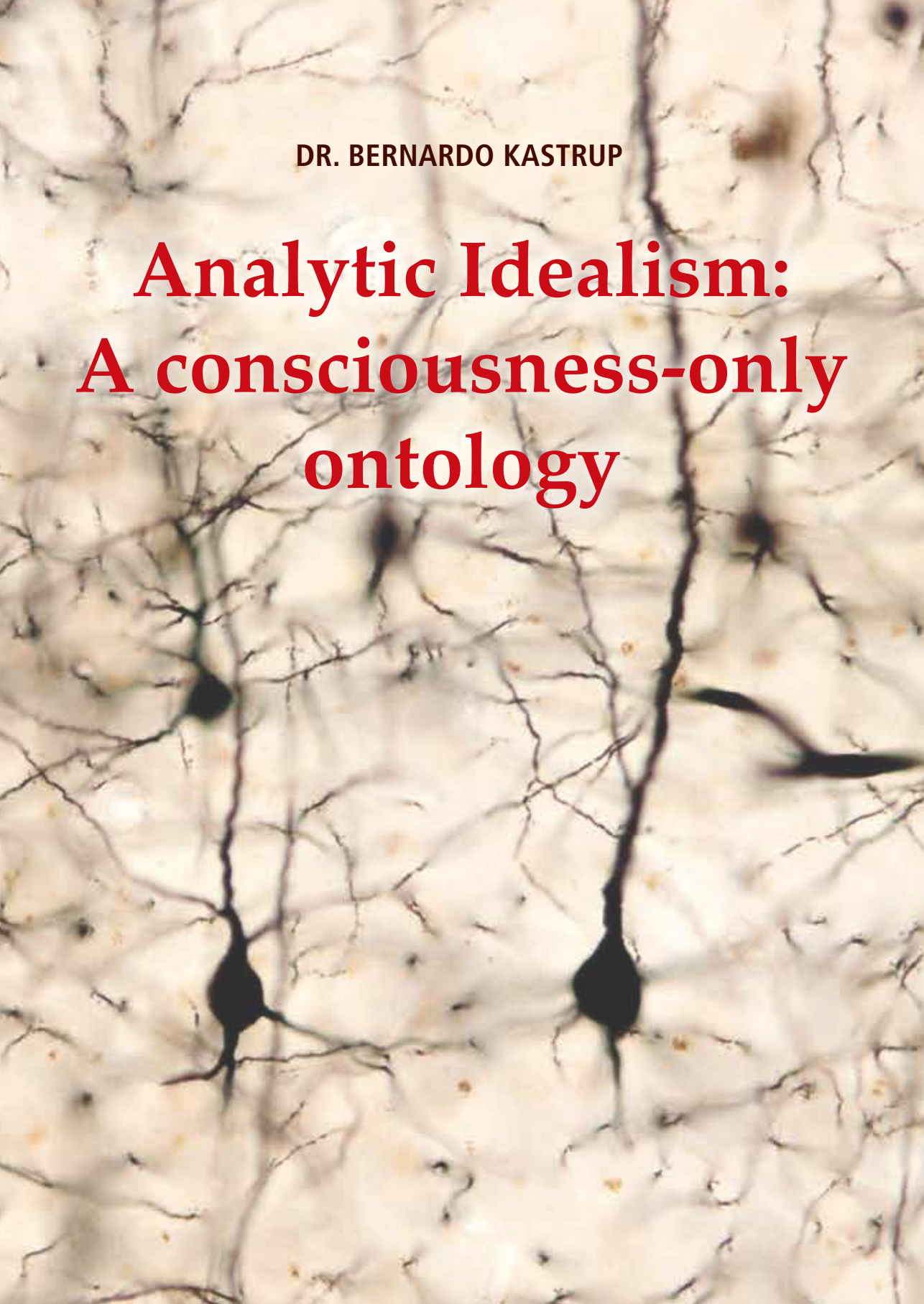

DR. BERNARDO KASTRUP

Analytic Idealism: A consciousness-only ontology

Analytic Idealism:
A consciousness-only ontology

dr. Bernardo Kastrup

Kastrup, Bernardo
Analytic Idealism: A consciousness-only ontology
ISBN/EAN: 978-94-028-1400-2

Analytische Idealisme:
Een ontologie met alleen bewustzijn

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken
volgens besluit van het college van decanen
te verdedigen op maandag 29 april 2019
om 14.30 uur precies

door

dr. Bernardo Kastrup
geboren op 21 oktober 1974
te Niterói, Rio de Janeiro, Brazilië.

Promotiecommissie

Promotor:
prof. dr. M. V. P. Slors

Copromotor:
dr. ing. L. C. de Bruin

Manuscriptcommissie:

prof. dr. P. J. J. M. Bakker

Prof. B. J. Carr, MA, PhD
Queen Mary University of London, Verenigd Koninkrijk

Dr. S. Coleman
University of Hertfordshire, Verenigd Koninkrijk

dr. V. A. Gijsbers
Universiteit Leiden, Nederland

dr. D. W. Strijbos

Table of Contents

Acknowledgments

I would like to thank, first of all, my supervisors, prof. dr. Marc Slors and dr. Léon de Bruin, for their guidance and the productive discussions we have had in the run-up to the completion of this dissertation.

The critical feedback and suggestions of anonymous reviewers have significantly improved the papers that constitute the core of this dissertation.

The gracious support of Prof. David Chalmers has also been instrumental. Not only has David critically reviewed key parts of my material, he has also given me the opportunity to participate—with funding from the Global Institute for Advanced Studies of New York University, which I gratefully acknowledge—in a workshop focused on idealism, in Shanghai, late in the spring of 2017.

Discussions with, and critical feedback from, Prof. Galen Strawson have helped sharpen key points made in this dissertation, for which I am very grateful to Galen.

Discussions with many other colleagues have also been valuable. With the risk of leaving important names out, I would like to explicitly thank philosophers Itay Shani, Daniel Stoljar, Miri Albahari, Michael Pelczar, Barry Dainton and Philip Goff.

The feedback and encouragement I received from researchers of the Department of Psychiatry and Neurobehavioral Sciences, University of Virginia School of Medicine, during my visit there in the spring of 2016—for which I also gratefully acknowledge funding—have been key to the effort that eventually resulted in this dissertation. I am particularly grateful to Prof. Edward F. Kelly for his continuing encouragement and guidance.

Neuroscientist Anil Seth has been of much help by pointing out to me the relatively recent literature on so-called 'no-report paradigms' in consciousness research.

For the past five years, I have also been very fortunate to be able to count on the informal guidance of Menas C. Kafatos, the Fletcher Jones Endowed Chair Professor of Computational Physics, Chapman University.

Without the continuing trust of Michael D. Lemonick, Chief Opinion Editor at *Scientific American*, the ideas discussed in this dissertation would not have had the level of mainstream media exposure they now have.

Last but not least, without the loving and reassuring presence of my partner, Claudia Damian, this dissertation would never have come true.

1. Introduction

1.1 In a nutshell

This dissertation elaborates on a modern, analytic version of the ontology of idealism, according to which (a) phenomenal consciousness, as an ontological category, is fundamental; and (b) everything else in nature can ultimately be reduced to, or grounded in, patterns of excitation of phenomenal consciousness. It posits a reduction base consisting of a single element: spatially unbound, universal phenomenal consciousness. Its key challenge is then to explain how the seemingly distinct phenomenal inner lives of different subjects of experience can arise within this fundamentally unitary phenomenal field. This is sometimes called the "decomposition problem" in the literature (Chalmers 2016a) and it is the core problem this dissertation attempts to tackle. Along the way, a variety of other challenges are addressed, such as: how we can reconcile idealism with the fact that we all inhabit a common external world; why this world unfolds independently of our personal volition or imagination; why there are such tight correlations between measured patterns of brain activity and reports of experience; etc.

Idealism has had its heyday in Western philosophy in the 18th (Berkeley) and early 19th (Hegel) centuries. Though it has enjoyed popularity amongst continental philosophers, analytic philosophers have, by and large, failed to take idealism seriously, perhaps because of its association with religious traditions in both East and West. With this dissertation, I hope to help change this by offering a strictly analytic, conceptually clear articulation of idealism. I also hope to offer empirical neuroscientific evidence suggesting that idealism may be better suited to make sense of the data than mainstream physicalism or constitutive panpsychism.

The core of this dissertation consists of five papers—each a piece of the larger jigsaw puzzle assembled in this volume—published in academic journals. They are reproduced here without any change of substance. The order in which they are presented is meant to help more effectively convey the overarching argument of which each individual paper is a part. In this Introduction, I shall summarize this overarching argument so to help readers place each paper in its broader context, as they make their way through the dissertation. In other words, here I shall attempt to sketch the final, completed picture of the jigsaw puzzle before each piece is explored in more detail.

For this reason, however, readers should not expect the highly summarized argument presented in this Introduction to be strictly rigorous or complete. The goal is to first convey the general idea behind this dissertation, before elaborating on it with the rigor of the subsequent chapters. The appropriate literature reviews, as well as discussions on how the work presented here is situated in the context of prior efforts, are also comprised in the papers ahead (chapter 2 to 6), not in this Introduction.

1.2 The big picture

It has now become trite to point out that mainstream physicalism fails—unless one subscribes to its eliminative formulation, a view whose absurdity I shall not bother to argue for here[1]—to account for the sole given fact of reality: the existence of experience (e.g. Chalmers 2003). Physicalism is also arguably irreconcilable—insofar as it presupposes physical realism—with results emerging from physics laboratories around the world (e.g. Kim *et al*. 2000, Gröblacher *et al*. 2007, Romero *et al*. 2010, Lapkiewicz *et al*. 2011, Ma *et al*. 2013, Manning *et al*. 2015, Hensen *et al*. 2015, etc.), as elaborated upon in Section A.3 of Appendix A.[2] So both in terms of its explanatory power and its consistency with empirical observations, our mainstream ontology is found wanting.

I mention this merely to highlight the need for an alternative ontology, such as that offered in this dissertation. Other than a brief review of the 'hard problem of consciousness' in Chapter 3, I shall not focus on discussing the untenability of mainstream physicalism. This has already been done in the literature (e.g. Levine 1983, Chalmers 1996, Rosenberg 2004: 13-30, Strawson *et al*. 2006: 2-30, etc.). What I *shall* attempt in the next chapter is something more ambitious and—hopefully—more constructive: to point out the failures and internal contradictions *of the very thought processes that underlie mainstream physicalism* and related ontologies. Only by understanding these implicit, unexamined failures and contradictions can we hope to reform our thinking and eventually solve (or circumvent) the associated dilemmas and paradoxes.

In this context, Chapter 2 discusses what is perhaps the root of key unresolved problems in contemporary analytic philosophy: the tendency to try to make sense of nature by replacing concrete observations with theoretical abstractions. Such attempts often consist of mere word games, played in thought with a rich and shifting phantasmagoria of concepts. The process tends to unfold so implicitly that many don't seem to even notice how many steps of—epistemically unreliable—conceptual abstraction their reasoning entails. Chapter 2 attempts to make these word games explicit. It also suggests more epistemically reliable lines of reasoning that avoid unnecessary conceptual abstractions.

By pursuing these more reliable lines of reasoning, Chapter 3—the core of this dissertation—elaborates on an analytic formulation of idealism. It argues that the best categorical explanation for the facts of nature entails that these facts

[1] Interested readers can peruse, instead, an excellent recent essay by Galen Strawson, who has put it best: https://www.nybooks.com/daily/2018/03/13/the-consciousness-deniers/.

[2] For a less technical approach, see two of my essays on *Scientific American*'s *Observations* blog: https://blogs.scientificamerican.com/observations/should-quantum-anomalies-make-us-rethink-reality/ and https://blogs.scientificamerican.com/observations/coming-to-grips-with-the-implications-of-quantum-mechanics/.

are essentially *phenomenal*. All of existence consists, it is claimed, solely of *ideas*—thoughts, emotions, perceptions, intuitions, imagination, etc.—even though not one's *personal* ideas alone.

The ontology articulated in Chapter 3 can be summarized thus: there is only universal phenomenal consciousness. We, as well as all other living organisms, are *dissociated alters* of this universal consciousness, analogously to how a person with Dissociative Identity Disorder (DID) manifests multiple disjoint centers of subjectivity also called 'alters.' We, and all other living organisms, are surrounded by the *trans*personal phenomenal activity of universal consciousness, which unfolds beyond the dissociative boundary of our respective alter. The inanimate world[3] we perceive around us is the 'extrinsic appearance'—i.e. the phenomenal image imprinted from *across* our dissociative boundary—of this activity. The living organisms we share the world with are the extrinsic appearances of other alters.

Instead of the mainstream physicalist postulate of an ontological category fundamentally outside and independent of mind,[4] Chapter 3 offers a different categorical interpretation of what we call 'matter.' Indeed, according to it the living brain is merely a *phenomenal appearance* of a person's conscious inner life—her thoughts, feelings, fantasies, beliefs, etc.—as presented on the screen of perception of e.g. another person. And since the brain is made of matter, this is what the matter in a living brain *is*. Chapter 3 then goes further and, in the spirit of parsimony that underlies this entire dissertation, argues that this is also what *all* matter is: the phenomenal appearance of equally phenomenal activity unfolding across a dissociative boundary. The matter constituting the inanimate universe is, thus, what *trans*personal experiences unfolding outside the alters *look like* from the point of view of an alter, just as a living brain is what *personal* experiences look like. By construing all matter to be a phenomenal appearance of equally phenomenal activity, analytic idealism requires nothing more than phenomenality to offer a coherent categorical interpretation of nature.

Many criticisms can be—and have been—made against such a consciousness-only ontology. Indeed, because of the formidable cultural momentum behind the notion of an objective physical world distinct from mind, one can promptly leverage a ready-made, culturally sanctioned list of objections against idealism. Chapter 4 lists many of these objections and tackles them one by one. It attempts to show that they are often based on logical fallacies such as question-begging, unexamined assumptions, misunderstandings of the implications of analytic idealism, etc.

---

[3] Throughout this dissertation, I use the word 'inanimate' in the sense of non-living; i.e. as that which is not biology.

[4] Throughout this dissertation—*except in Chapter 5*, where they are defined in a different way—I use the words 'mind' and 'mentation' as synonyms of phenomenal consciousness and phenomenal activity, respectively.

One objection is exceptional because it poses some legitimate difficulties: a necessary implication of the ontology proposed in Chapter 3 is that an organism's metabolism—*all* of it—is the extrinsic appearance of the organism's conscious inner life. This is reasonable enough for certain patterns of brain activity known to correlate with experiences accessible through introspection, but what about metabolism beyond the brain, such as e.g. liver and kidney function? And what about the metabolic activity taking place in, say, a person's left big toe? If the ontology proposed in Chapter 3 is correct, then liver, kidney and even toe function must all correspond to experiences as well. Yet, try as we might, these experiences do not seem to be accessible through introspection.

Moreover, even if we were to look at the brain alone, ignoring the metabolism in the rest of the body, recent studies in psychology suggest the presence of seemingly *un*conscious mental processes in the brain (e.g. Hassin 2013). This, if true, would already contradict analytic idealism.

Chapter 5 bites these bullets and argues that, despite appearances to the contrary, there is no clear reason to believe that *any* mental process is truly unconscious. Instead, it attempts to show that there are, in fact, very good reasons to think that what we regard as unconscious mental processes correspond merely to an *illusion* of unconsciousness, which results from dissociative states or lack of metacognition. And once these two mechanisms—dissociative states and lack of metacognition—are identified, they can explain why experiences corresponding to areas of the living body beyond the nervous system can't be accessed through introspection.

The last piece of the puzzle is that of empirical evidence. Chapter 6 compiles and discusses a broad list of instances of brain function *impairment* that are accompanied by *enrichment* of conscious inner life and an *expansion* of one's sense of identity. The list includes cases as varied as asphyxiation, physical trauma to the head, the consumption of psychoactive substances that dampen brain activity, etc.

Such correlations between impaired brain function and enriched conscious inner life are at least counterintuitive under the mainstream physicalist notion that conscious inner life is constituted or generated by brain activity. Under analytic idealism, on the other hand, they are to be expected: if normal brain function is part of the extrinsic appearance of a *dissociated* alter of universal consciousness, then *some* forms of reduction or impairment of normal brain function should be the extrinsic appearance of a reduction or impairment of the *dissociation*. And, of course, from a first-person perspective a reduction of dissociation must be experienced as an enrichment of conscious inner life: reintegrated memories, the recovery of a broader sense identity, renewed access to previously dissociated insights and emotions, reintegration of previously dissociated skills, etc. Contrary to physicalism, analytic idealism can thus not only accommodate, but also make sense of, the evidence discussed in Chapter 6.

Naturally, the argument in Chapter 6 is not that *all* impairment of brain function should be accompanied by enriched inner life. Otherwise, the smartest

and most creative people would be those with the most damaged brains. This is clearly not the case. *But neither does analytic idealism require it to be the case.* Allow me to elaborate.

As discussed in Chapter 3, a living organism corresponds to a dissociated alter of universal consciousness. As such, each person can be regarded as a segment of universal consciousness—meant here generically, without implying that universal consciousness necessarily has spatiotemporal extension—comprising its own dissociated phenomenal states. Segments comprising many phenomenal states can be referred to as 'big alters,' whereas segments comprising few phenomenal states can be referred to as 'small alters.' It is reasonable to say, for instance, that human beings correspond to bigger alters than, say, insects.

Notice that, in principle, *both big and small alters can be equally well dissociated*. In other words, the relative amount of phenomenal states encompassed by an alter does not bear relevance to how well dissociated these phenomenal states are from the rest of universal consciousness. It is entirely coherent, within the logic of analytic idealism, that a small alter could be more strongly dissociated from universal consciousness than a big alter, or the other way around.

Now, since brain activity is part of the extrinsic appearance of an alter's dissociated phenomenal states, it stands to reason that some—even *most*—types of brain function impairment should correspond simply to a reduction of the phenomenal states of the alter. These types of brain function impairment will not disrupt the dissociation itself, but only stifle whatever is circumscribed by the dissociative boundary. The alter will become smaller, cognitively compromised, but still equally well dissociated. This is why, under analytic idealism, many or even most types of brain function impairment should still come accompanied by cognitive deficit, not awareness expansion.

Only some specific types of brain function impairment, which somehow affect *the dissociative mechanisms themselves*—as opposed to the phenomenal states encompassed by the alter—should correlate with an enrichment of conscious inner life. They make the dissociative boundary 'porous,' so to speak. At present, however, it is not yet known what precise aspects of brain function correspond to these dissociative mechanisms, even though some tantalizing indications are discussed in Chapter 6. For this reason, it is currently impossible to predict with accuracy what types of brain function impairment should lead to what type of effect: awareness expansion or cognitive deficit.

What distinguishes the predictions of analytic idealism from those of mainstream physicalism is this: under analytic idealism, *some* types of brain function impairment should, in principle, lead to enriched conscious inner life. Under mainstream physicalism, however, this is much more difficult to argue, as elaborated upon in Chapter 6. More rigorously put, my claim is this: *there are some types of brain function impairment—which under mainstream physicalism should correlate with cognitive deficit and under analytic idealism*

with enriched inner life—that have been shown to be accompanied by enriched inner life.

