microsubjects contain my full experience, though because they all experience exactly the same phenomenal content, I think of myself as one. There are billions of ‘me’ in my brain thinking the same thought – ‘I really went overboard with this view’ – but because they all think it at the same time, I have the illusion that there is only one ‘me’. It is like listening to a recording of the same melody played on a billion identical pianos in unnaturally perfect sync and pitch. It would sound the same as if it were played on only one piano.

What prevents my proposal from being pure mereological nihilism is precisely this: in both DMP and GMP, the relational structure has an endpoint in one or more microsubjects which experience themselves as one, as a proper whole. This allows for a restricted or ‘moderate’ version of mereological nihilism on which the only proper wholes are organisms. Peter van Inwagen (1990: 115) supports this view: parts compose a whole if and only if the activity of the parts constitutes a life. Similarly, Trenton Merricks (2001) argues for a restricted form of mereological nihilism in which only conscious beings can be regarded as proper

wholes, which is perhaps more pertinent to my proposal. That is, I am presenting two ways, through DMP and GMP, in which conscious beings can be regarded (and regard themselves) as wholes.

Finally, regarding which version of monadic panpsychism I prefer, I opt to argue for a disjunction of the two. Regardless of whether DMP or GMP is more convincing, they are both steps in the right direction when it comes to forming a new conception of panpsychism. However, I do find the idea of a billion or more instances of ‘me’ less intuitively appealing than the idea of the dominant role switching dynamically. Still, I will defend both theories from objections, which will further reveal their potential viability.

### **3. Objections and Replies**

Now, I will extensively discuss some potential objections to monadic panpsychism. Starting from the general objection that monadic panpsychism is counterintuitive, perhaps more so than other versions of panpsychism. Then, I will address what challenge the quality combination problem presents to my view. The question of what the ‘I’ in monadic panpsychism refers to will be explored next. After that, I will discuss the structure combination problem with regard to monadic panpsychism. Finally, I will examine objections specific to DMP and GMP, regarding individuation and symmetry. The aim of this chapter is to defend monadic panpsychism from objections, and hopefully demonstrating that the theory remains viable.

### 3.1 Intuitions Against Monadic Panpsychism

Like with all forms of panpsychism, the immediate objection is that my proposal is counterintuitive or just downright unbelievable. When people first hear of panpsychism, in its general form, the first question is commonly whether rocks are conscious, followed by a stare of disbelief. However, monadic panpsychism might appear as even more counterintuitive than other forms of panpsychism since it claims that human consciousness obtains at the level of microsubjects. Considering that panpsychists also raise objections against physicalism based on intuition, I will take this charge seriously.

The sense of scale is probably what is most suspicious: it is just crazy to say that my rich human experience is contained within one physical ultimate, as per DMP, or that every physical ultimate in my brain is fully me, as per GMP. However, mereologically nihilistic views about the physical which deny that there are proper physical wholes are not seen *as* controversial as panpsychism, so why should mereological nihilism about consciousness, which denies that there are proper mental wholes, be seen as more controversial? In fact, is it not more parsimonious to expect that the mental behaves in alignment with the physical? This is not an endorsement of mereological nihilism about the physical. The point is merely that mereological nihilism is a counterintuitive view, yet it does not face the same level of incredulity as panpsychism does. So, I freely admit that monadic panpsychism is indeed a deeply counterintuitive view, but that does not mean that it is not true.

If we are more ready to accept that there are no tables but just particles arranged table-wise (van Inwagen 1990), then we should grant the same level of acceptance to the idea that there are no non-fundamental subjects but consciousness-involving ultimates arranged human consciousness-wise, forming the relevant causal structure in which microphenomenal structuralism obtains. Monadic panpsychism postulates that the phenomenal character of

microsubjects depends on the relations those microsubjects feature in, within a relevant causal system, and if that is enough for rich human-level consciousness to be the end result, then the view is preferable to forms of panpsychism which face the subject-summing combination problem or depend on emergence. That alone is enough for the theory to be taken seriously, despite it being counterintuitive.

### 3.2 Quality Combination: The Palette Problem

While monadic panpsychism manages to avoid the subject-summing version of the combination problem, it still has to face the issue of how qualities combine. The most pressing aspect of the quality combination problem is the *palette problem*, the question of how the limited palette of microqualities can produce “the vast array of macroqualities, including many different phenomenal colors, shapes, sounds, smells, and tastes” (Chalmers 2017: 183). This problem is especially relevant to Russellian and monadic panpsychists since both argue that the phenomenal qualities experienced by microsubjects correspond to the small number of kinds of physical ultimates, so that there is quark<sup>24</sup>-type experience, muon-type experience, gluon-type experience, and so on<sup>25</sup>. It might be even more serious for monadic panpsychism considering that I accept the possibility that fundamental consciousness is basic and bare, similar in kind to consciousness\* as Tye (2021b) discusses it. The overarching problem pertains to the question of how qualities combine at all, while the more specific palette problem asks how the small number of fundamental qualities leads to the rich qualities experienced in human consciousness.

---

<sup>24</sup> Assuming for the sake of discussions that quarks, muons, and gluons truly are fundamental particles.

<sup>25</sup> To clarify, Chalmers (2017: 183, 189–90) frames it in terms of microqualities corresponding to microphysical properties, such as mass, charge, spin, while I prefer to frame it in terms of fundamental qualities corresponding to physical ultimates. Both formulations ultimately end up with the same asymmetry of quantity: few basic qualities, many rich conscious experiences.

The most straightforward way for the monadic panpsychist to avoid the palette problem is by saying that while there might be few *intrinsic* phenomenal properties, there can be a plethora of *relational* phenomenal properties. This extension of the original Russellian commitment can be done without giving up on the idea that fundamental phenomenal qualities correspond to physical ultimates and without infringing on the causal closure of the physical. While it might be true that a quark in isolation has a quark-type of experience, there is nothing that would violate causal closure in the idea that there is also a quark-type-of-experience-when-part-of-*this*-causal-structure or when part of *that* causal structure. If the causal structure itself is what constitutes and fixes the experience of the dominant microsubject, then there is no possibility of that experience differing if the brain state remains the same. Identical causal structures, such as identical brain states, will produce identical experiences. So, the claim is that there are many relational phenomenal qualities corresponding to the many possible causal structures such as brain states.

There are few intrinsic but many relational phenomenal properties and there is nothing problematic in principle with the idea of the phenomenal properties changing through the interaction of physical ultimates. The causal structure is precisely what determines how the phenomenal qualities will interact and which microsubject's experience will be modified, directly or indirectly, so that it becomes dominant. This solution to the palette problem is directly tied to the metaphysics of monadic panpsychism, where the production of rich human-level consciousness via the thesis of microphenomenal structuralism happens solely at the fundamental level. In the case of DMP, microsubjects constitute the experience of the dominant microsubject by relationally affecting its original experience, but not so that the dominant microsubject experiences the same qualities that those microsubjects have in addition to its original quality. In the case of GMP, the result is that all of the microsubjects in my brain co-constitute each other, so that the experience of each microsubject is relationally affected

indirectly (i.e. not by literally transferring phenomenal qualities from one microsubject to the other) in an equal manner, which results in all of them having the same rich human experience.