Notice that, because of the inherent limitations of gauging consciousness from a second- or third-person perspective, the playing field isn't level: many types of brain function impairment may cause both an enrichment of conscious inner life *and* compromise the subjects' ability to *report* this enrichment. For instance, language or motor centers, memory pathways or a variety of other communication-critical functions in the brain may be compromised, harming or eliminating the subjects' ability to speak or write. For all we know, many subjects could be lying in hospital with severe head trauma or other brain ailments, having unfathomable inner experiences, and yet be utterly incapable of relating any of it to family or medical staff. If brain areas essential to metacognition are compromised, subjects may not even be able to report their experiences *to themselves*, as discussed in Chapter 5. Consequently, the potential for evidence that corroborates analytic idealism is restricted by conflicting requirements: the corresponding brain function impairment must be sufficient to affect dissociative mechanisms—not just dampen the phenomenal states encompassed by the alter—whilst preserving enough cognitive function so subjects can report their expanded awareness. These conflicting requirements aren't trivial to meet concurrently. It is, thus, if anything, surprising that so many case reports exist in the literature that seem to corroborate analytic idealism, as discussed in Chapter 6.

Chapter 7 then discusses some important issues related to, but left insufficiently addressed by, chapters 2 to 6. It also points to potential areas of future investigation.

1.3 The appendices

The analytic case for idealism is laid out in chapters 2 to 6. Nonetheless, there are two topics that, despite not being part of the core argument of this dissertation, arise so forcefully from it that it would have been negligent to leave them unaddressed. I have thus added two papers—also originally published in academic journals—that tackle these topics in two appendices, A and B, respectively.

Appendix A addresses the following question: If analytic idealism is true, what are its implications? In other words, how does it change the way we look upon life and the world? Indeed, whereas mainstream physicalism denies the meaning of the world by construing it to be a mechanical contraption governed by blind laws and mere chance, analytic idealism regards the world as the extrinsic appearance of intrinsic, universal phenomenal activity. According to it, nature holds hidden but inherent semantic meaning: it points symbolically to something beyond its face-value appearance. This is unpacked in Appendix A, in an attempt to highlight the relevance of idealism to life.

Because it was originally published as a self-contained paper, Appendix A—more specifically, Section A.3—includes an empirical argument to substantiate

its starting hypothesis that the world is essentially phenomenal. It elaborates upon the experimental evidence for what is technically called 'contextuality' in physics: the notion that physical quantities are fundamentally dependent on observation and have no definite existence before being observed. The link between contextuality and idealism is also made explicit in Appendix A.

Attentive readers will notice that, throughout Appendix A, I use the word 'meaning' to denote 'sense' (as in the sense of a word or phrase), 'significance' (as in the significance of a historical moment) *and* 'purpose' (as in the purpose of an action), freely conflating all three usages. This conflation is intentional and implicitly reflects the very conclusion of the appendix: that the purpose of life is to unveil the sense and significance of the world. Thus the meaning of life in the world is simultaneously life's purpose *and* the world's sense and significance. Indeed, the very linguistic versatility of the word 'meaning' amplifies the argument in Appendix A: 'purpose' is intrinsically connected with 'sense' and 'significance.' Perhaps language captures and preserves—like a time capsule—ancient intuitions we have since allowed to escape us.

The second question that naturally arises if one finds the argument for analytic idealism compelling is this: What drove the formidable momentum behind the mainstream adoption of physicalism over the past 200 years or so, if a more plausible and viable alternative—unaffected by fundamental problems, such as the 'hard problem of consciousness'—has existed all along?

Appendix B argues that a key motivation for the development and mainstream adoption of physicalism has been *psychological*, as opposed to philosophical. This may come as a surprising assertion, for physicalism is often regarded as a purely fact-based interpretation of reality, untarnished by subjective biases or covert wish-fulfillment maneuvers. Appendix B argues that this may not be true, for there are compelling reasons to believe that the physicalist worldview protects and validates the ego, even in view of formidable threats such as death. Perhaps even more surprisingly, behind physicalism's apparent denial of meaning there operate—it is argued—psychological mechanisms that seek to *enhance* one's sense of meaning in life.

For these reasons, I argue that it is not surprising that physicalism, despite its inherent problems, has come to amass the formidable level of support it has today among the intellectual elites, particularly in academia. As is often the case with views that come to define a culture, there is more to physicalism's success than its philosophical merits.

1.4 Preempting misunderstandings

While discussing the ideas presented in this dissertation with other philosophers, it has become clear to me that a few observations should be made upfront, in order to facilitate the correct understanding—and, perhaps more importantly, preempt *mis*understandings—of what is claimed in the papers ahead.

A claim made already in the title of Chapter 3 is that, according to analytic idealism, the universe is *in* consciousness. This is liable to misinterpretation: insofar as analytic idealism entails that consciousness is the categorical basis—the underlying essence—of all that exists, shouldn't one say, instead, that the universe *is* consciousness?

This would indeed be so if I took the word 'universe' to denote 'all there is.' After all, if the universe is all there is and it is 'made of' consciousness, then to say that the universe is *in* consciousness would amount to saying that consciousness is in consciousness.

However, throughout this dissertation, I take the word 'universe' to denote *what we perceive and measure*. In other words, I am using the operational definition of 'universe' in physics, as opposed to a metaphysical one. The claim is then that this perceived universe is *in* consciousness since, according to analytic idealism, it consists of particular patterns of excitation *of* universal consciousness. The universe is thus in consciousness for the same reason that ripples are in water.

A second point prone to misunderstanding is the following: as discussed in Chapter 3, an important contribution of this dissertation is the notion that *dissociation*—at a *universal* scale—is what creates the appearance of fragmentation of universal consciousness into multiple disjoint centers of experience, such as you and me. Yet, at a *human* scale, dissociation is often thought of as presupposing intentionality, or 'aboutness.' Allow me to elaborate.

Many of our human phenomenal states entail intentionality: we think *about* buying the car we *saw* at the dealership; we feel bad *about* the news we *heard* on the radio; etc. These thoughts and feelings are thus about things or events in the 'world out there': they are anchored in some content of *sense perception* accessible through episodic memory (Chalmers 1996: 19). When dissociation happens at the human level, it is often episodic memory access that becomes compromised as a reaction to trauma: traumatic memories are no longer accessible through the chains of cognitive association that characterize regular psychic life (American Psychiatric Association 2013).

The possible misunderstanding is then this: *if* dissociation presupposed intentionality, then the argument in Chapter 3 would fail because it posits that dissociation—at a universal scale—is what enables intentionality to begin with, by creating a boundary between an alter and its surrounding environment. Without this boundary there would be—it is argued—no sense perception, no 'world out there' *about* which we could think or feel.

The misunderstanding lies in assuming that dissociation is *defined in terms of* intentional content, simply because it often happens, in humans, in connection with intentional content. However, dissociation entails merely the cessation of an otherwise normal cognitive association between two phenomenal states—say, a thought and a feeling—regardless of whether these states have intentional content or not.

A thought experiment should make this clear: it is possible to conceive of an infant kept from birth in an ideal sensory-depravation chamber. Such an infant would not only have abstract thoughts and feelings, but there would also be natural cognitive associations across these thoughts and feelings. If some of these cognitive associations were to cease, one would still be able to speak of dissociation, even though the infant would have never experienced sense perception. In other words, it seems perfectly possible that phenomenal states could become dissociated from each other, even if they don't have phenomenal content. Therefore, the argument in Chapter 3 holds.

The final point of possible misunderstanding has to do with the argument laid out in Chapter 2. There, I attempt to show that the notion of an ontological category outside and independent of phenomenal consciousness is not only a theoretical abstraction—as opposed to an empirical observation—it is also an *epistemically unreliable* abstraction. Readers of that paper have, however, tended to assume that I was seeking to make an affirmative *metaphysical* point based on this *epistemic* basis (see the Open Peer Commentaries in Kastrup 2018b). This is incorrect.

So let me be clear upfront: what Chapter 2 attempts is to highlight that different ontologies inherently carry different epistemic costs—i.e. degrees of epistemic confidence—even if these ontologies are both internally consistent and consistent with empirical observations. And whereas this is admittedly *not* a metaphysical argument, it undoubtedly has great relevance in informing one's choice of metaphysics, since all that is available for making such a choice is one's knowledge. The degree to which one's knowledge is reliable should be a factor—perhaps even a *defining* factor—in the choice.

Having made these upfront clarifications, I am now ready to begin elaborating on analytic idealism by, first, laying out its epistemic basis and motivation in the next chapter.

2. Conflating Abstraction with Empirical Observation: The False Mind-Matter Dichotomy

This paper first appeared in *Constructivist Foundations*, ISSN 1782-348X, Vol. 13, No. 3, in July 2018.

2.1 Abstract

The alleged dichotomy between mind and matter is pervasive. Therefore, the attempt to explain matter in terms of mind (idealism) is often considered a mirror image of that of explaining mind in terms of matter (mainstream physicalism), in the sense of being structurally equivalent despite being reversely arranged. I argue that this is an error arising from language artifacts, for dichotomies must reside in the same level of abstraction. Because matter outside mind is not an empirical observation but rather an explanatory model, the epistemic symmetry between the two is broken. Consequently, matter and mind cannot reside in the same level of abstraction. It then becomes clear that attempting to explain mind in terms of matter is epistemically more costly than attempting to explain matter in terms of mind. The paper highlights the primacy of perceptual constructs over explanatory abstraction on both epistemic and ontic levels.

2.2 Introduction

The (unexamined) assumption that mind and matter are jointly exhaustive and mutually exclusive concepts is pervasive today. In other words, many scholars implicitly take every aspect of existence to be either mental (e.g. thoughts, emotions, hallucinations) or physical (e.g. tables and chairs), mentality and physicality being polar opposites in some sense. Originating with René Descartes and Immanuel Kant (Walls 2003: 130), this dichotomy has been firmly entrenched in Western thought since at least the early nineteenth century. Eminent scholarly publications of the time, such as *The British Cyclopædia of Natural History,* lay it out unambiguously: "as mind is the opposite of matter in definition, the perfection of its exercise must be the opposite of that of the exercise of matter" (Partington 1837: 161). From the early twentieth century onwards, more nuanced formulations of the dichotomy were proposed. Alfred North Whitehead (1947), for instance, considered mind and matter *co-dependent* opposites. Even Henri Bergson, whose conception of an *élan vital* was meant to dilute the Cartesian split, was careful not to completely eradicate the dichotomy (Catani 2013: 94).

Indeed, this trend towards more nuanced formulations endures to this day. Philosopher David Chalmers, for instance, wrote that the "failure of materialism leads to a kind of *dualism*: there are both physical and nonphysical [i.e. mental] features of the world" (1996: 124). He speaks of *property* dualism (*ibid.*: 125) to

distinguish it from the discredited *substance* dualism of Descartes. Nonetheless, the essence of the dichotomy persists intact. Public endorsements of property dualism by influential science spokespeople, such as neuroscientists Christof Koch (2012a: 152) and Sam Harris,[1] lend academic legitimacy to it. Harris, for instance, claims that mind and matter each represent "half of reality," making the implicit assumption that they have comparable epistemic status (that is, that matter is as confidently knowable as mind). So pervasive is this assumption that it has become integral to our shared cultural intuitions.

Whilst a fundamental dichotomy between mind and matter is readily accepted by large segments of the population—perhaps for psychological reasons (Heflick *et al*. 2015)—in philosophical circles the corresponding dualism is properly regarded as unparsimonious. For this reason, philosophy has historically attempted to explain one member of the alleged dichotomy in terms of the other. The ontology of idealism, for instance, attempts to reduce "all sense data to mental contents" (Tarnas 2010: 335), whereas mainstream physicalism—perhaps better labelled as 'materialism,' but which I shall continue to refer to as 'mainstream physicalism' for the sake of consistency with some of the relevant literature—attempts to reduce all mental contents to material arrangements (Stoljar 2016). To be more specific, idealism entails that mind is nature's fundamental ontological ground, everything else being reducible to, or grounded in, mind, whereas mainstream physicalism posits that nature's fundamental ontological ground is matter outside and independent of mind, everything else being reducible to, or grounded in, matter.

The problem is that the ingrained cultural intuition that mind and matter have comparable epistemic status tends to creep—unexamined—even into philosophical thought, leading to the tacit conclusion that idealism and mainstream physicalism are mirror images of each other, in the sense of being structurally equivalent despite being reversely arranged. In the present essay, I contend that this tacit conclusion is false because it overlooks important epistemic considerations: we do *not*—and fundamentally *cannot*—know matter as confidently as we know mind. By incorrectly positing that idealism incurs an epistemic cost comparable to that of mainstream physicalism in at least some important sense, the tacit conclusion undervalues idealism and overvalues physicalism. This confusion may be a key enabler of physicalism's success in underpinning our present-day mainstream worldview. Once the tacit conclusion is properly examined and rectified, as attempted in this essay, idealism may emerge as a more plausible ontology than mainstream physicalism, at least in terms of its epistemic cost.

Like Gilbert Ryle (2009), I argue that mind and matter do *not* form a dichotomy. My argument, however, does not depend—as Ryle's controversially

[1] See Harris's video titled "You Are More Than Your Brain" on *Big Think*, 4 September 2016, available at https://www.facebook.com/BigThinkdotcom/videos/10153879575418527/.

does (Webster 1995: 483)—on equating mind with behaviours. Indeed, Ryle attempts to refute the alleged dichotomy by effectively relegating mind to the status of mere illusion (*ibid*: 461). My argument, instead, rests on the notion that mind and matter are not epistemically symmetrical—a concept I shall formally define in section 2.5—as members of a dichotomy must be. I do not deny mind, because it is epistemically primary: all knowledge presupposes mind.

That the notion of physically objective matter—that is, matter outside and independent of mind—is now largely taken for granted suggests cultural acclimatization to what is a mere hypothesis. After all, physically objective matter is not empirically observable, but a conceptual explanatory device *abstracted from* the patterns and regularities of empirical observations—that is, an *explanatory abstraction* (Glasersfeld 1987; more on this in section 2.4). Indeed, there seems to be a growing tendency in science today to mistake explanatory abstraction for what is available to us empirically. This has been extensively documented before, but mostly in regard to clearly speculative ideas such as superstring theory and multiverse cosmologies (Smolin 2007). When it comes to the everyday notion of physically objective matter, however, many fail to see the same conflation at work.

To illustrate and highlight the conflation with an admittedly extreme example, the next section briefly reviews the ontology of pancomputationalism, which posits ungrounded computation as the primary element of existence (Piccinini 2015). Indeed, the idea of replacing physicalism with ontic pancomputationalism should provide a visceral demonstration of the epistemic cost of substituting explanatory abstraction for empirical observation. In this context, my suggestion is that an analogous epistemic disparity exists between idealism and mainstream physicalism. In other words, if one is convinced that ontic pancomputationalism is absurd in comparison to physicalism, then—and on the same basis—one has reason to question the plausibility of mainstream physicalism in comparison to idealism.

Section 2.4 then elaborates more systematically on the different planes of abstract explanations used in science and philosophy. It provides the basis for the refutation of the alleged dichotomy between mind and matter later carried out in section 2.5, which forms the core of this essay. Finally, the Conclusion sums it all up.

Before we start, however, some terminology clarifications are needed. Throughout this essay, I use the word 'mind' in the sense of phenomenal consciousness. Following Thomas Nagel's (1974) original definition of the latter—which has since been further popularized by Chalmers (1996, 2003)—I stipulate that, if there is anything it is like to be a certain entity, then the entity is minded. As such, mind—as the word is used here—is epistemically primary, an assertion further substantiated in section 2.4. In this sense, mind does not necessarily entail higher-level functions such as metacognition—that is, the knowledge of one's knowledge (Schooler 2002: 340)—or even a conscious sense of self as distinct from the world. It necessarily entails only the presence of

phenomenal properties, in that it is defined as the substrate or ground of experience. Moreover, insofar as what we call 'concreteness' is itself a phenomenal property associated with the degree of clarity or vividness of experience, mind is the sole ground of concreteness. Anything allegedly non-mental cannot, by definition, be concrete, but is abstract instead, in the sense of lacking phenomenal properties.

I am well aware that the word 'mind' is used in entirely different ways—often decoupled from experience—in other contexts, such as philosophy of biology (Godfrey-Smith 2014) and artificial intelligence (Franklin 1997). Yet, I believe the usage I am defining here is adequate for the context of the present paper. And given this usage, experience can be coherently regarded as an excitation of mind, whereas mind can be coherently regarded as the substrate or ground of experience.

2.3 The epistemic cost of explanation by abstraction

By postulating a material world outside mind and obeying laws of physics, physicalism can accommodate the patterns and regularities *of* perceptual experience. *But it fails to accommodate experience itself*. This is called the 'hard problem of consciousness' and there is now a vast literature on it (e.g. Levine 1983, Rosenberg 2004: 13-30 and Strawson *et al*. 2006: 2-30). In a nutshell, the qualities of experience are irreducible to the parameters of material arrangements—whatever the arrangement is—in the sense that it is impossible, even in principle, to deduce those qualities from these parameters (Chalmers 2003).

As I elaborate in section 2.5, the "hard problem" is not merely hard, but fundamentally insoluble, arising as it does from the very failure to distinguish explanatory abstraction from empirical observation discussed in this paper. As such, it implies that we cannot, *even in principle*, explain mind in terms of matter. But because the contemporary cultural ethos entails the notion that mind and matter constitute a dichotomy, one may feel tempted to conclude that there should also be a symmetrical 'hard problem *of matter*'—that is, that we should not, even in principle, be able to explain matter in terms of mind. The natural next step in this flawed line of reasoning is to look for more fundamental ontological ground preceding both mind and matter; a *third* substrate to which matter and mind could both be reduced.

A good example of this line of reasoning is brought by ontic pancomputationalism, which posits that ungrounded information processing is what makes up the universe at its most fundamental level (Fredkin 2003). As such, ontic pancomputationalism entails that computation precedes matter ontologically. But "if computations are not configurations of physical entities, the most obvious alternative is that computations are abstract, mathematical entities, like numbers and sets" (Piccinini 2015). According to ontic

pancomputationalism, even mind itself—psyche, soul—is a derivative phenomenon of purely abstract information processing.[2]

To gain a sense of the epistemic cost of this line of reasoning, consider the position of physicist Max Tegmark (2014). According to him, "*protons*, *atoms*, *molecules*, *cells* and *stars*" are all redundant "baggage" (*ibid*: 255). Only the mathematical parameters used to describe the behaviour of matter are real. In other words, Tegmark posits that the universe consists purely of numbers—ungrounded information—but nothing to attach these numbers to. The universe supposedly is a "set of abstract entities with relations between them," which "can be described in a baggage-independent way" (*ibid*: 267). He attributes all ontological value to a description while—paradoxically—denying the existence of the very thing that is described in the first place.

Clearly, ontic pancomputationalism represents total commitment to abstract mathematical concepts as the foundation of existence. According to it, there are only numbers and sets. But what are numbers and sets without the mind or matter where they could reside? It is one thing to state in language that numbers and sets can exist without mind and matter, but it is another thing entirely to explicitly and coherently conceive of what—if anything—this may mean. By way of analogy, it is possible to *write*—as Lewis Carrol did—that the Cheshire Cat's grin remains after the cat disappears, but it is another thing entirely to conceive explicitly and coherently of what this means.

Ontic pancomputationalism appeals to ungrounded information—pure numbers, mathematical descriptions—as ontological primitive, i.e., as the sole fundamental aspect of existence. But what exactly is information? Our intuitive understanding of the concept has been cogently captured and made explicit by Claude Shannon (1948): information is given by state differences discernible in a system. As such, it is a property *of* a system—associated with the system's possible configurations—not an entity or ontological class unto itself. Under mainstream physicalism—that is, materialism—the system whose configurations constitute information is a material arrangement, such as a computer. Under idealism, it is mind, for experience entails different phenomenal states that can be qualitatively discerned from one another. Hence, information requires a mental or material substrate in order to be even conceived of explicitly and coherently. To say that information exists in and of itself is akin to speaking of spin without the top, of ripples without water, of a dance without the dancer, or of the Cheshire Cat's grin without the cat. It is a grammatically valid statement devoid of any semantic value: a language game less meaningful than fantasy, for internally consistent fantasy can at least be explicitly and coherently conceived of and, thereby, known as such. But in what way can we know information uncouched in mind or matter?

One assumes that serious proponents of ontic pancomputationalism are well aware of this line of criticism. How do they then reconcile their position with

[2] See Fredkin's online draft paper titled "*On the Soul,*" available at: http://www.digitalphilosophy.org//wp-content/uploads/2015/07/on_the_soul.pdf.

it? A passage by Luciano Floridi—well-known advocate of information as ontological primitive—may provide a clue. In a section titled "The nature of information," he states:

> Information is notoriously a polymorphic phenomenon and a polysemantic concept so, as an explicandum, it can be associated with several explanations, depending on the level of abstraction adopted and the cluster of requirements and desiderata orientating a theory. … *Information remains an elusive concept*. (Floridi 2008: 117, emphasis added)

Such ambiguity lends ontic pancomputationalism a kind of conceptual fluidity that renders it impossible to pin down. After all, if the choice of ontological primitive is given by "an elusive concept," how can one definitely establish that the choice is wrong? In admitting the possibility that information may be "a network of logically interdependent but mutually irreducible concepts" (*ibid.*: 120), Floridi seems to suggest, even, that such elusiveness may be unresolvable.

While vagueness may be defensible in regard to natural entities conceivably beyond the human ability to apprehend, it is at least difficult to justify when it comes to a *human concept* such as information. *We invented the concept*, so we either specify clearly what we mean by it or our conceptualization remains too ambiguous to be ontologically meaningful. In the latter case, there is literally *no sense* in attributing ontological value to information and, hence, ontic pancomputationalism is—once again—strictly meaningless.

Although ontic pancomputationalism is an admittedly extreme example, an analogous attempt to reduce concreteness—that is, the felt presence of conscious perception (Merleau-Ponty 1964)—to mere explanatory abstraction lies behind both mainstream physicalism and the alleged mind-matter dichotomy, as I shall argue in the next section. At the root of this concerning state of affairs is a generalized failure to recognize that every step of explanatory abstraction away from the concreteness of conscious perception implies a reduction in epistemic confidence: we do not know that abstract conceptual objects exist with the same level of confidence that we *do* know that our perceptions—whatever their source or underlying ontic nature may be—exist. I do not know that subatomic particles outside and independent of mind exist with the same level of confidence that I *do* know that the chair I am sitting on, which I am directly acquainted with through conscious perception, exists. Worse still, with what confidence can we know that a loosely defined, possibly incoherent concept such as ungrounded information lies at the foundation of existence? As such, steps of explanatory abstraction can only be justified if the relevant empirical observations cannot be explained *without* them, lest we conflate science and philosophy with meaningless language games. This is an important claim, so allow me to dwell on it a little longer before proceeding to the next section.

It could be argued that the existence of perceptual illusions indicates that conscious perception entails *less* epistemic confidence than abstract formal systems. For instance, in the well-known "checker shadow" illusion created by

the Perceptual Science Group of the Massachusetts Institute of Technology, two identically coloured squares—*A* and *B*—of a checkerboard are initially perceived to be of opposite colours because of the different contexts in which they are perceived (see Figure 2.1). Should we then declare that conscious perception is fundamentally unreliable? Well, notice that *it is also conscious perception that eventually dispels the illusion*: by looking at one of the squares as it is moved to the other's context, one sees that it indeed has the same colour as the other square. So even in the case of perceptual illusions, it is still direct, concrete experience that provides us with the epistemic confidence necessary to recognize the illusion for what it is.

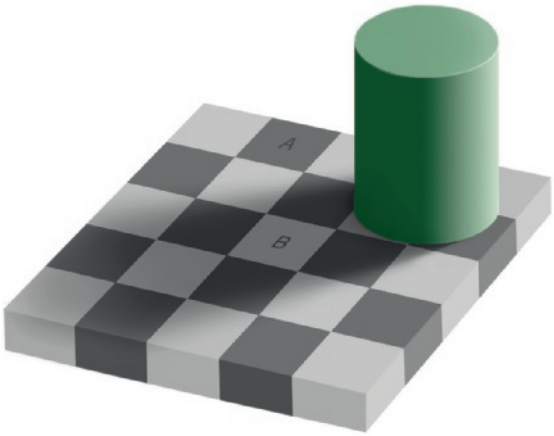

Figure 2.1: The "checker shadow" illusion. Despite appearances to the contrary, squares A and B are the same shade of grey.

Further supporting the claim that abstracting away from direct experience implies a reduction in epistemic confidence is the anti-realist view in philosophy of science. According to it, abstract theoretical entities—such as subatomic particles, invisible fields and any other postulated entity that escapes our ability to *directly* perceive—are but "convenient fictions, designed to help predict the behaviour of things in the observable world" (Okasha 2002: 61; see also van Fraassen 1990). In other words, the best we can say about subatomic particles and other abstract entities is that the observable world behaves *as if* these abstract entities existed. This does not entail or imply that the entities exist as such, which we cannot be certain of either way (van Fraassen 1980). In this sense, explanatory abstraction again implies reduction in epistemic confidence, insofar as we do not know that subatomic particles and invisible fields exist with the same level of confidence that we do know that the world we consciously perceive exists.

2.4 Levels of explanatory abstraction

Like ontic pancomputationalism, mainstream physicalism is no stranger to the epistemic cost of explanatory abstraction: the existence of a material world outside and independent of mind is a theoretical inference arising from *interpretation* of sense perceptions within a framework of complex thought, not an empirical observation. After all, what we call the world is available to us solely as 'images'—defined here broadly, so as to include any sensory modality—on the screen of perception, which is itself mental. Even physicist Andrei Linde, one of the founders of the theory of cosmic inflation, acknowledged this in a 1998 talk titled "Universe, Life, Consciousness," delivered at the Center for Theology and the Natural Sciences (CTNS), Berkeley, California:[3]

> Let us remember that our knowledge of the world begins not with matter but with perceptions. I know for sure that my pain exists, my 'green' exists, and my 'sweet' exists ... everything else is a theory. Later we find out that our perceptions obey some laws, which can be most conveniently formulated if we assume that there is some underlying reality beyond our perceptions. This model of material world obeying laws of physics is so successful that soon we forget about our starting point and say that matter is the only reality, and perceptions are only helpful for its description.

Now, we know that mind is capable of autonomously generating the imagery we associate with matter: dreams and hallucinations, for instance, are often qualitatively indistinguishable from the so-called 'real world.' Therefore, the motivation for postulating an objective material world must go beyond the mere existence of this imagery. And indeed, what the notion of objective matter attempts to make sense of are certain *patterns and regularities observable in the imagery*, such as:

- The correlations between observed brain activity and reported inner life (see, e.g. Koch 2004 for a scientific take on the neural correlates of consciousness, but consider also the obvious effects of e.g. alcohol consumption and head trauma—both of which disrupt regular brain activity—on inner experience);
- The observation that we all seem to inhabit the same world; and
- The observation that the dynamics of this world unfold independently of our personal volition.

After all, if mind is not a product of objective arrangements of matter, how can there be such tight correlations between brain activity and experience? If the world is not made of matter outside our individual minds, how can we all share the same world beyond ourselves? If the world is not independent of mind, why can we not change the laws of nature simply by imagining them to be different?

[3] At the time of this writing the transcript of this talk was available online at: http://web.stanford.edu/~alinde/SpirQuest.doc.

Clearly, thus, the non-mental world posited by physicalism is largely an attempt to make sense of these three basic observations. As such, it is an *explanatory abstraction*, not itself an observation. We conceptually *imagine* that there is a non-mental world underlying our perceptions—and in some sense isomorphic to these perceptions—because doing so helps explain the basic observations (see Figure 2.2). Nonetheless, whatever ontological class is pointed to by this conceptual abstraction remains perforce epistemically inaccessible, a recognition already present in Immanuel Kant's *Critique of Pure Reason*.

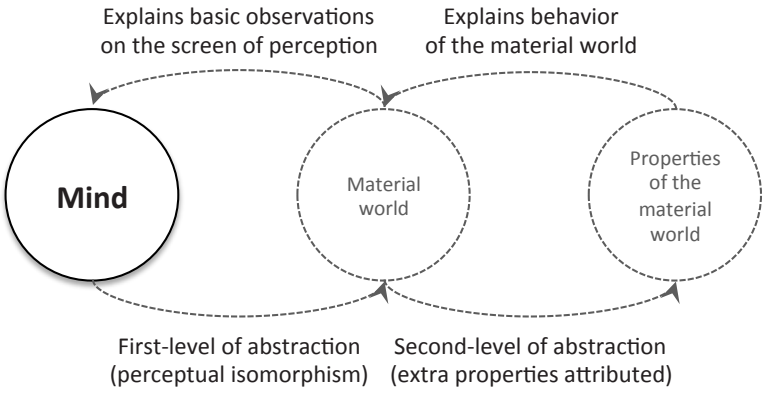

Figure 2.2: Levels of explanatory abstraction. Grey and dotted parts represent steps of abstraction.

Explanatory abstraction does not stop at this first level. After imagining a non-mental world isomorphic to our perceptions, we are left with the task of explaining how and why this world behaves the way it does. Why do objects fall when dropped? Why does a piece of amber attract chaff when rubbed? How can certain metals magnetically attract other metals? To answer these questions, we must attribute to the material world certain properties that go beyond perceptual isomorphism. We say, for instance, that matter has the properties of mass, charge and spin. These properties constitute a second-level of explanatory abstraction beyond direct experience (see Figure 2.2 again).

Naturally, there can be even more levels of explanatory abstraction involved. Superstring theory, for instance, attempts to explain the properties of matter through the particular modes of vibration of imagined hyper-dimensional strings (Greene 2003). But the two levels illustrated in Figure 2.2 are sufficient for the discussion that follows.

The defining characteristic of explanation by abstraction is a progressive movement away from Edmund Husserl's (1970) "life-world," from the concreteness of direct experience. First, one posits a world devoid of qualities (Varela, Thompson and Rosch 1993) and, as such, devoid of concreteness too,

for concreteness is a quality of experience. Then, one progressively loads this world with properties that entail no direct isomorphism to experience. For instance, we do not see electric charge or spin; we only see the behaviour of matter that these abstract properties supposedly explain, such as attraction and repulsion. Similarly, we do not feel mass; we only feel the weight and inertia of objects, which the property of having mass supposedly explains (Okasha 2002: 58-76).

Because concreteness is the intuitive foundation of what we consider *real*, each step in this movement away from concreteness takes us farther from what we intuitively sense to be real (Merleau-Ponty 1964). One may then become lost in a forest of intellectually appealing but ultimately arbitrary conceptualizations. This, again, is the epistemic cost of explanation by abstraction.

2.5 Dispelling the mind-matter dichotomy

By definition, the two members of a dichotomy are jointly exhaustive and mutually exclusive. Ontologically, this means that if one member is the case, then the other is necessarily *not* the case, and vice-versa. For instance, in the context of biological organisms, if life is *not* the case, then death is necessarily the case. In the context of a job application, if success is the case (i.e. the applicant gets the job), then failure is *not* the case. And so on. As such, a *single test* suffices to acquire knowledge about the ontological status of *both* members of a dichotomy. If I can perform a test to determine if a person is alive, then I will automatically know whether the person is dead, without having to test for death separately. If I can set a criterion for success, then that same criterion will automatically determine whether failure is the case, without my having to set a separate criterion for failure. And so on. I shall call this property of a dichotomy *epistemic symmetry*. When two concepts are epistemically symmetrical, knowledge of one implies knowledge of the other.

Now notice that *epistemic symmetry can only hold for concepts residing in the same level of explanatory abstraction*. If they do not, then there necessarily is at least one extra inferential step necessary to know whether one of the concepts obtains. This breaks the symmetry, for then we cannot acquire knowledge of the ontological status of both concepts with a single test.

Here is an example: the presence of a negative feeling can be tested for directly through introspection—thus entailing no inferential steps—whereas testing for the presence of a positive electric charge requires an inference by observation of the associated behaviour of matter. Because of this need for an extra inferential step, knowing the negative feeling cannot imply knowledge of the positive electric charge. The negative feeling and the positive electric charge are not, therefore, epistemically symmetrical and cannot constitute a dichotomy.

Conversely, positive and negative electric charges are both properties of matter, residing in the second level of explanatory abstraction illustrated in Figure 2.2. As such, they are epistemically symmetrical and can constitute a dichotomy. Indeed, every level of explanatory abstraction can encompass dichotomies. For

instance, the size of material objects is isomorphic to perceptual qualities: we can subjectively test whether an object is big or small in relation to another object. As such, bigness and smallness both reside in the first level of explanatory abstraction and are epistemically symmetrical; they can constitute a dichotomy (see Figure 2.3).

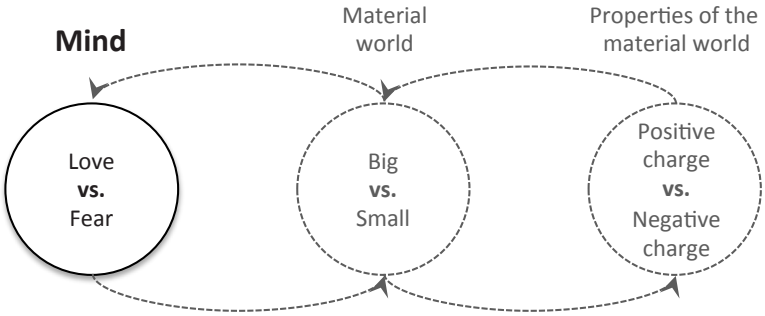

Figure 2.3: Dichotomies in their respective levels of explanatory abstraction.

But—and here is the key point—*mind and matter do not reside in the same level of explanatory abstraction*. Mind—as defined in Section 2.2—is the ground within which, and out of which, abstractions are made. Matter, in turn, is an abstraction *of* mind (see Figure 2.2 again). This breaks the epistemic symmetry between them: we do not know matter in the same way that we know mind, for—as cogently argued by Linde in the earlier quote—matter is an inference and mind a given. Consequently, although mind can encompass polar opposites—such as the feelings of love and fear in the context of a situation where someone feels passionate about a particular aspect of someone else (assuming that other passions, such as hate, which is arguably a form of fear, are particular instances of love or fear)—it cannot itself be the polar opposite of matter or matter's properties. It follows that we have no reason to conclude that reducing matter to mind is as challenging as reducing mind to matter, and there is thus no substantiation for a 'hard problem of mind.' Stronger still, insofar as what we call 'matter' can be parsimoniously construed as phenomenal patterns of excitation of mind, matter is on an epistemic par with mind and can, in principle, be reduced to the latter, for both already reside in the same ontological domain. This move takes mind itself to be an ontological primitive and eliminates any conceivable 'hard problem of mind,' since mind now does not need to be reduced.

The notion of a dichotomy between mind and matter arises from language. In order to speak of the substrate of experience we must give it a name, such as 'mind' or 'consciousness,' thereby linguistically objectifying the subject. Then, we conflate language with what language attempts to describe, implicitly assuming that mind is an object just as matter allegedly is. We forget that there is no epistemic symmetry between the two.

Indeed, because the concept of mind-independent matter, as an explanatory abstraction, arises *in mind*, as an 'excitation' *of* mind, to say that mind and matter constitute a dichotomy is akin to saying that ripples and water constitute a dichotomy.[4] Dichotomies can exist only between different kinds of ripples—say, those that flow mostly to the right versus those that flow mostly to the left—not between ripples and the substrate where they ripple. Mind is the substrate of the explanatory abstraction we call matter, so when we speak of a mind-matter dichotomy we fall into a fundamental "category mistake," as Ryle (2009) put it. *However, contrary to what Ryle suggests, it is matter that is the abstraction, not mind.*

The notion that idealism and mainstream physicalism are mirror images of each other arises from a failure to grasp this point. Lucid contemplation of these ontologies shows that idealism attempts to reduce an explanatory abstraction (physically objective matter) to that which articulates and hosts the abstraction in the first place (mind). This is *prima facie* eminently reasonable. Mainstream physicalism, in turn, attempts to reduce mind to mind's own explanatory abstractions, an obvious paradox that constitutes the crux of the 'hard problem.'

There would be no 'hard problem' if one did not conflate explanatory abstractions with concrete ontological primitives, if one did not attempt to paradoxically reduce mind to abstractions *of* mind. The 'hard problem' is not something empirically observed but the salient result of internal contradictions in a logico-conceptual schema; contradictions that I hope to have helped make explicit with the present paper.

Naturally, circumventing the 'hard problem' in the way suggested above ultimately forces us to make do with mind alone as an ontological primitive and thereby entertain some form of idealism—more specifically, a form of idealism wherein mind is the experientially given ground of existence, whose manifestations comprise the concrete phenomenality you and I undergo in everyday life. And whereas idealism in the West has had its heyday in the eighteenth (e.g. Berkeley) and early nineteenth (e.g. Hegel) centuries, it is now enjoying renewed interest (Chalmers 2018) for having been updated and revitalized with compelling new formulations (e.g. Kastrup 2017b and 2017e,[5] Yetter-Chappell 2018, as well as Fields *et al*. 2017, insofar as the latter can be construed as a form of idealism). These are sometimes proposed under new names, such as 'cosmopsychism' (e.g. Shani 2015, Nagasawa and Wager 2016), which, as the name suggests, posits that the cosmos as a whole is essentially

[4] Allow me to insist on an observation already made in Section 1.4: this is an *epistemic* point. Mind-independent matter can only be known as a concept that arises in mind, for we have no direct access to the hypothesized ontological category the concept denotes. Therefore, *insofar as we can know it*, mind-independent matter is an excitation of mind and, as such, cannot form a dichotomy with mind for the same reason that ripples cannot form a dichotomy with water.

[5] See also Chapter 3 of this dissertation.

phenomenal. Even 'radical constructivism' can be construed as a form of idealism, insofar as its claims are not merely epistemic, but ontic: "Radical constructivism … develops a theory of knowledge in which knowledge does not reflect an 'objective' ontological reality, but *exclusively* an ordering and organization of a world *constituted by* our experience" (Glasersfeld 1987: 24, emphasis added). Finally, the strongest objections usually leveraged against idealism have recently also been tackled (Kastrup 2017c[6]).

Having said all this, it should be noted that, in and of itself, the argument provided in this paper, despite being supportive of idealism, does not necessarily *imply* idealism. I have focused on epistemic cost considerations and did not show whether or how idealism can account for all relevant empirical observations we make of nature. Indeed, an articulation of an idealist ontology is not within the scope of this paper. But if it is demonstrated—as some of the papers cited above claim to do—that idealism *can* account for all empirical observations that mainstream physicalism allegedly accounts for, then epistemic cost considerations certainly tilt the balance in favour of idealism, due to the latter's lack of reliance on inflationary, epistemically unreliable, paradoxical abstractions. As such, the core claim of this essay is not so much the validity of idealism as that physically objective matter is a doubtful *cognitive construct*, in the strict constructivist sense: insofar as we believe to see matter outside and independent of mind when we look at the world around ourselves, we are conflating a rational-linguistic construction with what is empirically observed.

2.6 Conclusion

The pervasive but unexamined assumption that mind and matter constitute a dichotomy is an error arising from language artifacts. Members of dichotomies must be epistemically symmetrical and, therefore, reside in the same level of abstraction. Physically objective matter—as an explanatory model—is an abstraction *of* mind. We do not *know* matter in the same way that we know mind, for matter is an inference and mind a given. This breaks the epistemic symmetry between the two and implies that mainstream physicalism and idealism cannot be mirror images of each other.