To further illustrate the simplicity and appeal of monadic panpsychism, I will compare it to a popular reply to the palette problem. Patrick Lewtas postulates that at least “some basic physical objects [...] simultaneously experience instances of more than one distinct basic experience type” (2016: 757). Using quarks as a stand-in for whatever is fundamental, he further argues: “the quark has a plurality of distinct and wholly separate conscious properties (e.g. red experience, taste-of-brine experience) just as it has a plurality of distinct and wholly separate physical properties (e.g. mass, spin)” (Lewtas 2016: 757). To avoid problems with causal closure, he argues that consciousness is non-causal, at least not in an *active* sense, yet physical nature has the capacity to change itself in response to *passive* conscious properties:

“It sits there as is, and the physical responds to it. The experience doesn’t exert force upon the physical the way a flying brick exerts force on a windowpane. Instead, the physical ‘detects’, ‘reads’, or otherwise ‘picks up on’ the nature/content of the experience and changes its own state accordingly. This keeps the actively causal/functional wholly on the physical side.” (Lewtas 2018: 143–4)

Therefore, if physical ultimates can have many different basic phenomenal qualities and if consciousness is non-causal in the active sense, then there is no threat from the palette problem and causal closure is preserved. This is a non-starter for both the Russellian and the monadic panpsychist since they accept the view that every fundamental physical ultimate *type* has its corresponding fundamental phenomenal *type* – every quark has the same quark-type experience. However, my aim here is not to explicitly argue against Lewtas since I merely want to compare my solution to the palette problem to his proposal. The point is that Lewtas’ view,



within the purview of panpsychism, depends on both a non-standard view of fundamental phenomenal qualities as well as on a non-standard view of causation.

Others went in the completely opposite direction. Roelofs (2014) argues that there is nothing problematic with the notion of a very small number of basic qualities, while Turausky (manuscript) makes the claim that there might be only *one* basic phenomenal quality that can account for human-level consciousness. I will not go into further detail since I only want to make the point that monadic panpsychism does not need to accept any special view of fundamental qualities or causality in order to avoid the palette problem, while also avoiding the subject-summing version of the combination problem. Because of this, I believe that my proposal has the edge.

### 3.3 Structure Combination: The Structural Mismatch Problem

Another major objection to monadic panpsychism is the *structure combination problem*, the most pressing version of which is the *grain problem*. Experiences seem to be smooth and continuous, such as an expanse of red in our visual experience, which is at odds with the discrete and particularised structure of the brain, involving “transfers of or interactions among large numbers of electrons, ions, or the like” (Maxwell 1978: 398). Being smooth and being particulate or discontinuous are structural properties that are mutually incompatible, so at least some mental events seem to “exemplify structural properties that are not exemplified by any brain event” (Maxwell 1978: 398). How can we reconcile this discrepancy? Monadic panpsychists think that the structure of conscious experience is isomorphically related to physical structure, so how does this multitude of microsubjects entwined in a complex connected network interact in order to produce this smooth, continuous expanse? Why are those relations between microsubjects revealed to us through experience? This is different from

the quality combination problem since it asks how phenomenal and physical structural properties are connected, as opposed to asking how qualities combine at solely the phenomenal level.

One potential reply to this objection was discussed by Michael Lockwood (1993), who claims that the grain problem disappears when we consider more recent trends in physics, as opposed to classical physics. He presents an argument based on quantum mechanics:

“[T]here are, in quantum mechanics, no observables, or sets thereof, which are *a priori* privileged. In particular, there is, in terms of quantum-mechanical observables, no rock-bottom level of structure to be discerned in the world. [...] In quantum mechanics there is a sense in which all observables, and in particular observables corresponding to every level of structure, are to be regarded as equal in the sight of God, as are different frames of reference, relativistically conceived.” (Lockwood 1993: 288)

That is, according to Lockwood (1993: 288–9), the world can be structured at other levels too, not just at the level of elementary particles. So, the idea that there is one specific structure which needs to be revealed in conscious experience, or correspond to macrophenomenal experience, is misleading since it relies on classical physics, without considering the possibilities that quantum mechanics provides. Specifically, Lockwood argues that one state can present itself in different ways, under different aspects, which is how he avoids the structural mismatch problem:

“My claim is that, by appealing to the quantum-mechanical concept of an observable, we can render it intelligible, as with the grain of the wood, that a common underlying structure should manifest itself in superficially very different ways. On the side of introspection, moreover, such a conception

removes the need to appeal to any inner representation, distinct from the state itself. For to be directly acquainted with a 'cross-section' of something is *a fortiori* to be directly acquainted with the thing itself, not merely some cognitive surrogate of it – in spite of the fact that what is thereby revealed to consciousness is revealed only under a certain aspect.” (1993: 290)

The brain activity producing a certain phenomenal state is revealed to the subject *as it is in itself* but under a certain point of view (Lockwood 1993: 289). Admittedly, this is very abstract and potentially difficult to understand, but the general argument is that there might not be any structural mismatch if we look at quantum mechanics. This is somewhat similar to the argument presented by William Seager (1995), that particles in the state of quantum entanglement might be responsible for the production of higher-level consciousness. Presumably, quantum entanglement scenarios would rely on a completely different notion of structure than classical physics and could also be considered as different manifestations of the same phenomenon.

Philip Goff (2017) discussed a reply similar to Lockwood since they both argue for a structural *match* between the structure of the brain and the structure of consciousness, though Goff does not rely on quantum mechanics in order to present his argument. Specifically, Goff argues that there is “a vast multiplicity of kinds of consciousness corresponding to a vast multiplicity of structures in the brain” and that this means that “the mystery as to why there is a form of consciousness mirroring a seemingly quite arbitrary brain-structure disappears” since “many macro-level brain structures correspond to phenomenology” (2017: 207). According to Goff, we do indeed find structure in the brain isomorphic with the structure of consciousness if we consider less basic kinds of brain structure, with consciousness corresponding both to more basic and to less basic brain structures (Goff, Seager & Allen-Hermanson 2017). For the proponent of monadic panpsychism, this reply coheres well with the idea that phenomenal and physical structure is isomorphic, especially since Goff seems to emphasise less basic brain

structures (considering that monadic panpsychism delegates consciousness to the fundamental (and only) level of reality).

Daniel Stoljar argues that the grain problem gets the phenomenology wrong. While it might seem plausible to say that the expanse of red in my visual experience is smooth and continuous, it does not follow from this “that the experience *itself* is smooth and continuous” since “an experience of red represents something as being red, but it itself is not red” (2001b: 276). The smoothness – the absence of grain – is a feature of something that experiences *represent* but not a feature of experiences themselves (Stoljar 2001b: 276). Moreover, Stoljar argues that “many acts or states of experiencing seem in a certain respect ‘diaphanous’ to introspection: introspection reveals the intentional objects of experiences to us, but not the experiences themselves” (2001b: 276). All that introspection reveals is that we often have experiences “which represent things as being smooth and continuous” (Stoljar 2001b: 276), but this is different from saying that the experiences *themselves* are smooth and continuous. I think this response to the grain problem could be attractive to all panpsychists who have to face the issue, considering that it shows how, at least in principle, we can provide an answer.

### 3.4 A Pilot in a Ship

This objection is based on a simple question: what does the ‘I’ refer to? In monadic panpsychism, there seem to be at least two things the ‘I’ can refer to: to the microsubject or to the organism as a whole. When I utter ‘I am thinking about London’, it seems that the ‘I’ refers to the microsubject, the thinking thing. In contrast, when I utter ‘I am raising my arm’, the ‘I’ seems to refer to the body, to the whole organism. If I am a microsubject – if the locus of my entire rich human experience is in a microsubject – it appears that I am correct when I say that I am *thinking* about something but mistaken when I say that I am *doing* something or when I

am describing my body. I might have the thought or the intention that I am raising my arm, but if I am just a microsubject in my brain, then strictly speaking *I* am not raising my arm. Conversely, if the ‘I’ refers to the organism, then the organism is not thinking about London. So, who am ‘I’? The microsubject, the brain, the head, the neural system, the body minus the right leg, etc. (Strawson 1959/2003)?

This objection relates to views frequently discussed in literature on personal identity. The debate is often framed as being between *animalists* (Olson 2007, Snowdon 2014), who identify the ‘I’ with the organism, and psychological continuity theorists (Parfit 1971, Shoemaker 1984), who argue that the persistence of the ‘I’ is based on direct psychological connections between mental states. However, it is important to note that the objection raised against monadic panpsychism is not strictly about persistence conditions for personal identity, but quite simply about what the subject is and what its use of ‘I’ refers to.

In monadic panpsychism, the microsubject is the locus of my entire rich human experience. There is the microsubject which, taken in isolation, is just a point of view with a faint glimmer of consciousness. There is also the constantly changing flux of phenomenal qualities that constitute my experience within that microsubject. When I say, ‘I am Nino’, am I referring to the microsubject, to the flux of experiences, or to the organism? What occupies the point of view is the microsubject itself to which the phenomenal qualities of the whole system are presented. Thus, if the ‘I’ refers to the microsubject, to the thing that has my point of view, then I *am* one of those microsubjects in DMP, and ‘I’ am the name for the billions of qualitatively identical microsubjects in GMP, as was discussed in Section 2.3.

So, when I say that I am raising my arm, it does not seem to be true since the ‘I’ refers to a microsubject, not to the body. In contrast, when I say I am thinking of London, it does seem to be true since the microsubject is the thinking, experiencing thing. How can the monadic

panpsychist respond to this objection? After all, a common intuition that we have about ourselves is that we are *embodied* subjects. Descartes presented this problem poignantly:

“Nature likewise teaches me, through these very feelings of pain, hunger, thirst, and so forth, that I am not present in my body only *as a pilot is present in a ship*, but that I am very closely conjoined to it and, so to speak, fused with it, so as to form a single entity with it. For otherwise, when the body is injured, I, who am nothing other than a thinking thing, would not feel pain as a result, but would perceive the injury *purely intellectually*, as the pilot perceives by sight any damage occurring to his ship.” (Descartes 2008: 57; emphasis mine)

The animalist’s argument seems to encapsulate this intuition that we are related to our bodies in a special way, though it goes a step further by identifying Descartes’ thinking thing with the animal itself:

1. A human animal is sitting in my chair.
2. That human animal is thinking.
3. I am the thinking animal.

Therefore, the human animal sitting in my chair is me.

This is the *thinking animal argument*, also known as the problem of *too many thinkers*. If the animalist is correct, when I am in pain, I do not perceive the damage merely intellectually – I *am* damaged. Psychological continuity theorists argue that such continuity is a persistence condition for personal identity, but their view is fully compatible with the idea of a subject being essentially embodied. In Maurice Merleau-Ponty’s phenomenology, the subject is inseparable from the body since the essence of subjectivity is “bound up with that of the body and that of the world [...] because my existence as subjectivity is merely one with my existence

as a body and with the existence of the world” (1962: 408). In contrast, P. F. Strawson (1959/2003: 103) claimed that the concept of a person should not be analysed in terms of an animated body or of an embodied *anima*, which is why he thought that disembodied subjectivity is conceivable. However, despite that, he argued that the concept of a person necessarily involves both mental and bodily aspects, both being the condition required for the successful ascription of subjectivity to ourselves and others (Strawson 1959/2003: 104). P. F. Strawson agreed with Descartes, that we are not pilots in ships, and argued in many ways that our perceptual experiences are causally dependent on our bodies (as reported by Lewis 1962: 6–7).

In monadic panpsychism, this intuition is not upheld since the ‘I’ does not refer to the organism, to the animal looking through its eyes and seeing the world. Rather, the ‘I’ refers to the dominant microsubject in DMP, buried somewhere in the brain, or to all of the microsubjects having the same synchronised experience in GMP. There is a vague sense in which the microsubject is embodied, as simply being a part of the organism, but it is not embodied in the sense of the aforementioned phenomenological intuition, as a living and breathing body that feels pain and pleasure directly. That is, it seems that the ‘I’ in monadic panpsychism is indeed a *pilot in a ship*.

What can the monadic panpsychist say in defence? Using the arm-raising and thinking-about-London examples again, I can think of three possible replies:

- (a) Biting the bullet and admitting that ‘I am raising my arm’ is not literally true, but that it has to be translated in some sense (e.g. ‘I am thinking or intending that I am raising my arm and then observing the body raising the arm’), while ‘I am thinking about London’ is literally true;

- (b) Arguing that the microsubject is embodied in a more robust sense since it causally features in the structure that makes up the brain (and the rest of the body);
- (c) Expanding what the ‘I’ refers to, such as saying that the microsubject in isolation alone is not the ‘I’, but that the relevant causal structure (the brain) is what the ‘I’ refers to, so that both arm-raising and thinking-about-London can be performed by the subject.

Regarding (a), it is possible to say that I am a microsubject that *has* a body. This sense of possession can be based on the microsubject’s causal connection to the body. After all, it features in the causal structure which is a part of the body, so the microsubject is *in* the body as a proper part. H. D. Lewis (1962: 5), in his commentary on P. F. Strawson’s *Individuals*, mentions that a person can say ‘I went to open the door’ in ordinary language, though we should properly distinguish the mental and the physical activity: ‘I was intending to open the door and my body moved towards it’. Similarly, he argues that the utterance ‘I am bald’ is perfectly apt for normal purposes, but that the strict truth is that ‘my *head* is bald’:

“I should also wish to add that my real self is my mind, and that it is only in a derivative and secondary sense that my body is said to be myself at all. In other words, in the strict sense I am not bald at all, and cannot be; it is only part of my body that can be bald, my body is not something that I *am* but something that I *have*, and here linguistic usage, if that were what we should appeal to, is on our side.” (Lewis 1962: 8)

This is admittedly a counterintuitive view when compared to our everyday normal phenomenology, but it is certainly an option for the monadic panpsychist. I do not think it avoids the ‘pilot in a ship’ since the microsubject is in the body that it has, just like the pilot is in the ship that they own. The microsubject then has a false belief that it is raising its arm, though that belief can be translated into a true statement, as suggested by H. D. Lewis.



The option (b) focuses on the fact that a pilot *commands* the ship. The microsubject is a part of the causal structure that causes the arm to raise. No matter how small its contribution, it at least ‘micro-causally’ affects the chain of events leading to the arm raising. Similarly, when the body is damaged, there is a causal chain from the point of impact to the microsubject, so that the microsubject *does* feel the pain rather than just perceiving the injury purely intellectually, as Descartes feared. In DMP, the causal contribution of the dominant microsubject might be minute, though in GMP, the causal contribution can be attributed to the whole causal system relevant for arm-raising. So, while the ‘I’ strictly refers to a microsubject or a group of them, the ‘I’ is embodied in a more robust sense, which makes it possible to say that I am both the thinking thing and the arm-raising thing. This is because the ‘I’ is a part of the causal chain leading to the arm-raising.

Option (c) is a further step up. The ‘I’ can be redefined to refer to the whole brain, assuming the brain is the relevant causal structure for both thinking and acting. While my rich human experience is located in one dominant microsubject in DMP and in a synchronised group of microsubjects in GMP, that experience necessarily depends on other microsubjects, as well as the relations that obtain between them. Therefore, what ‘Nino’ refers to cannot strictly be any microsubject taken in isolation, since what I call ‘Nino’ is constituted by the entire brain. The existence of the relevant causal system, such as the brain, is a necessary condition for ‘Nino’ to exist at any point in time, in both DMP and GMP. If the ‘I’ then refers to the brain, it seems fully coherent to say that I am both thinking about London and raising my arm; that the brain is thinking about London and that the brain is raising the arm. Similarly, Jeff McMahan (2002: 93–94) and Derek Parfit (2012: 6–7) argue against the thinking animal argument that the person or the ‘I’ is the brain since the brain is what truly, non-derivatively thinks, while the animal only thinks derivatively since it has a brain as its proper part.

Naturally, there are several objections that can be raised against all of these options. Option (a) offends our intuition, option (b) still keeps the pilot in the ship, though it gives them more control, while option (c) does not necessarily avoid the problem of too many thinkers (see Hershenov 2013). Still, I believe option (b) is at least defensible, especially for GMP due to the greater potential causal contribution of *more* microsubjects, while option (c) is attractive because it captures the natural intuition that our brain is the seat of our consciousness – what one refers to when one says ‘I’ or ‘me’ – quite well. However, the monadic panpsychist does not necessarily have to provide a direct reply to this objection. The problem of too many thinkers arises for many different theories of personal identity and the self (Hershenov 2013), so it is not necessarily a different or more serious issue for monadic panpsychism.