Failure to recognize that different levels of epistemic confidence are intrinsic to different levels of explanatory abstraction lies at the root not only of the false mind-matter dichotomy, but also of attempts to make sense of the world through increasingly ungrounded explanatory abstractions. Lest we conflate science and philosophy with hollow language games, we must never lose sight of the difference between an abstract inference and a direct observation. Keeping this distinction in mind allows us to construct useful predictive models of nature's *behaviour*—which ultimately is what science is meant to do—without restrictive and ultimately fallacious inferences about what nature *is*.

[6] See Chapter 4 of this dissertation.

This, in turn, liberates us from thought artifacts such as the “hard problem of consciousness” and opens up whole new avenues for making sense of self and world.

3. The Universe in Consciousness

This paper first appeared in the *Journal of Consciousness Studies*, ISSN 1355-8250 (print), Vol. 25, No. 5-6, pp. 125-155, in June 2018. A summary of its core idea has appeared in *Scientific American* on 18 June 2018.[1]

3.1 Abstract

I propose an idealist ontology that makes sense of reality in a more parsimonious and empirically rigorous manner than mainstream physicalism, bottom-up panpsychism and cosmopsychism. The proposed ontology also offers more explanatory power than these three alternatives, in that it does not fall prey to the hard problem of consciousness, the combination problem or the decombination problem, respectively. It can be summarized as follows: there is only cosmic consciousness. We, as well as all other living organisms, are but dissociated alters of cosmic consciousness, surrounded by its thoughts. The inanimate world we see around us is the extrinsic appearance of these thoughts. The living organisms we share the world with are the extrinsic appearances of other dissociated alters.

3.2 Brief introduction

This paper seeks to articulate an ontology that overcomes the principal limitations of the most popular alternatives. The first half of the paper comprises a detailed analysis of relevant literature, highlighting what advances have been made and what problems have been created or left unsolved by recent developments in analytic philosophy. In the second half, starting from what I consider to be the most promising current platform, I propose an idealist framework that may open viable new avenues for addressing the key questions left unanswered by this current platform. At the end, I hope to offer a coherent view of the nature of reality that accounts for all relevant facts without incurring any fundamental problem.

3.3 The mainstream physicalist ontology and its problems

The mainstream ontology of physicalism posits that reality is constituted by irreducible entities—which, like Galen Strawson (2006: 9), I shall call 'ultimates'—outside and independent of phenomenal consciousness. These ultimates, in and of themselves, do not instantiate phenomenal properties: there is nothing it is like to be an ultimate, the capacity for experience

[1] At the time of this writing, the *Scientific American* essay was freely available online at: https://blogs.scientificamerican.com/observations/could-multiple-personality-disorder-explain-life-the-universe-and-everything/.

emerging only at the level of complex arrangements of ultimates. They are also sometimes held to lack objective qualities: in and of themselves, ultimates may have no color, flavor, smell, etc. Indeed, according to mainstream physicalism qualities may exist only in the phenomenal field of the experiencer, which in turn is a product of the operation of a sufficiently complex nervous system. It is the specific arrangement of ultimates in a nervous system that, allegedly, somehow constitutes or generates its phenomenal properties.

The key problem of mainstream physicalism centers on how our subjective experience of qualities—what it is like to feel the warmth of fire, the redness of an apple, the bitterness of disappointment, etc.—can arise from mere arrangements of ultimates. These ultimates do possess abstract relational properties such as mass, spin, momentum and charge, but there is nothing about mass, spin, momentum or charge, or the relative positions and interactions across ultimates, in terms of which one could deduce what the warmth of fire, the redness of an apple or the bitterness of disappointment feel like, subjectively. As long as they fit with the broadly observed correlations between neural activity and reported experience, mappings between these two domains are entirely arbitrary: in principle, it is as (in)valid to state that spin up constitutes or generates the phenomenal property 'coldness' and spin down 'warmth' as it is to say the exact opposite. There is nothing intrinsic about spin—or about any other property of ultimates or arrangements thereof—that would allow us to make the distinction.

This central—and arguably insoluble—problem has been referred to by different names, such as the 'explanatory gap' (Levine 1983) and, more recently, the 'hard problem of consciousness' (Chalmers 1996, 2003): the qualities of experience are irreducible to the observable parameters of physical arrangements—whatever the arrangement may be—in the sense that it is impossible even in principle to deduce those qualities from these parameters. More generally, the argument here is that there is no *entailment* from facts about ultimates to facts about experience: there is no fact about ultimates that implies *a priori* a fact about experience.

Greg Rosenberg (2004: 13–30) articulated what is perhaps the best refutation of entailment from facts about ultimates to facts about experience. His argument begins with the recognition that all facts about ultimates are merely patterns of bare differences. This echoes Bertrand Russell's point (2009) that science can only characterize things and phenomena in terms of how they differ from other things and phenomena. For instance, an ultimate with positive electric charge is characterized in terms of how its relevant behavior differs from that of a negatively charged ultimate. Charge is thus a relational property defined on the basis of bare differences. Nothing can be scientifically stated about what a charge, in and of itself, intrinsically is. The same can be argued about all other facts about ultimates.

Rosenberg then proceeds to show that facts about experience—phenomenal properties—cannot be entailed by patterns of bare differences, even though qualitative differences between experiences can admittedly instantiate a

structure of bare differences. Therefore, phenomenal properties cannot be reduced to facts about ultimates. Allow me to unpack this.

There are qualitative differences across our experiences of various colors: what it is like to see yellow is different from what it is like to see red. These qualitative differences can even be graded along relevant dimensions: the qualitative difference between seeing yellow and red seems bigger than the qualitative difference between seeing yellow and orange. If one were to assign a number to represent each of these degrees of difference, one could abstract out a purely quantitative—that is, bare—difference structure from the experiences of seeing various colors. However, that a bare difference structure can be abstracted out from phenomenal properties does not imply that phenomenal properties are entailed by bare difference structures. Maintaining so inverts the logic of the situation: it is phenomenal properties that ground bare difference structures in the first place.

To bring this point home, Rosenberg offers the following thought experiment: imagine a field of tightly packed yellow and red dots. If one observes this field from a sufficient distance, one sees the color orange. It could then be argued that the phenomenal property 'orange' arises from a pattern of bare differences associated with the delta in wavelength between yellow and red photons, as well as the relative size and distribution of the dots. However, if one were to choose another pair of colors with the same delta in wavelength—say, yellow and green—and otherwise maintain the same relative structure of dots, a phenomenal property different from 'orange' would result. In other words, the same pattern of bare differences would yield a different phenomenal property. Hence, phenomenal properties are not entailed by patterns of bare differences and cannot be reduced to properties and arrangements of ultimates.

This and other arguments along similar lines render mainstream physicalism arguably untenable.

3.4 Consciousness as an irreducible property of matter

At least since the time of René Descartes, the most recognizable alternative to physicalism has been 'substance dualism': if one cannot reduce phenomenal properties to physical elements, then the phenomenal and the physical may be two distinct, fundamental ontological classes. There are different versions of substance dualism, but the most intuitive one is arguably 'interactionism': since phenomenal events seem to cause physical events (as in when felt pain causes me to move my arm) and vice versa (as in when a needle piercing my arm causes me to feel pain), then the phenomenal and the physical must be causally connected. However, a problem with interactionism is summarized by Chalmers (2016b: 23): if the physical domain is causally closed—as it seems to be in so far as we have been able to ascertain through the scientific method—then causal influences we intuitively attribute to the phenomenal domain must ultimately be, in fact, physical. There is arguably no place for phenomenal properties in the causal nexus. Possible dualist answers to this have been

proposed but, as acknowledged by Chalmers himself (who admits to sympathy towards dualism), "there is at least a prima facie case against dualism here" (*ibid*.: 24). Chalmers then posits an "Hegelian synthesis" (*ibid*.) between mainstream physicalism and substance dualism, in the form of the notion that ultimates themselves may be *fundamentally* conscious.

Indeed, under mainstream physicalism, ultimates are elementary subatomic particles—quarks, leptons, gauge bosons and scalar boson(s)—with certain fundamental properties. These properties are relational and abstract, such as mass, charge, spin and momentum. Mainstream physicalism's key problem, as we have seen, is its inability to account for phenomenal properties. So the most straightforward way out is to posit that at least some elementary particles *also* have fundamental phenomenal properties. In Strawson's words, "Assuming, then, that there is a plurality of physical ultimates, some of them at least must be intrinsically experiential, intrinsically experience-involving" (2006: 24).

I shall call these experiencing elementary particles 'phenomenal ultimates'. I shall also generally refer to the broad ontological outlook described above as 'bottom-up panpsychism,' even though I am aware that there are many variations of it that would be better discriminated from one another (such as 'panexperientialism,' 'constitutive micropsychism,' 'panprotopsychism,' 'deferential monadic panpsychism,' etc.). Be that as it may, the key general idea here is that, by positing phenomenal properties to be fundamental, bottom-up panpsychism evades the need to reduce these properties and thereby avoids the hard problem altogether. Moreover, bottom-up panpsychism places these new fundamental properties seamlessly alongside existing abstract relational properties, as the categorical basis of the latter. This neatly integrates phenomenal properties in the framework of scientific thinking, for they now occupy a proper place within the causal nexus.

To see why this seemingly elegant approach nonetheless fails, notice that, according to bottom-up panpsychism, the unitary phenomenal life of a human being is supposedly constituted by micro-level phenomenal parts. At some point in the remote past phenomenal ultimates

> organized into increasingly complex forms, both experiential and non-experiential, by many processes including evolution by natural selection. And just as there was spectacular enlargement and fine-tuning of non-experiential forms (the bodies of living things), so too there was spectacular enlargement and fine-tuning of experiential forms. (*ibid*.: 27)

However, the idea that micro-level phenomenal states can combine to form unitary macro-level phenomenal states is arguably incoherent. It leads to a variety of 'combination problems' (Chalmers 2016a), at least one of which is arguably as insoluble as the hard problem itself (Carruthers and Schechter 2006, Goff 2006, 2009).

The best argument against bottom-up panpsychism is perhaps Sam Coleman's (2014). As bottom-up panpsychists themselves seem to agree, "'experience is impossible without an experiencer,' a *subject of experience*" (Strawson 2006: 26,

emphasis added). Therefore, bottom-up panpsychism implies that each phenomenal ultimate, by virtue of bearing phenomenal properties, instantiates a micro-level *subject*. Moreover, it implies that macro-level subjects with a seemingly unitary perspective, such as you and me, must somehow arise through some form of bottom-up combination of micro-level subjects. This is called the 'subject combination problem.'

Coleman connects subjectivity with the presence of a perspective, or *point of view*:

> That a given subject has a particular phenomenological point of view can be taken as saying that there exists a discrete 'sphere' of conscious experiential goings-on corresponding to this subject, with regard to which other subjects are distinct in respect of the phenomenal qualities they experience, *and they have no direct (i.e. experiential) access to the qualitative field enjoyed by the first subject*. A subject, then, can be thought of as a *point of view* annexed to a *private* qualitative field. (Coleman 2014: 30, emphasis added)

Notice Coleman's emphasis on the *private* character of the qualitative field annexed to a subject. I shall return to this point later.

Bottom-up panpsychism attempts to model the combination of phenomenal states after the way ultimates combine physico-chemically. After all, the force and appeal of its argument rests on the analogous treatment of phenomenal properties and standard physical properties such as mass, spin and charge. Therefore, Coleman also makes explicit what *combination* means in this physico-chemical sense:

> Combination, thus, is 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. (*ibid*.)

For instance, an oxygen and two hydrogen atoms *combine* to form a water molecule: they become intrinsically altered in the process of forming covalent bonds with one another, but continue nonetheless to exist in the resulting molecule.

In this framework, bottom-up panpsychism implies that the private point of view of each phenomenal ultimate that constitutes you becomes intrinsically altered in the process of combining to form the private point of view you enjoy right now—that is, your "unique experiential portal to reality," as put by Itay Shani (2015: 399). *But each must nonetheless continue to exist in you*, just as quarks continue to exist in protons, protons continue to exist in oxygen atoms, and oxygen atoms continue to exist in water molecules.

However, Coleman argues, "points of view cannot combine" in this manner (2014: 32). If a first constituent lower-level subject sees, say, only blue, and a second sees only red, then only the *qualitative contents* of their respective perspectives can conceivably survive—possibly in altered form—as combined ingredients of the resulting higher-level subject's phenomenal field (e.g. if the latter sees purple). But the original constituent *points of view* cannot survive,

for they entail seeing *only* red and *only* blue, respectively. Since the resulting higher-level subject has, *ex hypothesi*, a single compound portal to reality, it cannot *both* see only red *and* only blue. At least one of the constituent lower-level points of view will thus necessarily disappear—in fact, *both* will disappear if the higher-level subject sees purple—which is not consistent with combination in the physico-chemical sense.

One may argue that what happens instead is that the phenomenal state of the higher-level subject "is a novel state which in some way 'absorbs' or supersedes the mental states of the constituents" (Seager 2010: 179). In this so-called "combinatorial infusion" (*ibid.*) scenario, the lower-level points of view cease to exist in the process of forming the compound higher-level one. By parting with combination in the physico-chemical sense, this scenario negates much of the force and appeal of the bottom-up panpsychist argument. But the panpsychist can be spared this regret, for—as Coleman argues—the scenario does not work anyway.

Coleman's reasoning is that, to avoid the appeal to magic entailed by brute or strong emergence, "lower-level properties must contribute to their novel product *in virtue of their metaphysical nature*, or, otherwise put, *while remaining true to what they are*" (2014: 35, original emphasis). But "a set of points of view have nothing to contribute as such to a single, unified successor point of view. Their essential property defines them against it: in so far as they are points of view they are experientially distinct and isolated" (*ibid.*: 37, original emphasis). So the resulting higher-level point of view cannot be explicated in terms of the lower-level constituent points of view.

In conclusion, bottom-up panpsychism fails because there is no explicit and coherent way to ground the existence of macro-level subjects in micro-level phenomenal ultimates. Subject combination arguably requires—just as mainstream physicalism does—the appeal to magic entailed by brute or strong emergence. Yet, it was precisely this requirement that, in the case of mainstream physicalism, motivated the conception of bottom-up panpsychism as an alternative in the first place.

3.5 The questionable logical bridge in bottom-up panpsychism

Bottom-up panpsychism is motivated by the idea that, since physics only models the *behavior* of physical entities and says nothing about their *intrinsic nature* (Russell 2009), phenomenal consciousness may be this intrinsic nature. This is eminently reasonable, since the only physical entity we are acquainted with 'from within' is our own nervous system, whose intrinsic nature surely seems to be phenomenal (Eddington 1928). But bottom-up panpsychism then makes an *extra* claim: that phenomenal consciousness has the same *fragmented structure* that matter has on the screen of perception. In other words, since our body is constituted by myriad elementary particles insofar as we can perceive it, our phenomenal inner-life must itself be constituted by micro-level phenomenal parts—or so the argument goes.

This extra claim rests on a questionable logical bridge: it attributes to *that which experiences* a structure discernible only in the *experience itself*. Allow me to elaborate.

The concept of elementary particles—ultimates—arises from experiments whose outcomes are accessible to us only in the form of conscious perception (even when delicate instrumentation is used, the output of this instrumentation is only available to us as conscious perception). Such experiments show that the images on the screen of perception can be divided up into ever-smaller elements, until we reach a limit. At this limit, we find the smallest discernible constituents of the images, which are thus akin to pixels. As such, ultimates are the 'pixels' *of experience*, not necessarily of the experienc*er*. The latter simply does not follow from the former.

Therefore, that human bodies are made of elementary particles does not necessarily say anything about the structure of the experienc*er*: a human body is itself an image on the screen of perception, and so will necessarily be 'pixelated' insofar as it is perceived. Such pixelation reflects the idiosyncrasies *of the screen of perception*, not necessarily the structure of the human *subject* itself. As an analogy, the pixelated image of a person on a television screen reflects the idiosyncrasies of the television screen; it does not mean that the person herself is made up of pixels.

As suggestive as it may be, the hypothesis that phenomenal consciousness is the intrinsic nature of the physical does not imply that the fragmented structure of matter on the screen of perception is the fundamental structure of phenomenal consciousness itself.

3.6 What counts as a fundamental concrete entity?

We have seen in the previous section that elementary particles are the building blocks or 'pixels' of what is perceived, not necessarily of the subject that perceives. But we can ask a yet deeper question: Are elementary particles fundamental concrete entities on their own merit? Both mainstream physicalism and bottom-up panpsychism, in taking ultimates to be the discrete building blocks of nature, seem to assume so.

There are, however, strong reasons to believe that at least the entire inanimate universe is one integrated whole without ultimate parts. Jonathan Schaffer, for instance, points out that,

> physically, there is good evidence that the cosmos forms an entangled system and good reason to treat entangled systems as irreducible wholes. Modally, mereology allows for the possibility of *atomless gunk*, with no ultimate parts for the pluralist to invoke as the ground of being. (2010: 32, original emphasis)

Terry Horgan and Matjaž Potrč (2000) also contended that only the universe as a whole can be considered a concrete entity on its own merit, which they called the 'blobject.'

The physical substantiation for this line of thought is not recent. As early as in the 1930s, John von Neumann (1996) reasoned that, when two inanimate quantum systems interact, no measurement is actually performed but, instead, the two systems become entangled with one another, forming an indivisible whole. If the resulting whole then interacts with a third system, they, too, become entangled, forming a new and larger whole; and so forth. These are the so-called 'von Neumann chains' and, since everything in the universe ultimately is a quantum system, the entire inanimate universe must constitute one single von Neumann chain—that is, one indivisible whole (von Neumann also remarked that observation by a conscious, living human being clearly breaks the chain, since living humans demonstrably *can* perform a quantum measurement. Therefore, conscious living beings must be left out of the present argument).

The implication is that, physically, there are arguably no such things as fundamental microscopic ultimates. Although this may violate popular assumptions and intuitions, it also points the way to a third avenue of enquiry that holds some promise as an alternative to both mainstream physicalism and bottom-up panpsychism.

3.7 The whole universe as a unitary conscious entity

The idea that the (inanimate) universe may be an indivisible whole has proven tempting to those seeking an alternative to bottom-up panpsychism, so to avoid the subject combination problem: they posit that "*the cosmos as a whole is the only ontological ultimate there is*, and that it is *conscious*" (Shani 2015: 408, original emphasis). With this, there is no longer any need to explicate how lower-level subjects combine to form higher-level subjects, for the highest-possible-level subject is already the starting point.

This general outlook is called 'cosmopsychism' (Mathews 2011, Jaskolla and Buck 2012, Shani 2015, Nagasawa and Wager 2016). The seminal insight that freed cosmopsychism from the limitations of bottom-up panpsychism was arguably that of Freya Mathews (2011): she realized that, even under the hypothesis that phenomenal consciousness is the intrinsic aspect of the physical, there is no need to attribute the fragmented structure of matter to phenomenal consciousness itself. In her words, "an extension of subjectivity to physical reality generally [i.e. "force fields and even space itself"], rather than its restriction merely to matter, does seem to be required" (*ibid*.: 144).

Now the problem cosmopsychists face is the 'decombination problem' (also called the 'decomposition problem' in Chalmers 2016a): How do seemingly separate lower-level subjects—which, from now on, I shall follow Shani (2015: 415) in referring to as 'relative subjects'—form within the conscious cosmos? To paraphrase Coleman (2014: 30), how do they acquire their *private* point of view, whose associated qualitative field other relative subjects have no direct—that is, experiential—access to? After all, I cannot read your thoughts and, presumably, neither can you mine.