### 3.5 Individuating the Dominant Microsubject in DMP

One objection aimed at Dynamic Monadic Panpsychism (DMP) specifically is the question of how to individuate the dominant microsubject? Let us say that the current dominant microsubject of my brain is a quark, assuming for the sake of discussion that quarks are fundamental particles. Presumably, that quark is no different from any other quark in the universe. According to panpsychism, that quark is also not different when it comes to experience – all quarks have the same quark-type experience. Panpsychists want an ontologically sufficient condition for consciousness in the form of non-observable entities that do not go against the general scientific worldview but rather do the metaphysical heavy lifting (Brüntrup 2016: 59–60). That is, panpsychists generally want to preserve a commitment to the causal closure of the physical.

However, according to DMP, the quark that plays the dominant microsubject role is experientially different from other quarks – it contains the totality of my conscious experience

at a specific time. How do we know which quark in my brain is the dominant microsubject? And if quarks can have different experiences at different times, does that go against the idea that all quarks share the same *type* of experience? More broadly, the worry is that this implies that the same brain state could be experienced differently at different times, thus going against causal closure.

Regarding the first question, I am willing to bite the bullet and simply say that we do not or cannot know which physical ultimate plays the dominant microsubject role – at least not ‘from the outside’, without being that dominant microsubject. The dominant microsubject is individuated from other microsubjects purely phenomenologically, ‘internally’, without any measurable external sign that it is dominant, apart from featuring in the currently active firing of neurons. So, the only way to know which physical ultimate is the dominant microsubject is by *being* that microsubject, though that is also indirect and opaque. Obviously, it does not reveal to us the specific quark, but only gives us the trivial ability to say: “it’s this one – me!”

At most, we can potentially distinguish which neuronal structures are correlated with general consciousness, the current experience of the subject, particular memories, etc. For example, there are reasons to think that the claustrum – “a thin, irregular, sheet-like neuronal structure hidden beneath the inner surface of the neocortex in the general region of the insula” – is connected to integrating information at a fast time-scale in the brain and thus relevant to the production of consciousness (Crick & Koch 2005: 1271). This does not tell us which quark is the dominant microsubject, of course, but it brings us a tiny step closer.

### 3.6 GMP and Symmetry

Regarding Global Monadic Panpsychism (GMP), the objection is this: why should we suppose that each microsubject has the same symmetrical relation to other microsubjects? If

all of the microsubjects of my brain reflect each other and so produce my consciousness, it seems necessary to postulate a symmetrical relation through which all of them can reflect each other, so that each and every one has the totality of my experience. Are there any empirical reasons to think that there is such a symmetry in the brain? It might be that some neurons (and thereby some physical ultimates serving as microsubjects) are simply not at all times connected to all other neurons. In fact, brain *asymmetry* at the structural, functional, and behavioural levels is a well-known phenomenon in cognitive science and neuroscience (Toga & Thompson 2003). How can the proponent of GMP ensure that there is phenomenal symmetry without going against causal closure, especially considering the commitment to the isomorphism of physical and phenomenal structure?

There are three ways of replying to this objection. The first one is denying that the brain as a whole needs to be symmetrical. It is asymmetrical and that is an empirical fact. Instead, only the physical ultimates involved in producing an experience at a given time, such as my current experience, need to be symmetrical in the relevant sense. While the whole brain contains the totality of my experience, such as my memories and inclinations, all of that is not *always* active, at the same time, when producing my current experience. The relevant portion of my brain producing the current experience is the only set of physical ultimates that needs to be symmetrical. However, neuronal activity involved in producing some experiences can encompass various regions of the brain in an asymmetrical manner. So, symmetry is not something the proponent of GMP can count on.

But why should symmetry be the crucial notion for microphenomenal structuralist relations to obtain in GMP? The second way of replying to the objection is by denying the need for symmetry. If such microphenomenal structuralism is conceived of as the direct or indirect sharing or ‘reflecting’ of experiences, all that it takes for it to happen is that experiences can ‘pass through’ all of the physical ultimates relevant to the production of *this* experience. That

is, the neurons that produce my current experience need simply be connected and that is ensured by the causal structure they feature in, the brain. The relations pertaining to microphenomenal structuralism can ‘ripple’ through the activated neurons, creating a feedback nexus that leads to the current experience being reflected in all of the involved physical ultimates. Even if the neurons are not obviously causally linked, the physical ultimates at the fundamental level could be.

This leads to the third reply: symmetry or connectedness at the fundamental level need not appear as symmetry or connectedness at the neuronal level. Microphenomenal structural relations, as I described them, are relations that happen at the fundamental level. The phenomenal structure involved in those relations is isomorphic with causal structure at the *fundamental* level, so there is no reason to think that causal structure at higher levels should necessarily *appear* to us as symmetrical or connected, nor to think that its apparent asymmetry or disconnectedness at higher levels necessarily implies asymmetry or disconnectedness at the fundamental level. In addition, if a purely physicalist worldview can have physical ultimates or neurons that appear asymmetrical or disconnected yet produce a unified conscious experience, then this is not a special problem for GMP. Indeed, the hope of finding the possible ‘seat of consciousness’ within neuronal structures which integrates information across the brain, such as the claustrum, is an ongoing research project in neuroscience. The future result need not affect GMP in any way different from how it would affect a purely physicalist theory of consciousness. If there is a connection between the regions of the brain involved in producing a specific conscious experience at a given time, then both the physicalist and the proponent of GMP can integrate that into their respective theories.

## 4. Discussion

In this section, I will briefly compare monadic panpsychism to Tye's representational panpsychism, as well as to the more speculative proposal of purely categorical panpsychism, which I have discussed in previous chapters. The aim of this section is to show how my proposal is advantageous over Tye, as well as to present a possible back-up plan, so to speak, in the form of purely categorical panpsychism. Though, the strength it will depend largely on how strong motivations for categoricism are, and the theory will also have to face all of the issues that categoricists face.

### 4.1 Tye's Representational Panpsychism vs. Monadic Panpsychism

Tye (2021b) distinguishes between consciousness\* and representational conscious states, which is similar to an option I also accept as plausible, in which there is basic consciousness that microsubjects have, as well as the phenomenal qualities that they experience which are fully relationally constituted. This distinction is also what Tye relies on in order to present his answer to the combination problem. He connects it to the global workspace theory, as I have discussed in Chapter III, and argues that a proper functional organization of phenomenal ultimates will result in the complex whole inheriting consciousness, if those phenomenal properties behave as a global workspace.

Tye's proposal certainly captures the intuition that consciousness is a sharp and determinate concept. Consciousness\* is non-representational and fundamental by definition, and it is tied to physical ultimates, so there is no danger of vagueness entering into the picture from anywhere. However, whether his proposal is successful against the pressure from the combination problem, particularly the subject summing problem, is questionable. As we have

seen in Chapter III, Tye (2021b) requires us to accept the counterintuitive and perhaps indefensible notion that consciousness is not necessarily possessed by persons, subjects, points of view, etc.

So, if monadic panpsychism can preserve the intuition that consciousness is a sharp concept whilst also avoiding the subject summing version of the combination problem, then it is a better option than Tye's representational panpsychism. I argue that it does both. As I stated earlier, it employs a similar strategy when it comes to differentiating between a basic form of consciousness and a qualitative or representational form of consciousness, at least as one of the options monadic panpsychism is compatible with, so it accommodates the sharpness intuition just as well as Tye does. In the case of the subject combination problem, the whole point of this chapter was to argue that monadic panpsychism avoids it. Microsubjects have a phenomenal character that is determined by the relations they feature in, in a relevant causal structure, so no combination, aggregation, fusion, or anything of the sort is necessary in order to explain how rich human consciousness comes about. Because of this, I conclude that monadic panpsychism is indeed preferable to representational panpsychism.

#### 4.2 Categorical Panpsychism and Monadic Panpsychism

Recall Chapter III. One of the ways in which I criticised purely categorical panpsychism as a way of responding to the combination problem is because of the problematic notion that laws of nature or necessitation relations should determine and thus solve how small subjects combine to form large subjects. However, microphenomenal structuralism provides a way of avoiding this objection: subjects do not need to combine, and the laws of nature or necessitation relations need only fix the microphenomenal structural relations between the

microsubjects, which would result in either DMP or GMP. So, the framework behind monadic panpsychism strengthens the case for categorical panpsychism as well.

However, even so, any proponent of categorical panpsychism would have to motivate the categoricalism itself first, which means providing an answer to the common objections raised against categoricalism. For example, some often criticised aspects of categoricalism are its dependence on a Humean view of causality, in some cases, as well as its commitment to the notion that categorical properties are only contingently related to the theoretical roles that they play. That would mean that the same conscious state might be connected to different theoretical roles in other possible worlds. For example, the feeling of pain might be related to smiling instead of wincing and crying. Still, this is a problem that the non-panpsychist categoricist has to face as well, so it is unclear why it would be a more serious objection for categoricist panpsychism. Ultimately, if the theory can provide a coherent account of consciousness that avoids the combination problem, preserves the intuition that consciousness is sharp and determinate, and integrates it fully into our general worldview, then it is worthy of considerations. I also consider this as a backup plan for monadic panpsychism, if its commitment to a hybrid theory of properties, where the categorical is consciousness-as-such/basic or bare consciousness/or consciousness\* as Tye (2021b) calls it, and where the dispositional is equivalent to the usual definition of dispositions in terms of stimulus conditions and manifestations, is proven to be problematic.

## **Conclusion**

This chapter began with the introduction of microphenomenal structuralism, which is a relation between microsubjects through which their phenomenal characters are relationally determined. Then, I presented the two viable forms of monadic panpsychism that result from



microphenomenal structuralism: dynamic and global monadic panpsychism. Through the process of discussing objections and replies, I have further clarified the theory, explaining its commitment to the isomorphism between the physical and the phenomenal, which proved to be a strong resource in overcoming several pressing issues. In addition, I have discussed several common replies to the quality and structure combination problems that the monadic panpsychist can also employ. Then, I examined certain objections that are specific to monadic panpsychism. Finally, I compared the view to Tye's representational panpsychism and argued that my proposal has distinct advantages.

To conclude, I think that the main advantage of my proposal is that it is able to adapt to various problems while still remaining appealing and coherent. Though it is admittedly a very counterintuitive and speculative view, it avoids the serious subject-summing version of the combination problem for panpsychism, which is why monadic panpsychism retains its appeal. Furthermore, it retains the intuition that consciousness is a sharp and determinate concept since it employs a strategy similar to Tye's (2021b) proposal, but without its more problematic commitments. This theory was also presented while following the constraints outlined in the previous chapters. The first chapter has shown that standard anti-physicalist arguments are inconclusive, so I have not relied on them to make my case for panpsychism. The vagueness argument demonstrated that we need to take the intuition that consciousness is a sharp and determinate concept seriously, which my theory accommodates. The second chapter presented various different arguments for panpsychism, resting on the full spectrum of ontologies of properties, and hopefully it has been successful in showing that there are many options for the panpsychist to make their case. Chapter III had the strongest constraint of them all: either the combination problem successfully refutes many versions of panpsychism, or it should be treated as successful, due to its force.

Based on these constraints, I presented monadic panpsychism as an all-in-one solution, offering two potentially viable formulations of the theory. These formulations can serve as bases for future discussions, which could move us away from combinatory and emergentist views. The contribution of my proposal is thus that it offers an alternative to already existing theories, a third way forward which avoids several crucial issues for panpsychism.

## General Conclusion

Throughout this work, I tried to defend panpsychism in various ways. Chapter I gave me the constraint that I cannot rely on standard arguments against physicalism, such as Nagel's perspective-based framework, Jackson's knowledge argument, and Chalmer's conceivability-to-possibility nexus. The phenomenal concept strategy, as a general move against anti-physicalist argument, motivated that constraint and, in turn, it led to the vagueness argument as an attractive option that the panpsychist can accept – one that is disconnected from the standard anti-physicalist arguments and one that does not seemingly suffer from the phenomenal concept strategy. This also informed my aims in the sense that I wanted to present a formulation of panpsychism that preserves the idea that consciousness is a sharp and determinate concept as well.

Chapter II hopefully demonstrated the strength of independent arguments for panpsychism, most notably Russellian monism, and Mørch (2020) and Dondoni's (2022) attempts at reconciling the standard argument with their respective views. Tye's (2021b) representational panpsychism was also discussed, as a unique take on the theory. Cosmopsychism, the top-down ontology of consciousness on which 'medium' minds, such as human minds, derive from global consciousness, was also discussed as a potential alternative to Russellian monism and Tye's (2021b) proposal, as well as to combinatory panpsychism in general. One interesting option in the form of purely categorical panpsychism was also explored, as a 'backup plan' of sorts, if other arguments for panpsychism fail.

The constraint from Chapter III was the most significant: the combination problem either succeeds or it should be treated as successful in blocking various combinatory forms of panpsychism. All of the various combination problems present a strong case against panpsychism, though only the subject summing version of it seems to be truly insurmountable. So, despite discussing all the various attempts at providing an answer to the combination

problem, I decided to formulate my original proposal based on the assumption that subjects cannot combine and that it is clear *why* they cannot combine. Here, I also mentioned purely categorical panpsychism, as a potentially simple answer to the combination problems, despite its heavy metaphysical commitments.

Based on that, in Chapter IV, I introduced monadic panpsychism, according to which a basic form of consciousness is fundamental, while representational or qualitative consciousness is construed relationally. That is, the phenomenal characters of microsubjects are determined by the relations they feature in. I have shown that this does not suffer from problems raised by the vagueness argument, quite simply because I adopt a similar strategy that Tye (2021b) uses in order to describe his theory of representational panpsychism, in which he distinguishes consciousness\* from representational consciousness. However, as I argued, since subjects do not combine at all under my view, while that either happens in the case of Tye's proposal or it demands for us to accept counterintuitive premises, my theory should be preferred. In addition, monadic panpsychism also avoids the combination problem since microsubjects are the true bearers of rich consciousness. Under monadic panpsychism, there is no need for subjects combining, at any level, for the rich human consciousness to obtain, for example. Wherever the relational chain that leads to a unified conscious experience ends – in whichever part of the brain or microsubject – it will necessarily be *experienced*. Any microsubject or ultimate can thus provide the basic consciousness or the most fundamental point of view necessary for experiencing as such.

Here is a brief summary of the argument that I presented, based on what was discussed in all chapters, though most importantly in Chapters III and IV:

1. Standard arguments against physicalism are inconclusive, so we need independent arguments for panpsychism.

The most forceful way of establishing this is through the phenomenal concept strategy. As I have argued, all anti-physicalist arguments suffer from the danger of collapsing into purely epistemological views. That is, the strategy blocks their inference from epistemological premises to an ontological conclusion.

2. Subjects do not combine; it is clear that subject summing is demonstrably incoherent.

I accepted the conclusion of the combination problem, siding with Coleman (2014), who argues that it can be seen how subjects *cannot* combine, and how that is incoherent and illogical.

3. The phenomenal character of any given microsubject is determined by its relations to other microsubjects featuring in the same relevant casual closure.

This is what I called *microphenomenal structuralism* and it is the basis on which I present monadic panpsychism, in both its dynamic and global forms. The motivation behind this is to create a version of panpsychism that is completely unphased by the combination problem, particularly by subject-summing, though I have also addressed how it avoids the quality combination problem as well.

4. Monadic panpsychism can accommodate the intuition that consciousness is sharp and determinate.

Just like Tye's (2021b) representational form of panpsychism, I tie consciousness to the fundamental, most basic level of reality. Similarly, I also distinguish between basic consciousness (or consciousness\*, as Tye calls it), and phenomenal qualities (or representational consciousness, as Tye calls it). The difference is that my proposal does not demand the acceptance of unattractive notions, such as that consciousness is not necessarily had by persons or subjects. So, monadic panpsychism preserves the intuition that consciousness is sharp in a clearer and more direct manner.