Before we address this problem, however, notice that there are at least two possible interpretations of cosmopsychism. The first one sticks to the bottom-up panpsychist view that a phenomenal ultimate has both phenomenal *and non-phenomenal* properties. This way, whereas it takes the cosmos as a whole to be the sole phenomenal ultimate there is, this interpretation grants that the abstract relational properties of the cosmos are not phenomenal. For this reason, I shall call this interpretation 'dual-aspect cosmopsychism.' According to it, the intrinsic aspect of the cosmos is phenomenal, but its extrinsic aspect—the physical structure we can objectively measure in a scientific sense—is non-phenomenal and circumscribes the cosmos's phenomenal field. In a sense, the extrinsic, physical aspect of the cosmos *bears* phenomenality within in.

Another interpretation of cosmopsychism entails that *the sole ontological primitive there is is cosmic phenomenal consciousness*—or simply 'cosmic consciousness' for ease of reference. Nothing exists outside or independent of cosmic consciousness. As such, under this interpretation one should say that the cosmos is *constituted by* phenomenality, as opposed to *bearing* phenomenality. In other words, here the perceivable cosmos is *in consciousness*, as opposed to being *conscious*.

The latter interpretation is Shani's (2015) position. Indeed, according to him the external, physical aspect of the cosmos is 'its *appearance* as an exterior complement to … subjective realities' (*ibid*.: 412, emphasis added). Appearances are, of course, phenomenal in nature. I shall thus call this interpretation 'idealist cosmopsychism,' since its reduction base is purely phenomenal.

Shani does still postulate a duality in cosmic consciousness to account for the clear qualitative differences between the outer world we, as relative subjects, perceive and measure and the inner world of our thoughts and feelings. He calls it the 'lateral duality principle' (*ibid*.: 410) and describes it thus:

> [Cosmic consciousness] exemplifies a dual nature: it has a *concealed* (or enfolded, or implicit) side to its being, as well as a *revealed* (or unfolded, or explicit) side; the former is an intrinsic dynamic domain of creative activity, while the latter is identified as the outer, observable expression of that activity. (*ibid*., original emphasis)

What is important to emphasize, though, is that this duality does not entail or imply two distinct ontological classes. Everything is still phenomenal.

Now, one must ultimately ground the revealed side of the cosmos in its concealed side, not only to eliminate what would otherwise be an arbitrary boundary, but also to accommodate the empirically undeniable causal links between the revealed order of the physical world we perceive and the concealed order of thoughts and feelings. After all, revealed physical things and phenomena—think of psychoactive drugs, bodily trauma, electromagnetic fields, etc.—causally affect our concealed thoughts and feelings. Causal links operating the other way around are also undeniable: our thoughts and feelings can lead to physical manifestations in the form of bodily behaviors. If the

revealed order were not grounded in the concealed, but constituted a separate phenomenal domain instead, how could these cross-influences take place? Indeed, Shani acknowledges as much when he writes that "the revealed order of reality is grounded in the concealed" (*ibid*.: 416).

Yet, Shani is not explicit in regards to *how* this grounding works. He states that the physical world we perceive is the way the structural patterns of the creative activity of cosmic consciousness are represented in relative subjects, such as you and me (*ibid*.: 412). This is fair enough as far as it goes, but what is the mechanism of representation whereby concealed phenomenal activity translates into revealed order from the perspective of relative subjects? How does the formation of a relative subject lead to such a significant qualitative transition as the representation of thoughts and feelings (the concealed order) in the form of perception (the revealed order)?

To tackle the decombination problem, Shani posits that the conscious perspective or point of view of each relative subject has both a specific and a generic character (*ibid*.: 423). Since a relative subject corresponds to a segment of cosmic consciousness, its specific character is derived from the *local pattern of phenomenal activity* taking place in that segment. Its generic character, in turn, is derived from the *intrinsically subjective, perspectival nature* of cosmic consciousness as a whole. Let me unpack this.

Shani posits two intrinsic features of cosmic consciousness as constituents of the generic character of each relative subject: *sentience* and *core-subjectivity* (*ibid*.: 426). In other words, each relative subject is phenomenally conscious by virtue of the fact that cosmic consciousness is itself intrinsically capable of experience. Also, each relative subject has 'ipseity, or I-ness, by which is meant an implicit sense of self which serves as the dative … of experience, namely, as *that to whom things are given*, or disclosed, from a perspective' (*ibid*., original emphasis). The claim is then that the sense of I-ness of each relative subject is the sense of I-ness intrinsic to cosmic consciousness as a whole. One could argue that sentience and core-subjectivity, so defined, are inextricable from one another. But even in this case, it is still useful to distinguish between these two cognitively salient aspects of what would admittedly be a single intrinsic feature of cosmic consciousness. So I shall continue to speak of sentience and core-subjectivity.

In summary, according to Shani a relative subject is grounded, on the one hand, in the intrinsic sentience and core-subjectivity of cosmic consciousness as a whole and, on the other hand, in the local patterns of phenomenal activity taking place in the particular segment of cosmic consciousness associated with the relative subject. The question now is: What are these local patterns of phenomenal activity that give rise to a *private* qualitative field, *inaccessible to other relative subjects*, as required by Coleman (2014)?

Shani posits that the smallest *cohesive* elements of nature correspond to the revealed appearance of micro-level relative subjects (2015: 415-16). In other words, *he returns to the bottom-up panpsychist view that elementary particles are subjects*. Shani motivates this with a metaphor:

> A relative [subject] is a 'vortex' surging from the oceanic background [of cosmic consciousness]. It is a cohesive system. (*ibid*.)

He uses the image of a vortex to refer to localization of phenomenal activity.

> Consider … the *most elementary* 'vortices'. [Their corresponding] localization process consists, then, in the intensification of experience, as well as in the concentration of focus, within limited and relatively well-defined boundaries … which serves to separate the system's inner reality from the inner reality of the ocean surrounding it … The result is an individual self (*however primitive*) engulfed in its own experiences and concerns while being ignorant of the deeper layers which bind it to the ground of all things … [T]*he theory implies that simple* [vortices] *are veritable subjects*. (*ibid*.: 418, emphasis added)

Having effectively returned to the idea of micro-level phenomenal ultimates, Shani then argues that macro-level relative subjects, such as you and me, *are formed by micro-level relative subjects coming together*. The rather technical core of his argument—which I shall not reproduce here, for it is not relevant to this paper—is that, by grounding the micro-level relative subjects in cosmic consciousness, he circumvents Coleman's (2014) attack on bottom-up panpsychism.

Even if the latter point is valid—and I have no reason to believe otherwise—I see multiple problems with this move. For one, once one starts from *cosmic* consciousness, it seems unnecessary and rather convoluted to descend all the way down to micro-level subjects, just to turn around again and go up to macro-level subjects. The only motivation I see for doing so is the arguably flawed notion, discussed earlier, that the 'pixels' discernible on the screen of perception must be the building blocks of the experien*cer*, as opposed to the *experience*. By making a concession to this physicalist intuition, Shani forces two problems upon himself: he has to explain (a) how the cosmic subject seemingly breaks up into myriad micro-level relative subjects, and then (b) how these micro-level relative subjects come together again to form macro-level relative subjects.

Moreover, recall that, as per Coleman's definition, subjects entail "a point of view annexed to a private qualitative field" (Coleman 2014: 30). So micro-level relative subjects must have *private* phenomenal fields *inaccessible to other subjects*. To tackle the decombination problem, one must explain how these private fields form within the ocean of cosmic consciousness. But Shani seems to address this only in a vague, tangential manner. For starters, it is unclear how or why a mere "localization process" in the ocean of cosmic consciousness would lead to local "intensification of experience" and "concentration of focus" (Shani 2015: 418). But even if we grant that it somehow does, a "concentration of focus within limited and relatively well-defined boundaries" does not seem sufficient "to separate the system's inner reality from the inner reality of the ocean surrounding it" (*ibid*.). By way of analogy, while my visual focus right now rests on the characters I am writing, I am not unaware of, or separate from, the contents of my peripheral vision; I still have direct—that is, experiential—

access to them. Moreover, it is hard to imagine that an electron could have sufficiently rich phenomenal properties to become "engulfed in its own experiences and concerns" (*ibid*.). It seems highly unlikely that there is enough cognitive complexity—if there can be cognition at all—at that microscopic level to justify such an appeal to mere self-absorption as the mechanism behind the separation of the electron's inner reality.

I do think Shani is on the correct general path here, but a more elaborate, explicit and precise case, with stronger *empirical* substantiation, seems to be necessary to tackle the decombination problem.

3.8 The key questions to be answered

The principle of parsimony implies that, of the two interpretations of cosmopsychism discussed above, idealist cosmopsychism is more economical and, therefore, should be preferred *if it can account for all relevant facts*. I shall thus take idealist cosmopsychism as my starting point and then attempt to address each of its problems and limitations. The goal is to account for all relevant facts with cosmic consciousness alone in the reduction base. Because I do not feel the need to invent new names for ideas that have historically established names, I shall call the resulting ontology simply *idealism*.

Specifically, here are the key problems of, and questions not sufficiently or explicitly addressed by, idealist cosmopsychism that I now set out to tackle:

a) *Grounding experience in cosmic consciousness*: How do myriad, ephemeral experiential qualities arise in one enduring cosmic consciousness?
b) *The decombination problem*: How do private phenomenal fields form within cosmic consciousness? Why can I not read your thoughts by simply shifting the focus of my attention?
c) *Reducing perception*: How can the revealed order of nature (the physical world we measure) be explained in terms of its concealed order (its underlying thoughts)? Why are the respective qualities so different?
d) *Explaining the correlations between brain function and inner experience*: If brain function does not constitute or generate phenomenality, why do they correlate so well?
e) *Explaining a seemingly shared, autonomous world*: If the world is imagined in consciousness, how can we all be imagining essentially the same world outside the control of our personal volition?

3.9 Experiences as excitations of cosmic consciousness

The first step is to clarify the relationship between cosmic consciousness and experience. After all, the two are not interchangeable: cosmic consciousness is, *ex hypothesi*, something relatively enduring and stable, whereas experiences are relatively ephemeral and dynamic. Yet, idealism posits that cosmic

consciousness is nature's *sole* ontological primitive, so how does the variety and dynamism of experience come into the picture?

I submit that (a) experiences are *patterns of self-excitation* of cosmic consciousness and that (b) cosmic consciousness has the inherent disposition to self-excitation. As such, experiences are not ontologically distinct from cosmic consciousness, just as a dance is not distinct from the dancer. There is nothing to a dance but the dancer in motion. In an analogous way, there is nothing to experience but cosmic consciousness 'in motion.'

Particular experiences correspond to *particular patterns* of self-excitation of cosmic consciousness, just as particular choreographies correspond to particular patterns of self-excitation of the dancer. These patterns can evolve in time and differ across different segments of cosmic consciousness. It is the variety and dynamics of excitations across the underlying 'medium' that lead to different experiential qualities. (One must be careful at this point: by referring to cosmic consciousness as a 'medium' I may appear to be objectifying it. Language forces me into this dilemma. But cosmic consciousness is subjectivity itself, not an object.) This way, even if the 'medium' is eternal and immutable, its self-excitations can come and go in myriad patterns.

This notion is entirely analogous to, and consistent with, how modern physics attempts to reduce the variety and dynamics of natural phenomena to an enduring primary substrate: quantum field theory, for instance, posits that all fundamental particles are particular modes of self-excitation of a quantum field, which is inherently disposed to self-excitation. Superstring theories posit essentially the same, but now the self-excited substrate is hyper-dimensional strings. Finally, according to M-theory the patterns of nature consist of modes of self-excitation of a hyper-dimensional membrane. Idealism, as I am formulating it here, essentially entails porting the evolving mathematical apparatus of modern physics to cosmic consciousness itself, as opposed to an abstract conceptual object. This should require but a straightforward and seamless transposition, implying no loss of predictive power.

3.10 Tackling the decombination problem

Cosmic consciousness comprises a variety of *phenomenal contents*—experiences, patterns of self-excitation—such as thoughts and feelings. If we take the human psyche as a representative sample of how cosmic consciousness operates—which is the best we can do, really—we can infer that, ordinarily, these phenomenal contents are *internally integrated* through *cognitive associations*: a feeling evokes an abstract idea, which triggers a memory, which inspires a thought, etc. These associations are *logical*, in the sense that, for instance, the memory inspires the thought because of a certain implicit logic linking the two. Ordinary phenomenal activity in cosmic consciousness can thus be modeled as a connected directed graph. See Figure 3.1a. Each vertex in the graph represents a particular phenomenal content and each edge a cognitive association logically linking contents together. Every phenomenal

content in the graph of Figure 3.1a can be reached from any other phenomenal content through a chain of cognitive associations.

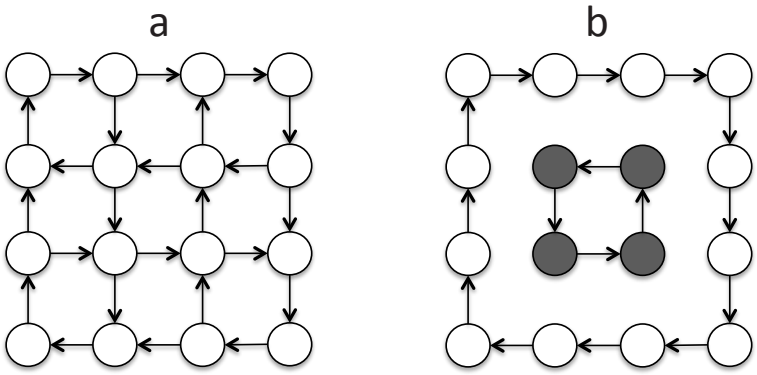

Figure 3.1: A connected graph (a) illustrating normal integration of phenomenal contents, and a disconnected graph (b) illustrating dissociation and the corresponding formation of an alter (inner subgraph in grey).

Each vertex in Figure 3.1 represents a particular pattern of self-excitation of cosmic consciousness. Each edge represents thus an association between two patterns of self-excitation, each pattern with its particular constituent harmonics. When the two patterns of self-excitation are concurrently present—that is, when the two associated phenomenal contents are experienced together—the association can be seen as a combination of the respective harmonics, like in a musical chord wherein multiple notes are played at the same time. When the association unfolds in temporal sequence—as e.g. in the case of a thought that fades away to make room for the experience of the memory it evokes—it can be visualized as a transition from the first to the second pattern of self-excitation, like notes played in sequence in a melody.

However, we know from the psychiatric literature that sometimes "a disruption of and/or discontinuity in the normal integration" of phenomenal contents can occur in the human psyche (Black and Grant 2014: 191). This is called *dissociation* and is well recognized in psychiatry today (American Psychiatric Association 2013). Dissociation entails that some phenomenal contents cease to be able to evoke others. A person suffering from a particularly severe form of dissociation, called Dissociative Identity Disorder (DID), exhibits multiple "discrete centers of self-awareness" (Braude 1995: 67) called *alters*. Each alter corresponds thus to a particular segment of the psychic space wherein it forms.

Dissociation can be visualized as what happens when the graph in Figure 3.1a becomes disconnected, such as shown in Figure 3.1b. Some phenomenal contents can then no longer be reached from others. The inner subgraph is thus a representation of an alter, corresponding to a particular segment of the originally integrated psychic space.

There is compelling empirical evidence that different alters can remain concurrently conscious. In Morton Prince's well-known study of the "Miss Beauchamp' case of DID, one of the alters "was a co- conscious personality in a deeper sense. When she was not interacting with the world, she did not become dormant, but persisted and was active" (Kelly *et al*. 2009: 318). Braude's more recent work (1995) corroborates the view that alters can be co-conscious. He points to the struggle of different alters for executive control of the body and the fact that alters "might intervene in the lives of others [i.e., other alters], intentionally interfering with their interests and activities, or at least playing mischief on them" (*ibid*.: 68). It thus appears that alters can not only be concurrently conscious, but that they can also vie for dominance with each other.

Clearly, the evidence indicates that different alters entail—to paraphrase Coleman (2014) again—different co-conscious points of view annexed to *private* qualitative fields, these private qualitative fields being carved out by virtue of dissociation. In other words, *different alters are different subjects*. The connected subgraph of phenomenal contents associated with an alter (see Figure 3.1b again) represents its private qualitative or phenomenal field. Moreover, alters form within a single overarching psyche, so the process of their formation entails a decomposition of an original subject into multiple lower-level subjects.

I submit that *dissociation in cosmic consciousness is what leads to the formation of relative subjects*. Each relative subject is thus an alter of cosmic consciousness, its private qualitative field corresponding to a segment of the latter's self-excitatory 'medium.'

By virtue of corresponding to a segment of cosmic consciousness, each alter retains—as Shani (2015) posited—the intrinsic features of *sentience* and *core-subjectivity*. But the local pattern of dissociative phenomenal activity in its respective segment is what bestows an alter its specific character, its unique perspective. In other words, the primary sense of I-ness of all alters is that of cosmic consciousness itself; the very consciousness of the alters, as an ontological 'medium,' is cosmic consciousness. But the particular phenomenal field of an alter, which defines its identity as a seemingly separate individual, is demarcated by a local dissociative process—analogous to DID—in the corresponding segment of the 'medium.' Naturally, because alters are fully grounded in cosmic consciousness, it is incoherent to say that they become separated from it; only an illusion of separation arises as a particular phenomenal content in the alter's dissociated qualitative field.

The key to my argument is the notion that dissociation can demarcate and carve out a private phenomenal field. This way, alters must become blind to all phenomenality taking place outside their respective field, which then explains why I cannot read your thoughts. And indeed, there is strong empirical evidence for the *literally blinding* power of dissociation: in 2015, doctors reported on the case of a German woman who exhibited a variety of alters (Strasburger and Waldvogel 2015). Peculiarly, some of her alters claimed to be

blind while others could see normally. Through EEGs, the doctors were able to ascertain that the brain activity normally associated with sight wasn't present while a blind alter was in control of the woman's body, even though her eyes were open. When a sighted alter assumed control, the usual brain activity returned. Clearly thus—if nothing else, for sheer empirical reasons—dissociation is a sufficiently powerful potential solution to the decombination problem.

3.11 At what level does cosmic dissociation occur?

The challenge we must now address is the so-called "boundary problem for experiencing subjects" (Rosenberg 2004: 77-90): What measurable structures in nature correspond to—that is, are the revealed appearance of—alters of cosmic consciousness? As we have seen, Shani (2015) posits that elementary particles are akin to micro-level alters, which in turn come together to compose higher-level relative subjects. However, as already mentioned, I believe this to be an unnecessarily convoluted notion. Instead, I submit that *cosmic dissociation happens precisely at the level of living beings with unitary consciousness, such as you and me*. You and I are alters of cosmic consciousness.

There are several arguments for this. The first has already been hinted at: given that we ordinarily experience an integrated phenomenal field, there is no direct reason to conjecture that this field is a composite of lower-level constituents.

Secondly, we have seen that von Neumann's reasoning regarding quantum measurement (1996) implies that the entire inanimate universe must be one unfathomable 'von Neumann chain'—that is, an entangled indivisible whole. As such, it is arbitrary—physically speaking—to carve out any segment of the inanimate universe and posit it to be the revealed appearance of an alter. Von Neumann did, however, exclude conscious living organisms from the embrace of von Neumann chains, since at least we, conscious human beings, clearly *can* perform quantum measurements. On this basis, only conscious living organisms can correspond to alters of cosmic consciousness, not elementary particles or any other subset of the inanimate universe.