Considering everything that has been said, I believe that monadic panpsychism truly is an interesting and significant contribution to the field, which might open up routes towards other new formulations of the view. If it is successful in avoiding the combination problem, then it is essentially successful in vindicating panpsychism as a serious alternative, considering how pressing the challenge from the combination problem is. If it is also successful in integrating the intuition that consciousness is sharp and determinate, that is merely a bonus that distinguishes it and gives it an advantage over other view. Moreover, I believe that the view I presented is a good overall theory of consciousness since it does not invoke a large range of premises. It is also a consistent and robust view that has the potential of being developed into many different types of views. So, ultimately, I hope to have achieved a broader goal: that of justifying panpsychism as a research project worthy of philosophical interest.

## References

- Alter, T., Coleman, S. (2020) 'Panpsychism and Russellian Monism', in W. Seager (ed.) *The Routledge Handbook of Panpsychism*, 230–42. New York: Routledge.
- Antony, M. V. (2006a) 'Vagueness and the Metaphysics of Consciousness', *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 128 (3): 515–538.
- Antony, M. V. (2006b) 'Papineau on the Vagueness of Phenomenal Concepts', *Dialectica*, 4: 475–483.
- Armstrong, D. (1997) *A World of States of Affairs*. Cambridge: Cambridge University Press.
- Armstrong, D. (1983) *What Is a Law of Nature?* Cambridge: Cambridge University Press.
- Armstrong, D. (1978) *A Theory of Universals*. Cambridge: Cambridge University Press.
- Bayne, T., Montague, M. (eds.) (2011) *Cognitive Phenomenology*. Oxford: Oxford University Press.
- Bayne, T. (2010) *The Unity of Consciousness*. Oxford: Oxford University Press.
- Bird, A. (2007) *Nature's Metaphysics: Laws and Properties*. Oxford: Oxford University Press.
- Brook, A., Raymond P. (2017) 'The Unity of Consciousness', *Stanford Encyclopedia of Philosophy*. <URL=<https://plato.stanford.edu/entries/consciousness-unity/>>.
- Brook, A., Raymond P. (2009) 'Unity of Consciousness', in B. P. McLaughlin, A. Beckermann and S. Walter (eds.) *The Oxford Handbook of Philosophy of Mind*. Oxford: Oxford University Press.
- Brüntrup, G. (2016) 'Emergent Panpsychism', in G. Brüntrup and L. Jaskolla (eds.) *Panpsychism: Contemporary Perspectives*: 48–71. New York: Oxford University Press.
- Carnap, R. (1928) *Der logische Aufbau der Welt*. Berlin: Bernary. Translated by R.A. George (1967) *The Logical Structure of the World*. Berkeley: University of California Press.
- Chalmers, D. (2020) 'Spatiotemporal functionalism v. the Conceivability of Zombies', *Nous*, 54 (2): 488–497.
- Chalmers, D. (2017) 'The Combination Problem for Panpsychism', in G. Brüntrup and L. Jaskolla (eds.) *Panpsychism*, 179–214. Oxford: Oxford University Press.
- Chalmers, D. (2013) 'Panpsychism and Panprotopsychism', *The Amherst Lecture in Philosophy*, Lecture 8: 1–35. <URL=<http://www.amherstlecture.org/chalmers2013/>>.
- Chalmers, D. (2010) *The Character of Consciousness*. New York: Oxford University Press.
- Chalmers, D. (2007) 'Phenomenal Concepts and the Explanatory Gap', in T. Alter, S. Walter (eds.) *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, 167–194. New York: Oxford University Press.
- Chalmers, D. (1996) *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- Choi, S., Fara, M. (2018) 'Dispositions', *Stanford Encyclopedia of Philosophy*. <URL=<https://plato.stanford.edu/entries/dispositions/>>.

- Clark, A. (2000) *A Theory of Sentience*. Oxford: Oxford University Press.
- Coleman, S. (2014) 'The Real Combination Problem: Panpsychism, Micro-Subjects, and Emergence', *Erkenntnis* 79 (1): 19–44.
- Coleman, S. (2012) 'Mental Chemistry: Combination for Panpsychists', *Dialectica*, 66 (1): 137–66.
- Crane, T. (1993) 'Reply to Pettit', *Analysis*, 53 (4): 224–227.
- Crane, T., Mellor, D. H. (1990) 'There is No Question of Physicalism', *Mind*, 99 (394): 185–206.
- Crick, F. C., Koch, C. (2005) 'What is the function of the claustrum?', *Philosophical Transactions of the Royal Society B*, 360: 1271–1279.
- Dainton, B. (2000) *Stream of Consciousness: Unity and continuity in conscious experience*. London and New York: Routledge.
- Dalal, N. (2021) 'Śaṅkara', *Stanford Encyclopedia of Philosophy*.  
<URL=<https://plato.stanford.edu/entries/shankara/>>.
- Descartes, R. (2008) *Meditations on First Philosophy with Selections from the Objections and Replies*. Translated with an Introduction and Notes by Michael Moriarty. Oxford: Oxford University Press.
- Dondoni, L. (2022) 'Towards a Unitary Case for Russellian Panpsychism', *Philosophia*, (50): 43–64.
- Eddington, A. (1928) *The Nature of the Physical World*. Cambridge: Cambridge University Press.
- Ellis, B. (2013) 'The categorical dimensions of the causal powers', in A. Bird, B. Ellis, & H. Sankey (eds.) *Properties powers structures*, 11–27. London: Routledge.
- Frankish, K. (2007) 'The Anti-Zombie Argument', *Philosophical Quarterly*, 57: 650–666.
- Frege, G. (1956) 'The Thought, A Logical Inquiry', *Mind*, 65 (259): 289–311.
- Fürst, M. (2011) 'What Mary's Aboutness is About', *Acta Analytica*, 26: 63–74.
- Goff, P., Coleman, S. (2020) 'Russellian Monism', in U. Kriegel (ed.) *Oxford Handbook of the Philosophy of Consciousness*. Oxford: Oxford University Press.
- Goff, P., Seager, W., Allen-Hermanson, S. (2017) 'Panpsychism', *Stanford Encyclopedia of Philosophy*. <URL= <https://plato.stanford.edu/entries/panpsychism/>>.
- Goff, P. (2017) *Consciousness and Fundamental Reality*. New York: Oxford University Press.
- Goff, P. (2016) 'The Phenomenal Bonding Solution to the Combination Problem', in G. Brüntrup & L. Jaskolla (eds.) *Panpsychism: Contemporary Perspectives*, 283–302. New York: Oxford University Press.
- Goff, P. (2009) 'Why Panpsychism Doesn't Help Us Explain Consciousness', *Dialectica* 63 (3): 289–311.
- Hawthorne, J. (2002) 'Advice For Physicalists', *Philosophical Studies*, 109 (1): 17–52.



- Heil, J. (2012) *The Universe As We Find It*. Oxford: Oxford University Press.
- Heil, J. (2003) *From an Ontological Point of View*. New York: Oxford University Press.
- Hempel, C. (1969) 'Reduction: Ontological and Linguistic Facets', in S. Morgenbesser, *et al.* (eds.) *Essays in Honor of Ernest Nagel*, 179–199. New York: St Martin's Press.
- Hershenov, D. B. (2013) 'Who Doesn't Have a Problem of Too Many Thinkers?', *American Philosophical Quarterly*, 50 (2): 203–8.
- Hilbert, D. R., Kalderon, M. E. (2000) 'Color and the Inverted Spectrum', in S. David (ed.) *Vancouver Studies in Cognitive Science*, 187–214. New York: Oxford University Press.
- Howell, R. J. (2015) 'The Russellian Monist's problems with mental causation', *Philosophical Quarterly*, 65, 22–39.
- Hurley, S. (1994) 'Unity and Objectivity', *Proceedings of the British Academy* 83: 49–77.
- van Inwagen, P. (1990) *Material Beings*. Ithaca, NY: Cornell University Press.
- Jackson, F. (1982) 'Epiphenomenal Qualia', *Philosophical Quarterly*, 32: 127–136.
- Jacob, P. (2019) 'Intentionality', *Stanford Encyclopedia of Philosophy*. <URL=<https://plato.stanford.edu/entries/intentionality/>>.
- James, W. (1911) *Some Problems of Philosophy*. New York: Longmans, Green & Co.
- James, W. (1909) *A Pluralistic Universe: Hibbert Lectures at Manchester College on the Present Situation in Philosophy*. New York: Longmans, Green and Co.
- James, W. (1890) *The Principles of Psychology*. New York: Dover Publications, Inc.
- Jaworski, W. (2016) *Structure and the Metaphysics of Mind*. Oxford: Oxford University Press.
- Kadić, N. (2017) 'The Grounding Problem for Panpsychism and the Identity Theory of Powers', *Croatian Journal of Philosophy*, 17 (49): 45-55.
- Kind, A. (2015) 'Pessimism about Russellian Monism', in T. Alter and Y. Nagasawa (eds.) *Consciousness in the Physical World: Perspectives on Russellian Monism*, 401–421. New York: Oxford University Press.
- Leibniz, G. W. (1902/2005) *Discourse on Metaphysics and The Monadology*. Translated by George R. Montgomery. Mineola, NY: Dover Publications, Inc.
- Leibniz GW (1704/1981) *New Essays on Human Understanding*. Trans: Remnant P., Bennett J. Cambridge: Cambridge University Press.
- Lewis, D. (1988) 'What Experience Teaches', *Proceedings of the Russellian Society* (University of Sydney), 13: 29–57. Reprinted in Lewis, D. (1999) *Papers in Metaphysics and Epistemology*, 262–290. Edited by Ernest Sosa et al. Cambridge: Cambridge University Press.
- Lewis, D. (1983) *Philosophical Papers*. Volume I. New York; Oxford: Oxford University Press.
- Lewis, D. (1972) 'Psychophysical and Theoretical Identifications', *Australasian Journal of Philosophy*, 50, 249-258.

- Lewis, H. D. (1962) 'Mind and Body: Some Observations on Mr. Strawson's Views: The Presidential Address', *Proceedings of the Aristotelian Society*, 1962–1963, vol. 63: 1–22.
- Lewtas, P. (2018) 'Passive Causation; Making Interactionism Work', *Journal of Consciousness Studies*, 25 (9-10): 139–162.
- Lewtas, P. (2016) 'Building minds: solving the combination problem', *Inquiry: An Interdisciplinary Journal of Philosophy*, 60 (7): 742–781.
- Loar, B. (1999) 'David Chalmers's *The Conscious Mind*', *Philosophy and Phenomenological Research*, 59 (2): 465–472. Review of D. Chalmers (1996) *The Conscious Mind*, Oxford: Oxford University Press.
- Lockwood, M. (1993) 'The Grain Problem', in Howard M. Robinson (ed.) *Objections to Physicalism*: 271–291. Oxford: Oxford University Press.
- Lowe, E. J. (2012) 'Against Monism', in Philip Goff (ed.) *Spinoza on Monism*, 92–112. Basingstoke: Palgrave Macmillan.
- Lycan, W. (1996) *Consciousness and Experience*. Cambridge, MA: MIT Press.
- Malatesti, L. (2011) 'Thinking About Phenomenal Concepts', *Synthesis Philosophica*, 26 (2): 391–402.
- Mandik, P. (2001) 'Mental Representation and the Subjectivity of Consciousness', *Philosophical Psychology*, 14 (2): 179–202.
- Martin, C. B., Heil, J. (1999) 'The Ontological Turn', *Midwest Studies in Philosophy*, 23: 34–60.
- Martin, C. B. (1996) 'Properties and dispositions', in T. Crane (ed.) *Dispositions: A Debate*, 71–86. London: Routledge.
- Mathews, F. (2011) 'Panpsychism as Paradigm', in M. Blamauer (ed.) *The Mental as Fundamental*. Frankfurt Heusenstamm: Ontos Verlag.
- Maung, H. H. (2019) 'Panpsychism, Conceivability, and Dualism Redux', *Synthesis Philosophica*, 67: 157–162.
- Maxwell, G. (1978) 'Rigid Designators and Mind-Brain Identity', in C. Savage (ed.) *Perception and Cognition: Minnesota Studies in the Philosophy of Science, Volume IX*, 365–404. Minneapolis: Minnesota University Press.
- McMahan, J. (2002) *The Ethics of Killing: Problems at the Margins of Life*. New York: Oxford University Press.
- Meinong, A. (1904) *Untersuchungen zur Gegenstandstheorie und Psychologie [Investigations in Theory of Objects and Psychology]*. Leipzig: J. A. Barth.
- Merleau-Ponty, M. (1962) *Phenomenology of Perception*. Translated by Colin Smith. London: Routledge.
- Merricks, T. (2001) *Objects and Persons*. Oxford: Oxford University Press.
- Molnar, G. (2003) *Powers: A Study in Metaphysics*. Oxford: Oxford University Press.
- Montero, B., Papineau, D. (2005) 'A Defence of the Via Negativa Argument for Physicalism', *Analysis*, 65 (3): 233–237.

- Mørch, H. H. (2020) 'The Argument for Panpsychism From Experience of Causation', in W. Seager (ed.) *The Routledge Handbook of Panpsychism*, 269–284. New York: Routledge.
- Mørch, H. H. (2018) 'Does Dispositionalism Entail Panpsychism?', *Topoi*, 39: 1073–1088.
- Mørch, H. H. (2014) *Panpsychism and Causation: A New Argument and a Solution to the Combination Problem*. PhD thesis, University of Oslo.
- Morris, K. (2017) 'The Combination Problem: Subjects and Unity', *Erkenntnis* 82: 103–120.
- Nagasawa, Y., Wager, K. (2016) 'Panpsychism and Priority Cosmopsychism', in G. Brüntrup & L. Jaskolla (eds.) *Panpsychism: Contemporary Perspectives*, 113–129. New York: Oxford University Press.
- Nagel, T. (1986) *The View from Nowhere*. New York: Oxford University Press.
- Nagel, T. (1979) *Mortal Questions*. Cambridge: Cambridge University Press.
- Nagel, T. (1974) 'What Is It Like to Be a Bat?', *The Philosophical Review*, 83 (4): 435–450.
- Nagel, T. (1971) 'Brain Bisection and the Unity of Consciousness', *Synthese* 22 (3): 396–413.
- Ney, A. (2008) 'Physicalism as an Attitude', *Philosophical Studies*, 138: 1–15.
- Nida-Rümelin, M., O Conaill (2019) 'Qualia: The Knowledge Argument', *Stanford Encyclopedia of Philosophy*. <URL= <https://plato.stanford.edu/entries/qualia-knowledge/>>.
- Nida-Rümelin, M. (1998) 'On Belief about Experiences. An Epistemological Distinction Applied to the Knowledge Argument against Physicalism', *Philosophy and Phenomenological Research*, 58 (1): 51–73.
- Norden, B. V., Jones, N. (2019) 'Huayan Buddhism', *Stanford Encyclopedia of Philosophy*. <URL= <https://plato.stanford.edu/entries/buddhism-huayan/>>.
- Olson, E. T. (2007) *What Are We? A Study in Personal Ontology*. New York: Oxford University Press.
- Papineau, D. (2007a) 'Phenomenal and Perceptual Concepts', in T. Alter, S. Walter (eds.) *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, 111–144. New York: Oxford University Press.
- Papineau, D. (2007b) 'Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness', *Notre Dame Philosophical Reviews* (4). Review of D. Stoljar (2006) *Ignorance and Imagination*, New York: Oxford University Press. <URL=<https://ndpr.nd.edu/reviews/ignorance-and-imagination-the-epistemic-origin-of-the-problem-of-consciousness/>>.
- Papineau, D. (2002) *Thinking About Consciousness*. Oxford: Oxford University Press.
- Papineau, D. (1993) *Philosophical Naturalism*. Oxford: Blackwell.
- Parfit, D. (2012) 'We Are Not Human Beings', *Philosophy*, 87 (1): 5–28.
- Parfit, D. (1971) 'Personal Identity', *Philosophical Review*, 80 (1): 3–27.
- Petrić, F. (1979) *Nova sveopća filozofija (Nova de Universis Philosophia)*. Zagreb: Sveučilišna naklada Liber.