Thirdly, as observed by Mathews, "the individuation of [inanimate] objects … is not consistently objectively determined … many of our individuations—of rocks and mountains, for instance—have basically nominal status" (2011: 144). Take what we call a 'car': though based on structural and functional reasoning that helps the business of transportation, its delineation is ultimately arbitrary. If one argues that, say, the spark plugs are integral to the car because without them the car cannot function, by the same token one would also have to include the fuel that makes its engine run, the environment air that allows combustion and cools the engine, the road gripped by the tires, the ground that sustains the road, the gravity that enables grip, and so on. The decision of where to stop is motivated by convenience. An analogous rationale applies to whether we distinguish the handle from the mug, the hood from the jacket, the river from the ocean, etc. This relative arbitrariness in the way we delineate

their boundaries renders inanimate objects problematic candidates for the revealed appearance of alters of cosmic consciousness. After all, in Mathews' words, "the boundaries between subjects are not nominal. The individuation of subjects, or centers of subjectivity, is objectively determined" (*ibid*.).

Mathews is giving us an important clue here. Indeed, the boundaries of our own body are not nominal. Our ability to *perceive* ends at the surface of the body: our skin, retinas, eardrums, tongue and the mucous lining of our nose. We cannot perceive photons hitting a wall or air pressure oscillations bouncing off a window, but we *can* perceive those impinging on our retinas and eardrums, respectively. Moreover, our ability to act through direct phenomenal intention also ends at the surface of the body: we can move our arms and legs simply by consciously *intending* to move them. However, we cannot do the same with tables and chairs. Clearly, thus, the delineation of our body is an *empirical fact*. I cannot just decide that the chair I am sitting on is integral to my body, in the way I *can* decide that the handle is integral to the mug. Neither can I decide that a patch of my skin is not integral to my body, in the way I *can* decide that the hood is not integral to the jacket. The criterion here is not merely a functional or structural one, but the range of phenomenality—sensory perception, intention—intrinsically associated with our body. Based on this objective criterion, there is no freedom to move boundaries at will.

What these considerations suggest is clear: *the physical boundary of the body is the revealed appearance of the dissociative boundary of our phenomenal field*. And in so far as we can assume that all living organisms have phenomenal inner life in some way akin to our own, the conclusion can be generalized: *living organisms are the revealed appearance of alters of universal consciousness;* they are carved out of their context by virtue of cosmic dissociation.

But *can* we assume that all living creatures have phenomenal inner life? I believe we can: insofar as it resembles our own, the extrinsic behavior of *all* metabolizing organisms is suggestive of their having dissociated phenomenal fields analogous to ours in some sense. This is obvious enough for cats and dogs, but what about plants and single-celled organisms such as amoebae? Well, consider this: "many types of amoeba construct glassy shells by picking up sand grains from the mud in which they live. The typical *Difflugia* shell, for example, is shaped like a vase, and has a remarkable symmetry" (Ford 2010: 26). As for plants, many recent studies have reported on their surprisingly sophisticated behavior, leading even to a proposal for a new field of scientific enquiry boldly called "plant neurobiology" (Brenner *et al*. 2006). Clearly, thus, even plants and single-celled organisms exhibit extrinsic behavior somewhat analogous to our own, further suggesting that they, too, have dissociated phenomenal fields. Of course, the same cannot be said of any inanimate object or phenomenon (those that have been engineered by humans to merely *simulate* the behavior of living beings, such as robots, natural language interfaces, etc., naturally don't count).

Finally, we have good empirical reasons to believe that normal metabolism is essential for the maintenance of our dissociated phenomenal fields, for when it

slows down or stops the dissociative boundary seems to become phenomenally porous (Kastrup 2017a[2]). So metabolism—the shared and differentiating characteristic of all living organisms—seems, again, to be the revealed appearance of alters of cosmic consciousness. The unique features of metabolism—think of DNA, morphogenesis, transcription, protein folding, mitosis, etc.—unify all life into a unique, clearly distinct natural category, despite the widely different forms organisms can take. This category provides the unambiguously demarcated "something in nature" that Rosenberg was looking for (2004: 77–90).

In conclusion, I posit that cosmic dissociation happens precisely at the level of individual living organisms. Each living organism is an alter of cosmic consciousness.

3.12 Reducing the revealed to the concealed order

Notice that the revealed side of nature relative to any given alter consists of images on the screen of the alter's conscious *perception*. Therefore, if one can reduce perceptions at the level of individual alters to non-perceptual phenomenal contents at the level of cosmic consciousness as a whole, one will have reduced nature's revealed to its concealed order.

Before we address this challenge, however, we need some further background on dissociation. By definition, phenomenal contents inside an alter cannot evoke phenomenal contents outside the alter, and vice versa. But they can still *influence* each other. Indeed, phenomenal impingement across a dissociative boundary is empirically known. John Lynch and Christopher Kilmartin (2013: 100), for instance, report that dissociated feelings can dramatically affect thoughts and corresponding behaviors, whereas David Eagleman (2011: 20-54) shows that dissociated expectations routinely mold our perceptions. Indeed, the entire clinical field of depth psychology is based on the notion that dissociated phenomenal contents in deeper layers of the psyche continuously impinge on the executive ego (Kelly *et al.* 2009: 301–34). We can visualize this as in Figure 3.2a, wherein the partial overlap of adjacent vertices internal and external to an alter represents impingement across its dissociative boundary.

Figure 3.2b illustrates the same thing according to a simplified representation unrelated to graph theory: the broader psychic space is represented as a white circle, with an alter represented as a grey circle within it. These circles are no longer graph vertices but represent sets of phenomenal contents. The dashed arrows represent the impingement of external and internal phenomenal contents—not explicitly shown—on each other, across the alter's boundary. For the avoidance of doubt, notice that these dashed arrows no longer represent cognitive associations. I shall use this simplified representation henceforth.

[2] See Chapter 6 of this dissertation.

Moreover, for simplicity's sake, from now on I shall also refer to all non-perceptual phenomenal contents simply as 'thoughts.'

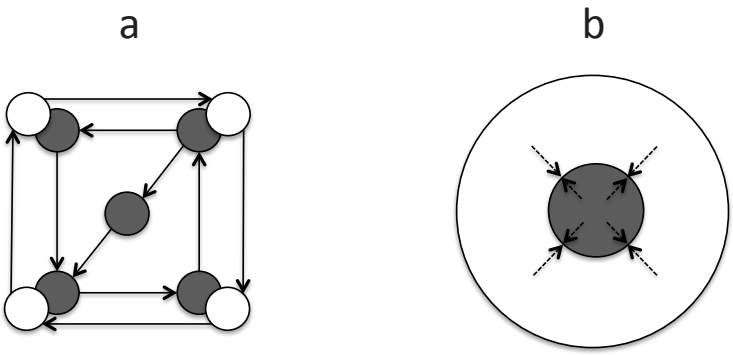

Figure 3.2: Phenomenal contents impinging on the dissociative boundary of an alter, illustrated in two different but equivalent ways, (a) and (b).

I submit that, before its first alter ever formed, the only phenomenal contents of cosmic consciousness were thoughts. There were no perceptions. The formation of the first alter then demarcated a boundary separating phenomenal contents *within* the alter from those *outside* the alter. This newly formed boundary is what enabled perceptions to arise relative to an alter: the thoughts surrounding the alter impinged on its dissociative boundary from the outside. And since phenomenal contents are particular patterns of self-excitation of cosmic consciousness, this impingement can be regarded as an *interference pattern* between excitations within and outside the dissociative boundary, respectively (see Figure 3.2a again). What we call perception, or the revealed side of nature, is the alter's experience of this interference pattern (*cf.* Kastrup 2017e). It follows that the revealed side of nature can be grounded in its concealed side: the former arises from excitatory interference between dissociated but mutually impinging *thoughts*. Indeed, I submit that the formation of dissociative boundaries is what partitioned the cosmos into revealed and concealed sides. See Figure 3.3.

The thoughts of an alter can also impinge on its dissociative boundary from the inside and thereby influence the surrounding phenomenal activity of cosmic consciousness (not shown in Figure 3.3). This corresponds to the effects on the world of the presence and actions of a living organism within it.

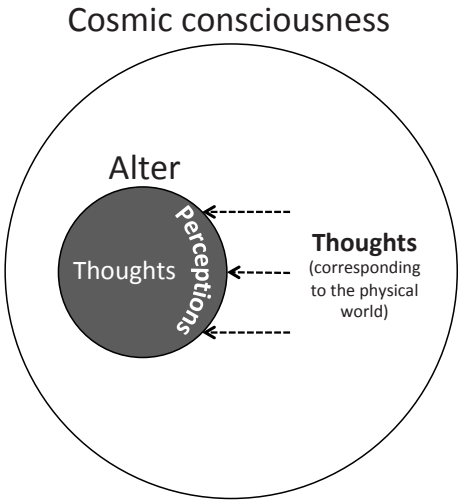

Figure 3.3: Thoughts in cosmic consciousness cause perceptions in an alter.

The revealed or extrinsic appearance of an alter's boundary is an organism's sense organs. In our case, these are our skin, eyes, ears, nose and tongue. Therefore, even if the outside stimulation is very faint and subtle, evolution has had billions of years to optimize the sensitivity of our sense organs—our alter's boundary—to pick up on these faint signals.

But how can a mere dissociative boundary give rise to a *qualitatively different* category of experience? If you recall, this is a question I raised earlier, motivated by the fact that perceptions *feel* undoubtedly very distinct from thoughts.

To answer it, let us first consider Donald Hoffman's interface theory of perception (2009): it asserts that evolution emphasizes perceptual qualities conducive to *fitness*, not to truth. In other words, we have evolved to perceive not the phenomenal contents that are really out there—that is, outside our alter—but just a phenomenal *representation* thereof that helps us survive and reproduce. Hoffman uses the analogy of a computer desktop: although a computer file is represented in it as, for instance, a blue rectangle, this does not mean that the file itself has the qualities of being blue and rectangular. As a matter of fact, the actual file does not have those qualities at all: it is a pattern of open and closed microscopic switches in a silicon chip. In an analogous way, my hypothesis is that the qualities we experience on the screen of perception—colors, sounds, flavors, textures, etc.—are *not* the qualities experienced by the segment of cosmic consciousness that surrounds our alter, but their 'desktop representation' instead. Our perceptions do not feel like the thoughts of cosmic consciousness because a *qualitative transition* between these two experiential categories has helped our ancestors survive and reproduce.

The work of Friston, Sengupta and Auletta (2014) has similar implications but, significantly, is derived from an entirely different line of reasoning. Their results are based on abstract mathematical considerations and, therefore, can in principle be leveraged under any ontology. They show that whenever a Markov blanket (Pearl 1988) defines the boundary of an individual organism, internal states of the organism will evolve to optimize for two conflicting goals: (a) to reflect external states of the world beyond the Markov blanket; and (b) to minimize their own entropy or dispersion. Goal (a) is about allowing the organism to know what is going on in the world outside, so it can take suitable actions to survive in that world. Goal (b) is about preventing the organism from losing its internal structural and dynamical integrity because of the second law of thermodynamics. In our case, the dissociative boundary of an alter is the Markov blanket, whose revealed appearance is our skin and other sense organs.

The key insight of Friston, Sengupta and Auletta can be paraphrased as follows: a hypothetical organism with perfect perception—that is, able to perfectly *mirror* the phenomenal states of the surrounding external world in its internal states—would not have an upper bound on its own internal entropy, which would then increase indefinitely. Such an organism would dissolve into an entropic soup. To survive, organisms must, instead, use their internal states to actively *represent* relevant states of the outside world in a *compressed, coded* form, so to know as much as possible about their environment while remaining within entropic constraints compatible with their structural and dynamical integrity. This way, my hypothesis is that the qualities of perception experienced by an alter are just compressed, coded representations of how surrounding thoughts of cosmic consciousness are experienced from the concealed perspective. As such, while there must be a *correspondence* between perception and surrounding thoughts, the respective experiential qualities do not need to be the same. In fact, they will be very different if it helps organisms resist entropy. Our perceptions do not feel like thoughts because they are coded representations thereof.

3.13 Explaining the correlations between brain function and inner experience

A principal argument for the mainstream physicalist position that the material brain somehow constitutes or generates consciousness is the empirically undeniable correlation between measurable brain function and inner experience (e.g. Koch 2004). The way the idealist ontology proposed here accommodates this fact was already implicit in the previous section: *a metabolizing body—which includes a functioning brain—is simply the revealed appearance of the dissociated phenomenal field of an alter*. The former correlates with the latter simply because the former is what the latter *looks like* from across a dissociative boundary. Indeed, this can be empirically substantiated in a rather direct manner.

In a 2014 study of dissociation (Schlumpf *et al.*), doctors performed functional brain scans on both DID patients and actors simulating DID. The scans of the actual patients displayed clear differences when compared to those of the actors, showing that *dissociation has an identifiable extrinsic appearance*. In other words, there is something rather particular that dissociative processes *look like*. This further substantiates the notion that living organisms such as you and me are the revealed appearance of cosmic-level dissociative processes. After all, we now know empirically that dissociation is identifiable when observed from across the dissociative boundary. Metabolizing bodies are to dissociation in cosmic consciousness as certain patterns of brain activity are to DID patients.

Let me elaborate further on this important point. For any given alter A_1 of cosmic consciousness, it is the phenomenal contents surrounding A_1 that cause its perceptions of the world around it. Dissociated phenomenal contents corresponding to another alter A_2 can be part of the phenomenal environment surrounding A_1. As such, the inner experiences of A_2 can also indirectly stimulate A_1's boundary—by impinging on their shared phenomenal environment—and thereby cause A_1's perceptions of A_2. This is what gives A_1 access to the revealed appearance of the inner experiences of A_2 in the form of A_2's metabolizing body. See Figure 3.4. And since A_2's brain is integral to its body, it follows that A_2's inner experiences cause the perception by A_1 of the activity in A_2's brain. This causal link explains the correlations between inner experience and corresponding patterns of brain activity.

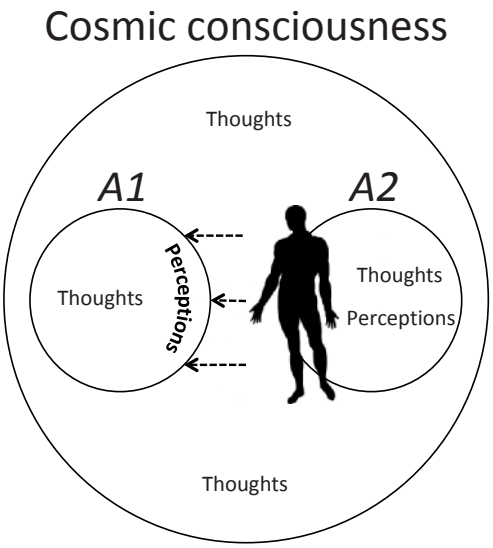

Figure 3.4: A metabolizing body is the revealed appearance of an alter's dissociated phenomenal field.

In essence, the claim here is that there is nothing to a metabolizing body but the revealed side—the extrinsic appearance—of the corresponding alter's inner experiences. Yet, one may object to this by arguing that many parts of the body seem entirely unrelated to inner experience: whereas certain patterns of brain activity correlate with subjective reports of experience, a lot seems to go on in the brain that subjects have no introspective access to (Westen 1999, Hassin, Ulleman and Bargh 2005, Dijksterhuis and Nordgren 2006, Augusto 2010, Hassin 2013). Moreover, what kind of inner experience does, say, liver function correspond to? What about big-toe function?

The answer to this objection is precise and compelling, but elaborate and specialized enough to have required its own paper (Kastrup 2017d[3]). Here, I shall simply remind the reader that a subject's lack of *metacognitive access* to an experience precludes reporting of the experience to self or others, but does not imply *absence* of the experience from the subject's qualitative field. With the emergence of no-report paradigms in neuroscience (Vandenbroucke *et al.* 2014, Tsuchiya *et al.* 2015), we now know that much is experienced that cannot be reported even to self, for subjects are often not aware *that* they have certain experiences. Moreover, as mentioned earlier, there are normal internal dissociations in the human psyche—the foundational claim of depth psychology—that render much of its phenomenal contents inaccessible to the reporting ego (Kelly *et al.* 2009: 301-34). So the hypothesis I am positing here is not defeated by the objection: all bodily metabolism—yes, even liver and toe function—can still correspond to concealed phenomenal contents, even though these contents may not be introspectively accessible.

3.14 Explaining our shared world

The final explanatory burden that needs to be addressed is the undeniable empirical fact that we all inhabit seemingly the same environment, and that the laws that govern the dynamics of this environment operate independently of our personal volition. After all, if the world is imagined—as implied by idealism—how come we are all imagining seemingly the same autonomous world?

Notice that the existence of a phenomenal environment wherein all metabolizing organisms are immersed—a shared world—is a direct implication of the argument already developed. To bring this out, we simply need to extend Figure 3.3 to multiple alters, as illustrated in Figure 3.5. All alters are immersed, like islands of a single ocean, in the thoughts that constitute the concealed side of the inanimate cosmos. These thoughts surround all alters and cause their mutually-consistent perceptions by impinging on their respective dissociative boundaries. And since the volition of an alter is a phenomenal content also dissociated from the rest of cosmic consciousness, it follows that alters cannot

[3] See Chapter 5 of this dissertation.

change the laws of nature. From the dissociated perspective of alters, the world is thus autonomous.

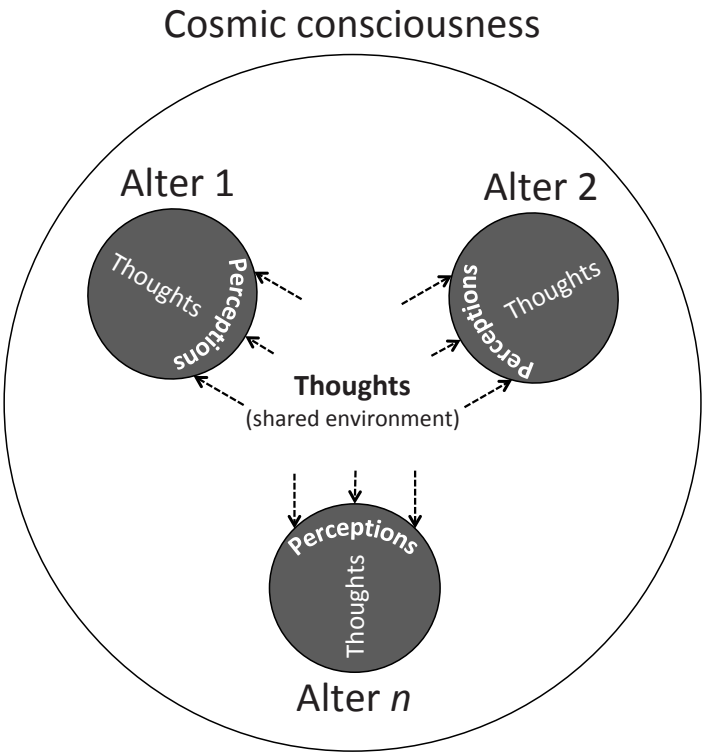

Figure 3.5: Alters are immersed in a common phenomenal environment.

3.15 Conclusions

I have elaborated on an idealist ontology that can be summarized as follows. There is only cosmic consciousness. We, as well as all other living organisms, are but dissociated alters of cosmic consciousness, surrounded by its thoughts. The inanimate world we see around us is the revealed appearance of these thoughts. The living organisms we share the world with are the revealed appearances of other dissociated alters. This idealist ontology makes sense of reality in a more parsimonious and empirically rigorous manner than mainstream physicalism, bottom-up panpsychism and cosmopsychism. It also offers more explanatory power than these three alternatives, in that it does not fall prey to the hard problem of consciousness, the combination problem or the decombination problem, respectively.