- Pfeifer, K. (2016) 'Pantheism as Panpsychism', in A. Buckareff, Y. Nagasawa (eds.) *Alternative Concepts of God: Essays on the Metaphysics of the Divine*, 41–49. Oxford: Oxford University Press.
- Purnell, F. (2004) 'Francesco Patrizi', *Stanford Encyclopedia of Philosophy*. <URL=<https://plato.stanford.edu/entries/patrizi/>>.
- Robinson, W. (2018) 'Russellian Monism and Epiphenomenalism', *Pacific Philosophical Quarterly*, 99 (1): 100–117.
- Roelofs, L. (2014) 'Phenomenal Blending and the Palette Problem', *Thought: A Journal of Philosophy*, 3 (1): 57–70.
- Royce, J. (1881) 'Mind Stuff and Reality', *Mind* 6 (23): 365–377.
- Russell, B. (1927/1992) *The Analysis of Matter*. London: Routledge.
- Schaffer, J. (2010) 'Monism: The Priority of the Whole', *The Philosophical Review* 119 (1): 31–76.
- Schacter, E. (2012) 'The switch model of split-brain consciousness', *Philosophical Psychology* 25 (2): 203–226.
- Schlick, M. (1959) 'Positivism and Realism', in A. J. Ayer (ed.) *Logical Positivism*, 82–107. New York: The Free Press.
- Schroeter, L. (2021) 'The 2D Argument Against Materialism', *Stanford Encyclopedia of Philosophy*. <URL=<https://plato.stanford.edu/entries/two-dimensional-semantics/supplement.html>>.
- Schwitzgebel, E., Gordon, M. S. (2000) 'How Well Do We Know Our Own Conscious Experience? The Case of Human Echolocation', *Philosophical Topics* 28 (2): 235–246.
- Seager, W. (2016) 'Panpsychism Infusion', in Brüntrup & Jaskolla (eds.) *Panpsychism: Contemporary Perspectives*: 229–248. New York: Oxford University Press.
- Seager, W. (2010) 'Panpsychism, Aggregation and Combinatorial Infusion', *Mind and Matter* 8 (2). <URL=<https://philpapers.org/archive/SEAPAA.pdf>>.
- Seager, W. (2006) 'The 'Intrinsic Nature' Argument for Panpsychism', *Journal of Consciousness Studies*, 13 (10–11): 129–145.
- Seager, W. (1995) 'Consciousness, Information and Panpsychism', *Journal of Consciousness Studies* 2 (3): 272–288.
- Sergent, J. (1990) 'Furtive Incursions into Bicameral Minds', *Brain* 113 (2): 537–68.
- Shani, I. (2015) 'Cosmopsychism: A Holistic Approach to the Metaphysics of Experience', *Philosophical Papers*, 44 (3): 389–437.
- Shani, I. (2010) 'Mind Stuffed with Red Herrings: Why William James' Critique of the Mind-Stuff Theory Does not Substantiate a Combination Problem for Panpsychism', *Acta Analytica* 25: 413–434.
- Shoemaker, S. (2006) 'The Frege-Schlick View', in J. Thomson and A. Byrne (eds.) *Content and Modality: Themes from the Philosophy of Robert Stalnaker*, 18–33. New York: Oxford University Press.

- Shoemaker, S. (1984) 'Personal Identity: A Materialist Account', in S. Shoemaker and R. Swinburne (eds.) *Personal Identity*: 67–132. Oxford: Basil Blackwell.
- Shoemaker, S. (1982) 'The Inverted Spectrum', *The Journal of Philosophy*, 79 (7): 357–381.
- Simon, J. A. (2017) 'Vagueness and zombies: why 'phenomenally conscious' has no borderline cases', *Philosophical Studies*, 174 (8): 2105–2123.
- Snowdon, P. F. (2014) *Persons, Animals, Ourselves*. Oxford: Oxford University Press.
- Spurrett, D., Papineau, D. (1999) 'A Note on the Completeness of 'Physics'', *Analysis*, 59 (1): 25–29.
- Stoljar, D. (2021) 'Physicalism', *Stanford Encyclopedia of Philosophy*. <URL=<https://plato.stanford.edu/entries/physicalism/>>.
- Stoljar, D. (2006) *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness*. New York: Oxford University Press.
- Stoljar, D. (2001a) 'The Conceivability Argument and Two Conceptions of the Physical', *Philosophical Perspectives*, 15: 393–413.
- Stoljar, D. (2001b) 'Two Conceptions of the Physical', *Philosophy and Phenomenological Research*, 62 (2): 253–281.
- Strawson, G. (2008) 'The Identity of the Categorical and the Dispositional', *Analysis* 68: 271–282.
- Strawson, G. (2006) 'Realistic Monism, Why Physicalism Entails Panpsychism', *Journal of Consciousness Studies*, 13 (10–11): 3–31.
- Strawson, P. F. (1959/2003) *Individuals*. London and New York: Routledge.
- Taylor, H. (2018) 'Powerful Qualities and Pure Powers', *Philosophical Studies*, 175: 1423–1440.
- Strickland, L. (2014) *Leibniz's Monadology: A New Translation and Guide*. Edinburgh: Edinburgh University Press.
- Thompson, E. (2015) 'Dreamless Sleep, the Embodied Mind, and Consciousness - The Relevance of a Classical Indian Debate to Cognitive Science', in T. Metzinger and J. M. Windt (eds). *Open MIND*. Frankfurt am Main: MIND Group.
- Toga, A. W., Thompson, P. M. (2003) 'Mapping Brain Asymmetry', *Nature Reviews Neuroscience*, 4: 37-48.
- Turausky, K. (manuscript) 'Picturing Panpsychism: New Approaches to the Combination Problem'. <URL= [https://turausky.files.wordpress.com/2017/07/combo-combination-problem\\_ket-2.pdf](https://turausky.files.wordpress.com/2017/07/combo-combination-problem_ket-2.pdf) >.
- Tye, M. (2021a) 'Qualia', *Stanford Encyclopedia of Philosophy*. <URL=<https://plato.stanford.edu/entries/qualia/>>.
- Tye, M. (2021b) *Vagueness and the Evolution of Consciousness*. Oxford: Oxford University Press.
- Tye, M. (2003) *Consciousness and Persons: Unity and Identity*. Cambridge, MA: MIT Press.

Tye, M. (2000) 'Knowing What it is Like: The Ability Hypothesis and the Knowledge Argument', in Tye, M. (ed.), *Consciousness, Color, and Content*, 3–20. Cambridge, MA: MIT Press.

Tye, M. (1996) 'Is Consciousness Vague or Arbitrary?', *Philosophy and Phenomenological Research*, 56 (3): 679–685.

Yablo, S. (1993) 'Is conceivability a guide to possibility?', *Philosophy and Phenomenological Research*, 53: 1–42.

Zahavi, D. (2014) *Self and Other: Exploring Subjectivity, Empathy, and Shame*. Oxford: Oxford University Press.

Zahavi, D. (2010) 'Minimal self and narrative self. A distinction in need of refinement', in T. Fuchs, H. Sattel, and P. Heningsen (eds.), *The Embodied Self: Dimensions, Coherence, and Disorders*, 3–11. Stuttgart: Schattauer Verlag.