4. On the Plausibility of Idealism: Refuting Criticisms

This paper first appeared in *Disputatio: International Journal of Philosophy*, ISSN: 0873-626X, Vol. 9, No. 44, pp. 13-34, in May 2017.

4.1 Abstract

Several alternatives vie today for recognition as the most plausible ontology, from physicalism to panpsychism. By and large, these ontologies entail that physical structures circumscribe consciousness by bearing phenomenal properties within their physical boundaries. The ontology of idealism, on the other hand, entails that all physical structures are *circumscribed by* consciousness in that they exist *solely as* phenomenality in the first place. Unlike the other alternatives, however, idealism is often considered implausible today, particularly by analytic philosophers. A reason for this is the strong intuition that an objective world transcending phenomenality is a self-evident fact. Other arguments—such as the dependency of phenomenal experience on brain function, the evidence for the existence of the universe before the origin of conscious life, etc.—are also often cited. In this essay, I will argue that these objections against the plausibility of idealism are false. As such, this essay seeks to show that idealism is an entirely plausible ontology.

4.2 Introduction

The mainstream physicalist ontology posits that reality is constituted by irreducible physical entities—which Strawson (2006: 9) has called "ultimates"—outside and independent of phenomenality. According to physicalism, these ultimates, in and of themselves, do not instantiate phenomenal properties. In other words, there is nothing it is like to *be* an ultimate, phenomenality somehow emerging only at the level of complex arrangements of ultimates. As such, under physicalism phenomenality is not fundamental, but instead reducible to physical parameters of arrangements of ultimates.

What I will call 'microexperientialism,' in turn, posits that there is already something it is like to be at least some ultimates (Strawson *et al*. 2006: 24-29), combinations of these experiencing ultimates somehow leading to *more complex* experience. As such, under microexperientialism phenomenality is seen as an irreducible aspect of at least some ultimates. The ontology of *pan*experientialism (Griffin 1998: 77-116, Rosenberg 2004: 91-103, Skrbina 2007: 21-22) is analogous to microexperientialism, except in that the former entails the stronger claim that *all* ultimates instantiate phenomenal properties.

Micropsychism (Strawson *et al*. 2006: 24-29) and panpsychism (Skrbina 2007: 15-22) are analogous—maybe even identical—to microexperientialism and panexperientialism, respectively, except perhaps in that some formulations of

the former admit cognition—a more complex form of phenomenality—already at the level of ultimates, as an irreducible aspect of these ultimates.

While microexperientialism, panexperientialism, micropsychism and panpsychism entail that bottom-up combinations of simple subjects give rise to more complex ones, such as human beings, cosmopsychism (Nagasawa and Wager 2016) takes the opposite route. Indeed, "the first postulate of cosmopsychism is that *the cosmos as a whole is the only ontological ultimate* there is, and that it is *conscious*" (Shani 2015: 408, original emphasis).

Finally, the ontology of idealism is characterized by a combination of two propositions: (a) phenomenal consciousness is irreducible; *and* (b) everything else—the whole of nature—is reducible to a unitary and universal phenomenal consciousness (henceforth, I shall refer to phenomenal consciousness simply as 'consciousness').

Idealism may be consistent with—even identical to—certain interpretations of cosmopsychism. According to Shani, for instance, cosmopsychism entails that "an omnipresent cosmic consciousness is the single ontological ultimate there is" (2015: 390). This perfectly embodies the defining tenet of idealism insofar as it implies that everything—*including the physical*—can be reduced to the phenomenal. Shani also writes that matter is the cosmos "in its *appearance* as exterior complement to the *subjective* realities of created selves" (2015: 412, emphasis added). The notion that matter is the phenomenal appearance of equally phenomenal dynamics is also eminently idealist. Therefore, these interpretations of cosmopsychism are essentially indistinguishable from idealism and I shall, henceforth, refer to them simply as idealism.

Other possible interpretations of cosmopsychism entail that the cosmos as a whole bears phenomenal properties—that is, has inner life—but also has an aspect—the physical universe we can measure—that is irreducible to these phenomenal properties. Naturally, this implies a form of dual-aspect monism, *a la* Spinoza (Skrbina 2007: 88). Indeed, under these views the cosmos can still be said to be *conscious*, but not *in consciousness*. In the former case, the cosmos *bears* phenomenality; in the latter—which is the idealist view—the cosmos is *constituted by* phenomenality. Interpretations of cosmopsychism that are not consistent with idealism shall not be further addressed in this paper.

In what follows, I will attempt to rebut the most common objections to the plausibility of idealism. I will seek to show that these objections are based on circular reasoning, conflation, unexamined assumptions and several other misconceptions.

4.3 The felt concreteness objection

English writer Samuel Johnson is said to have argued against Bishop Berkeley's idealism by kicking a large stone while exclaiming: "I refute it *thus*!" (Boswell 1820: 218). Johnson was clearly appealing to the felt concreteness of the stone to suggest that it could not be just a figment of imagination. Indeed, the felt

concreteness of the world is probably the main reason why people intuitively reject the notion that reality unfolds in consciousness. If a truck hits you, you will hurt, even if you are an idealist.

However, notice that appeals to concreteness, solidity, palpability and any other quality that we have come to associate with things outside consciousness are still appeals to phenomenality. After all, concreteness, solidity and palpability are *qualities of experience*. What else? A stone allegedly outside consciousness, in and by itself, is entirely abstract and has no qualities. If anything, by pointing to the *felt* concreteness of the stone Johnson was implicitly suggesting the primacy of experience over abstraction, which is eminently idealist.

We have come to automatically *interpret* the felt concreteness of the world as evidence that the world is outside consciousness. But this is an unexamined artifact of subliminal thought-models. Our only access to the world is through sense perception, which is itself phenomenal. The notion that there is a world outside and independent of the phenomenal is an *explanatory model*, not an empirical fact. No phenomenal quality can be construed as direct evidence for something outside phenomenality.

4.4 The private minds objection

As discussed in Section 4.2, under idealism there is only one universal consciousness. Yet, at a personal level, our mental lives are clearly separate from one another. I do not have direct access to your thoughts and feelings and, presumably, neither do you to mine. Moreover, I do not seem to be aware of what is happening across the galaxy and, presumably, neither are you. So, if all reality is reducible to one universal consciousness, how can there be separate private minds such as yours and mine?

To make sense of this under idealism, we need to review a mental condition called *dissociation* (Braude 1995, Kelly *et al*. 2009: 167-174 & 348-352, Schlumpf *et al*. 2014, Strasburger and Waldvogel 2015). Indeed, it is now well established in psychiatry that mental contents can undergo "a disruption of and/or discontinuity in [their] normal integration" (Black and Grant 2014: 191). This normal integration of mental contents takes place through chains of cognitive associations: a perception may evoke an abstract idea, which may trigger a memory, which may inspire a thought, etc. These associations are *logical*, in the sense that e.g. the memory inspires the thought because of a certain *implicit logic* linking the two. Integrated mentation can thus be modeled, for ease of visualization, as a connected directed graph. See Figure 4.1a. Each vertex in the graph represents a particular mental content and each edge a cognitive association logically linking mental contents together. Every mental content in the graph of Figure 4.1a can be reached from any other mental content through a chain of cognitive associations. Dissociation, in turn, can be visualized as what happens when the graph becomes disconnected, such as shown in Figure 4.1b. Some mental contents can then no longer be reached from others.

Following the psychiatric convention, I shall refer to the subgraph with grey vertices as a (dissociated) *alter*.

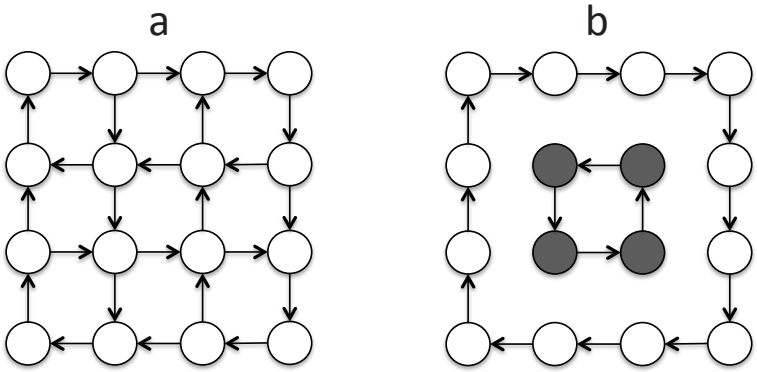

Figure 4.1: A connected graph (a) illustrating normal integration of mental contents, and a disconnected graph (b) illustrating dissociation and the corresponding formation of an alter (subgraph in grey).

Because cognitive associations are essentially logical, as opposed to spatio-temporal, the scheme of representation in Figure 4.1 allows for the *simultaneous* experience of multiple mental contents linked together in a connected subgraph. This is empirically justifiable: a perception, for instance, can be experienced *at the same time* as the thoughts it evokes and the emotions evoked by these thoughts. Moreover—and by the same token—the two disconnected subgraphs in Figure 4.1b can also represent two *concurrently conscious* subjects of experience. The substantiation for this is again empirical: there is compelling evidence that different alters of the same psyche can be co-conscious (Kelly *et al*. 2009: 317-322, Braude 1995: 67-68).

An alter loses direct access to mental contents surrounding it, *but remains integral to the underlying consciousness that constitutes it*. The disconnection between an alter and surrounding mental contents is logical, not ontic. As an analogy, a database may contain entries that are not indexed and, therefore, cannot be reached, but this does not physically separate those entries from the rest of the database.

Dissociation can coherently explain how seemingly separate but concurrently conscious subjects of experience—such as you and me—can form under idealism: each is an alter of universal consciousness. And because each alter becomes unable to evoke the mental contents of another, their respective inner lives acquire a seemingly private character, even though they remain integral to the underlying consciousness that constitutes them.

4.5 The stand-alone world objection

If all there is is consciousness, does the world continue to exist when not consciously observed by a living being? A negative answer to this question seems extremely implausible yet difficult to avoid under idealism. Bishop Berkeley has famously attempted to circumvent it by appealing to a divinity, as captured in Ronald Knox's limerick, *God in the Quad*:

There was a young man who said "God
Must find it exceedingly odd
To think that the tree
Should continue to be
When there's no one about in the quad."

Reply:

"Dear Sir: Your astonishment's odd;
I am always about in the quad.
And that's why the tree
Will continue to be
Since observed by, Yours faithfully, God."

Legitimate as an appeal to a divinity might have been in Berkeley's time, today more rigor is expected from a viable ontology. So how do we solve the problem of a stand-alone world under idealism?

With reference to the discussion in the preceding section, notice that, by definition, mental contents inside an alter of universal consciousness cannot directly evoke mental contents outside the alter, or vice-versa. But they can still *influence* or *impinge on* each other. Indeed, mental impingement across a dissociative boundary is empirically known. Lynch and Kilmartin (2013: 100), for instance, report that dissociated feelings can dramatically affect our thoughts, while Eagleman (2011: 20-54) shows that dissociated expectations routinely mold our perceptions. We can visualize this as in Figure 4.2a, wherein the partial overlap of adjacent vertices internal and external to the alter (*cf.* Figure 4.1b) represents mental impingement across its dissociative boundary.

Figure 4.2b illustrates the exact same thing according to a simplified representation: the broader consciousness is represented as a white circle with an alter represented as a grey circle within it. The dashed arrows represent the impingement of external and internal mental contents on each other, across the alter's boundary. I will henceforth use this simplified representation.

Now notice that mental contents of universal consciousness that surround—*but remain external to*—an alter can impinge on the alter's boundary from the outside. Under idealism, it can be coherently argued that this is what gives rise to sense perceptions: the physical world around us is the extrinsic *appearance* on the screen of perception of phenomenality surrounding our respective alter. See Figure 4.3.

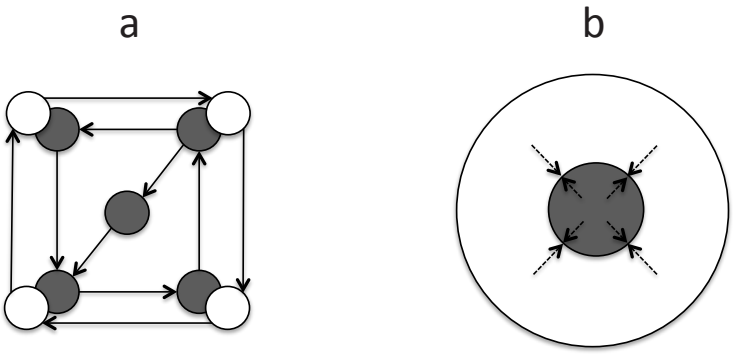

Figure 4.2: Mental contents impinging on the dissociative boundary of an alter, illustrated in two different ways (a) and (b).

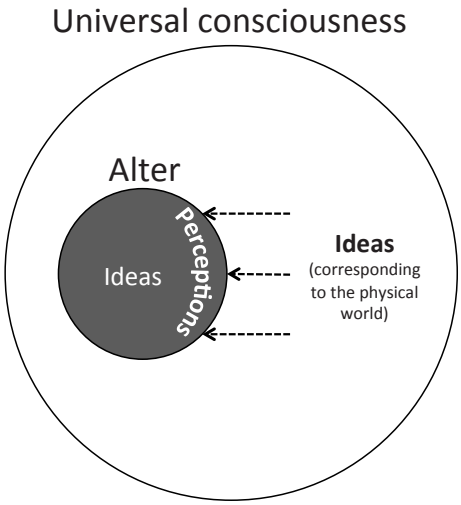

Figure 4.3: Mental contents of universal consciousness surrounding an alter can cause the alter's sense perceptions by impinging on its dissociative boundary.

The stand-alone character of the world can thus be coherently explained: the world is a perceptual representation of phenomenality dissociated from our personal psyche and, as such, independent of our personal inner life. That which underlies the physical world we perceive continues to exist—in the form of phenomenality outside our respective alter—even as we sleep.

4.6 The autonomy of nature objection

A closely related objection is this: nature unfolds according to patterns and regularities—the 'laws of nature'—independent of our personal volition. Human beings cannot change these laws. But if nature is in consciousness, should that not be possible by a mere act of imagination?

This objection can be rebutted along the same lines as the previous one. However, there is a more direct and intuitive refutation. Notice that the implicit assumption here is that all mental activity is acquiescent to volition, which is patently false even in our own personal psyche. After all, by and large we cannot control our dreams, nightmares, emotions and even many of our thoughts. They come, develop and go on their own terms. At a pathological level, schizophrenics cannot control their visions and people suffering from obsessive-compulsive disorder are constantly at the mercy of oppressive thoughts. There are numerous examples of conscious activity that escapes the control of volition. Often, we do not even recognize this activity as our own; that is, we do not identify with it. It unfolds as autonomous, seemingly external phenomena, such as dreams and schizophrenic hallucinations. Yet, all this activity is unquestionably within consciousness. We perceive it as separate from ourselves because the segment of our psyche that gives rise to this activity is dissociated from the ego, the segment with which we do identify.

So that there is activity in universal consciousness that we do not identify with and cannot control is entirely consistent with idealism. This activity is simply dissociated from our ego and its sense of volition.

4.7 The shared world objection

If all reality is in consciousness, then the world is akin to a dream. As such, idealism implies that we are all partaking in roughly the same dream. Yet, since our bodies are separate, we cannot be sharing a dream; or so the objection goes.

The objection begs the question by implicitly assuming that the body circumscribes dreaming consciousness, as opposed to the other way around. Only under this assumption does the impossibility of sharing a dream follow from the fact that bodies are separate. But under idealism, it is the body that is in universal consciousness, not consciousness in the body. Once this is properly understood according to the framework developed in the preceding sections, the rebuttal of this objection becomes rather straightforward: we all seem to inhabit the same world because our respective alters are surrounded by the same universal field of phenomenality, like whirlpools in a single stream. See Figure 4.4, which simply extends Figure 4.3 to multiple alters.

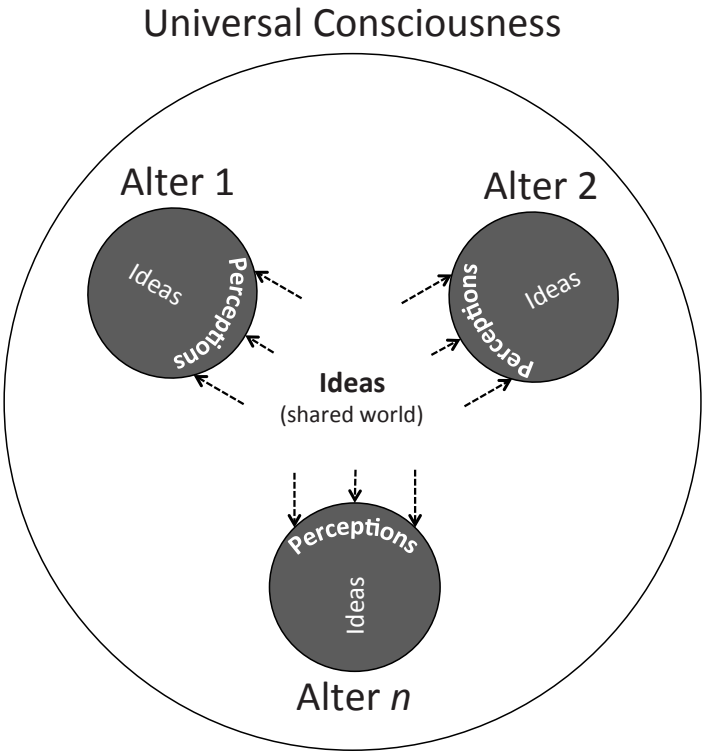

Figure 4.4: Alters of universal consciousness and their shared world.

4.8 The natural order objection

The world we perceive around ourselves is governed by stable and orderly natural laws. Therefore, if the contents of perception are a representation of phenomenality in universal consciousness, then this phenomenality must be stable and orderly at root. But our own personal thoughts and emotions are notoriously unstable and disorderly. So how plausible is it that the order and stability we discern in the laws of nature represent thoughts or emotions in universal consciousness?

The misconception here, of course, is that of anthropomorphization: to attribute to universal consciousness as a whole cognitive characteristics known only in small dissociated segments of it, such as human beings. Nothing in idealism precludes the possibility that phenomenality in universal consciousness unfolds according to very stable and orderly patterns and regularities, whose extrinsic appearance corresponds to the laws of nature. That our human thoughts and emotions seem rather reactive and unstable is a product of evolution under the pressures of natural selection within a particular planetary ecosystem. At a universal level, consciousness has not undergone such evolutionary pressures.

Under physicalism, the laws of nature are seen as irreducible causal patterns *somehow* built into the fabric of the cosmos. It is the dynamic unfolding of these patterns that leads to the order and complexity we see around ourselves. Under idealism, such irreducible causal patterns are posited to be *somehow* built into universal consciousness itself, instead of an objective fabric of spacetime. Yet, beyond this distinction, *they are the same patterns* that physicalism entails, as inherent to consciousness as physical laws are allegedly inherent to the fabric of spacetime. Idealism poses no extra difficulty than physicalism in this regard.

This can be better understood with a simple terminology move. Certain schools of psychology speak of 'psychological archetypes': innate, built-in templates according to which mental dynamics unfold (Jung 1991). As such, we can say that, under idealism, the laws of nature are the *archetypes of universal consciousness*. They are built-in templates according to which the 'vibrations' of universal consciousness—that is, phenomenality—develop, analogously to how the physical constraints of a vibrating surface determine its natural modes of vibration.

4.9 The equivalence objection

As we have seen in Sections 4.5 to 4.7, idealism acknowledges that there is a world outside *personal* psyches, since personal psyches are but dissociated segments of a broader universal consciousness. The objection, then, is that the notion of a broad stream of phenomenality outside personal psyches is equivalent to the physicalist postulate of a world outside consciousness.

Except for solipsism, any viable ontology must entail at least one inference beyond direct experience. This is necessary to make sense of the fact that we all inhabit the same world beyond ourselves and are unable to change its governing laws. For this reason, physicalism infers the existence of a universe outside consciousness, which we all inhabit. Idealism, on the other hand, infers simply that consciousness itself extends beyond its face-value personal boundaries. This way, while physicalism postulates a fundamentally new ontological class next to experience, idealism simply *extrapolates* the boundaries of consciousness—the sole undeniable ontological class and primary datum of existence—beyond those we can probe directly. To put it metaphorically, while idealism makes sense of reality by inferring that the Earth extends beyond the visible horizon, physicalism does so by inferring the existence of an isomorphic but ontologically distinct 'shadow' Earth. Clearly, the former is a more parsimonious inference and, as such, not equivalent to the latter.

More importantly, the implications of idealism are radically different from those of physicalism. For instance, while physicalism implies that consciousness ends upon the death of the body, idealism implies merely the end of the corresponding *dissociation*, not of consciousness proper. I have

elaborated on further differences in implications elsewhere (Kastrup 2015: 185-198).

4.10 The primacy of brain function objection

Not only are there (a) clear correlations between specific patterns of brain activity and reported inner experience (Koch 2004), we know that (b) physical interference with the brain—such as head trauma and the use of psychoactive drugs—can influence one's inner life rather dramatically. This may seem to suggest an arrow of causation pointing from a physical body outside consciousness to phenomenality, which would contradict idealism.

To make sense of observation (a), we need to briefly recapitulate earlier discussions. As we have seen in Section 4.4, under idealism private minds—such as our own human psyche—can be explained as dissociated alters of universal consciousness. We have also seen in Section 4.5 that the standalone world around us can be explained as the extrinsic appearance of phenomenality surrounding—but outside—our respective alter. Now, from the point of view of a given alter *A*, nothing prevents the dissociated mental activity of an alter *B* from being part of the phenomenality surrounding *A*. *B* is then part of *A*'s world and, as such, must also have an extrinsic appearance on *A*'s screen of perception. In other words, there must be something alters *look like* from a second-person point of view. And since we know from direct experience that our private inner life extends only to the boundaries of our metabolizing body—after all, we cannot perceive things that do not impinge on our skin or other sense organs, or move anything beyond our own body through direct intention—metabolizing bodies seem *prima facie* to be the extrinsic appearance of dissociated alters of universal consciousness. If so, this means that all living beings have private inner lives in some way analogous to our own, but tables and chairs do not. The latter are simply aspects of the *inanimate* universe, which, *as a whole*, is the extrinsic appearance of phenomenality *outside all alters*.

Brain activity, of course, is integral to a metabolizing human body. Therefore, under idealism, brain activity is simply part of what one's private inner experiences—self-reflective and otherwise, as I will elaborate upon in the next section—*look like* from across a dissociative boundary. To put it another way, one's brain activity is part of a phenomenal representation of one's inner life. And of course, a representation *must* correlate with the phenomenal process it is the appearance of, without requiring anything ontologically distinct from consciousness. That this correlation is empirically observed is thus entirely consistent with idealism.

A possible counterargument here is this: the patterns of neural activity one can measure with functional brain scanners can be enormously complex in terms of information content; perhaps more complex than the contents of consciousness we have introspective access to. What does the extra complexity then correspond to? The key to answering this question is in the next section,

wherein a distinction will be made between contents of consciousness we have introspective access to—that is, can self-reflect upon—and contents of consciousness that, despite still being experienced, fall outside the reach of introspection. The extra complexity, insofar as it indeed is the case, corresponds to the latter.

Regarding observation (b) of the objection, the suggested arrow of causation is based on an unexamined but pervasive assumption: that the physical is in some sense distinct from, yet causally effective upon, the phenomenal. This is precisely what idealism denies. Under idealism, the physical is simply the contents of perception, a particular type of phenomenality. As such, what we call 'physical interference with the brain' is the extrinsic appearance of phenomenality external to an alter that disrupts the inner experiences of the alter from across its dissociative boundary. The disruption 'pierces through' the boundary, so to speak. And that certain types of phenomenality disrupt other types of phenomenality is not only entailed by idealism, but also empirically trivial. After all, our thoughts disrupt our emotions—and vice-versa—every day. For the same reason that thoughts disrupt emotions, 'physical interference with the brain' disrupts an organism's inner life. None of this contradicts idealism.

4.11 The unconscious mentation objection

In Libet's now famous experiments (1985), neuroscientists were able to record, a fraction of a second *before* subjects reported making a decision to act, mounting brain activity associated with the initiation of a simple voluntary action. At first sight, this would seem to indicate that decisions are made in a neural substrate outside consciousness, thereby contradicting idealism. I use Libet's experiments here merely as an example, for today we know of many other instances of seemingly unconscious mentation, such as moving one's foot halfway to the brake pedal before one becomes aware of danger ahead (Eagleman 2011: 5). Under idealism, since everything is in consciousness, there cannot be such a thing as unconscious mentation. So what is going on?

The misconception here is a conflation of consciousness proper with a particular *configuration of* consciousness. Indeed, to report an experience—such as making a decision to act or seeing danger ahead—to another or to oneself, one has to *both* (a) have the experience *and* (b) know *that* one has the experience, which Schooler (2002) called a "re-representation." In other words, one can only report phenomenality that one is self-reflectively aware of at a metacognitive level. But self-reflection is just a particular configuration of consciousness, whereby consciousness turns in upon itself to experience knowledge of its own phenomenality (Kastrup 2014: 104-110). Nothing precludes the possibility that phenomenality takes place outside the field of self-

reflection. In this case, we cannot report the phenomenality—not even to ourselves—because we do not know *that* we experience it.[1]

The argument above is not idiosyncratic, for the existence of unreportable phenomenality is well established in neuroscience today (Tsuchiya *et al.* 2015, Vandenbroucke *et al.* 2014). Indeed, as elaborated upon by Schooler (2002), reportability is an *extra* function at a metacognitive level, on top of phenomenality proper. So the possibility that presents itself to us is that *all* mentation is actually conscious, even though we cannot report much of it. As such, the decisions made by Libet's subjects could well have been made in consciousness, but outside the field of self-reflection. The corresponding phenomenality then entered this field a fraction of a second later, thereby becoming reportable. Analogously, drivers may consciously see danger ahead before they can tell themselves *that* they see danger ahead. The *appearance* of unconscious mentation due to unreportability does not contradict idealism.

4.12 The unconsciousness objection

Along similar lines, the idea here is that, when we e.g. faint or undergo general anesthesia, we become seemingly unconscious. Yet, we do not cease to exist because of it, which may seem to contradict the idealist tenet that our body is the extrinsic appearance of conscious inner life.

Let us consider this more carefully. Imagine that you wake up in the morning after hours of deep sleep. You may remember nothing of what happened during those preceding hours, concluding that you were unconscious all night. Then, later in the day, you suddenly remember that you actually had a very intense dream. So you were not unconscious all night, you simply could not *remember* your experiences.

Indeed, all we can assert with confidence upon coming round from episodes of *seeming* unconsciousness is that we cannot *remember* phenomenality occurring during those episodes. The actual *absence* of phenomenality is impossible to assert with confidence. As a matter of fact, many things we have traditionally associated with unconsciousness are now known to entail intense experiences. For instance, fainting caused by e.g. asphyxiation, strangulation or hyperventilation is known to correlate with euphoria, insights and visions (Neal 2008: 310-315, Rhinewine and Williams 2007, Retz 2007). G-force-induced loss of consciousness (G-LOC) is also known to correlate with "memorable dreams" (Whinnery and Whinnery 1990). There is even evidence for "implicit perception" during general anesthesia (Kihlstrom and Cork 2007).

Sleep, of course, is known to correlate with dreams. But even during phases of sleep wherein electroencephalogram readings show no dream-related neural activity, there are other types of activity that may correlate with non-recallable phenomenality distinct from dreams. Indeed, this is precisely what a recent

[1] This is elaborated upon in much more detail in Chapter 5 of this dissertation.

study points out: "there are good empirical and theoretical reasons for saying that a range of different types of sleep experience, some of which are distinct from dreaming, can occur *in all stages of sleep*" (Windt, Nielsen and Thompson 2016: 871, emphasis added). The authors identify three different categories of sleep experiences distinct from dreams: (a) non-immersive imagery and sleep thinking, (b) perceptions and bodily sensations, and (c) 'selfless' states and contentless sleep experiences that may be similar to those reported by experienced meditators.

As such, what the empirical data shows is that episodes of seeming unconsciousness are associated with an impairment of memory access, but not necessarily with absence of phenomenality. As a matter of fact, there are strong indications, as mentioned above, that the opposite is true.

4.13 The solipsism objection

Some conflate idealism with solipsism, the notion that the world is one's *personal* dream, all other living creatures being just figments of one's personal imagination. Under solipsism, there is nothing it is like to be other people; they have no inner life; they exist only as appearances in the personal psyche of the dreamer. As such, whatever empirical evidence one brings to bear and whatever one says to a solipsist must be regarded by the solipsist as figments of his or her own imagination, which renders solipsism unfalsifiable. So the objection here is that, by being unfalsifiable, solipsism—and therefore idealism—is beneath philosophical debate.

Naturally, idealism is not solipsism. Under idealism, there *is* something it is like to be other living creatures; they also have private inner lives. So idealists take other people seriously as legitimate sources of reported experiences and views, not just as figments of one's own imagination. Moreover, idealists acknowledge that *there is a world outside and independent of their personal (dissociated) psyche*, as discussed in Sections 4.5 to 4.7. They simply do not acknowledge that this world is ontologically distinct from consciousness itself. Indeed, by acknowledging that dissociation in universal consciousness implies a world outside their own personal mentation, idealists look upon this world in a way entirely compatible with naturalism and scientific inquiry.

Unlike solipsism, idealism has the burden to explain observations non-trivially. Consider three basic facts that are often used to justify physicalism: (a) the laws of nature are independent of our personal volition; (b) we all seem to inhabit the same world; and (c) there are tight correlations between observable brain activity and reported inner life. Solipsism trivializes all three facts *in lieu* of actually making sense of them: the solipsist allegedly dreams them all up, rather arbitrarily. The idealist, on the other hand, by acknowledging the inner lives of other people and the autonomous nature of the world, has the burden to reconcile these three facts with the notion that reality unfolds in consciousness. If idealism is correct, (a) how come we cannot simply imagine a different and better world? If the world is akin to a dream in consciousness, (b)

how come we are all having the same dream? If consciousness is not generated by the brain, (c) how come there are such tight correlations between brain activity and inner experience? These questions have already been answered in Sections 4.6, 4.7 and 4.10, respectively. The important point here is this: idealism is falsifiable in that, if it cannot answer these and other questions in terms of universal consciousness alone, it must be discarded.

4.14 The cosmological history objection

There is overwhelming evidence for the existence of the universe before conscious life arose. Therefore—or so the objection goes—it is untenable to say that the universe exists in consciousness. This may strike some readers as obviously question-begging—which, of course, it is—but please bear with me for the sake of completeness.

The implicit assumption here is that consciousness arises only with biology, as a product of biology. Naturally, this is precisely what idealism denies. Under idealism, biology is merely the extrinsic appearance of *dissociated, local differentiations of consciousness* (that is, alters), not the constituent or generator of consciousness. There was universal consciousness before such dissociated, local differentiations arose. And there was phenomenality in this universal consciousness corresponding to the inanimate universe prior to the origin of life.

4.15 The implausibility of cosmic inner life objection

The last objection I will address in this essay is, like the first, purely intuitive. It asks rhetorically: How plausible is it that the inanimate universe as a whole is the extrinsic appearance of some kind of universal inner life? The intuitive appeal of the question is understandable. After all, we only have introspective access to our own (dissociated) personal inner life, so to gauge the presence of other or broader inner life we depend on perceivable external indicators. In other people and animals, these indicators are their behavior. But within the extremely small range of space and time in which we live our lives—and even in which human history as a whole has unfolded—we simply cannot perceive any intuitively-appealing indicator of universal inner life.

Yet, we can approach the question from a different angle. Consider a living brain exposed by surgeons during an operation. It is a very concrete object that can be seen, touched, cut, cauterized, etc. It is composed of the same types of atoms and force fields that make up the universe as a whole. There is nothing magical about a brain insofar as we can gauge on the screen of perception. And neither can we discern any intuitively-appealing indicator of inner life by simply looking at an exposed brain.

Nonetheless, we all know that 'behind' the living brain lies the entire inner life of a person, with love affairs and heartbreaks, successes and disappointments, great adventures and quiet introspective insights, great joy and indescribable

suffering. 'Behind' that very concrete object under the surgeon's scalpel there lies a world of phenomenality. Counterintuitive or not, this is the way nature is: what we call physical structures—such as living brains—*can* correspond in some way to rich phenomenality. We may not know *how* this is so, but we do know *that* it is so.

Therefore, unless we solve the 'hard problem of consciousness' (Chalmers 2003) and explain what makes brains different from the inanimate universe as a whole in this regard, if brains correspond to inner life it is not at all implausible that the inanimate universe as a whole could as well. After all, brains are made of the same 'stuff' that the rest of the universe is also made of.

One could argue at this point that only particular structural and functional organizations of this 'stuff,' as found in brains, are conducive to the kind of information processing associated with human inner life. For instance, Tononi (2004) has shown that *reportable* experiences correlate only with complex networks of information integration in the brain. Although it has recently been shown that there are structural similarities between brains and the universe at its largest scales (Krioukov *et al.* 2012),[2] it is implausible that analogous information integration takes place at a universal level. The distances and signal propagation times involved do not permit it (Siegel 2016).

However, the hypothesis offered here is not that the universe has *human-like* cognition and associated information integration. As a matter of fact, the hypothesis is not even that the universe has *cognition*, defined as the capacity to acquire knowledge or understanding. Instead, the claim is simply that there is *raw experience*—qualia, pure and simple—associated with the universe as a whole, which does not require anything like the kind of information integration underlying human self-reflection.

4.16 Conclusions

Idealism is a unique ontology in that, unlike physicalism and panpsychism, it asserts that physical structures are circumscribed by consciousness, as opposed to the other way around. Yet, analytic philosophy has traditionally considered idealism implausible. In this essay, I have argued that the alleged implausibility of idealism is based on misconceptions, such as:

- Unfounded intuition—e.g. taking the concreteness of the world to indicate its independence from consciousness, or asserting the implausibility of universal inner life;
- Lack of philosophical imagination—e.g. assuming that multiple private minds and a stand-alone world cannot be coherently reduced to a single universal consciousness;

[2] This conclusion has been confirmed and amplified by a later study done by Franco Vazza and Alberto Feletti (2017).

- Demonstrably wrong assumptions—e.g. that all mental activity is acquiescent to volition;
- Question-begging—e.g. arguing that different people cannot share a dream because their bodies are separate, and arguing that the universe cannot be in consciousness because it existed before conscious life first arose;
- Anthropomorphization—e.g. taking all conceivable processes in consciousness to necessarily be unstable and disorderly;
- Failure to understand the implications of idealism—e.g. asserting that a field of phenomenality outside personal psyches is equivalent to a physical world outside phenomenality;
- Unexamined assumptions—e.g. that the physical is in some sense distinct from, yet causally effective upon, the phenomenal;
- Conflation—e.g. conflating consciousness proper with self-reflection, conflating unconsciousness with failure to recall phenomenality, and conflating idealism with solipsism.

As such, idealism is an entirely plausible ontology that may offer the most parsimonious and explanatorily powerful option yet to make sense of reality.

5. There Is an 'Unconscious,' but It May Well Be Conscious

This paper first appeared in *Europe's Journal of Psychology*, ISSN: 1841-0413, Vol. 13, No. 3, pp. 559-572, in August 2017. A summary of this paper has also appeared in *Scientific American* on 19 September 2017.[1]

5.1 Abstract

Depth psychology finds empirical validation today in a variety of observations that suggest the presence of causally effective mental processes outside conscious experience. I submit that this is due to misinterpretation of the observations: the *subset* of consciousness called 'meta-consciousness' in the literature is often mistaken for consciousness proper, thereby artificially creating space for an 'unconscious.' The implied hypothesis is that *all* mental processes may in fact be conscious, the *appearance* of unconsciousness arising from our dependence on self-reflective introspection for gauging awareness. After re-interpreting the empirical data according to a philosophically rigorous definition of consciousness, I show that two well-known phenomena corroborate this hypothesis: (a) experiences that, despite being conscious, aren't re-represented during introspection; and (b) dissociated experiences inaccessible to the executive ego. If consciousness is inherent to all mentation, it may be fundamental in nature, as opposed to a product of particular types of brain function.

5.2 Introduction

The foundational theoretical inference of the clinical approach called 'depth psychology'—whose origins can be traced back to the works of Frederic Myers, Pierre Janet, William James, Sigmund Freud and Carl Jung—is that the human psyche comprises two main subdivisions: a conscious and an unconscious segment (Kelly *et al*. 2009: 301-334). The conscious segment comprises mental activity to which one has introspective access. The so-called 'ego' is the felt sense of personal self that arises in association with a subset of this introspectively-accessible activity—e.g. some bodily sensations, images, thoughts, beliefs, etc.—and it is in this sense that I use the word 'ego' throughout this paper. In contrast, the unconscious segment comprises mental activity to which one has no introspective access. Inaccessible as it may be, depth psychologists contend that mental activity in the 'unconscious'—a term often used as a noun—still can and does influence one's conscious thoughts, feelings and behaviors. A more modern articulation of the notion of a *mental* unconscious—as opposed to what has historically been called "unconscious

[1] At the time of this writing, the *Scientific American* essay was freely available online at: https://blogs.scientificamerican.com/observations/consciousness-goes-deeper-than-you-think/.

cerebration" (Kelly *et al*. 2009: 340-352)—can be found in the writings of Kihlstrom (1997), for example.[2]

Recent empirical results seem to corroborate the hypothesis of a mental unconscious by revealing the presence of mental activity individuals cannot access through introspection, but which nonetheless causally conditions the individuals' conscious thoughts, feelings and behaviors (e.g. Westen 1999, Augusto 2010, Eagleman 2011). Hassin goes as far as insisting, "unconscious processes can carry out every fundamental high-level function that conscious processes can perform" (2013: 196). He reviews empirical evidence indicating that the unconscious is capable of cognitive control, the pursuit of goals, information broadcasting and even reasoning (Hassin 2013: 197-200). This echoes Dijksterhuis and Nordgren, whose experiments indicate that the unconscious can encompass "all psychological phenomena associated with thought, such as choice, decision making, attitude formation and attitude change, impression formation, diagnosticity, problem solving, and creativity" (2006: 96). Even practitioners of cognitive therapy, who have traditionally ignored the unconscious, have more recently found clinical value in interpreting possible indirect manifestations of inaccessible mental activity in the form of dreams (Rosner, Lyddon and Freeman 2004). This new scientific approach to the hypothesis of an unconscious has been called "the new unconscious" (Hassin, Uleman and Bargh 2005).

Clearly, there is significant evidence for the presence of causally-effective mental activity that we ordinarily cannot access through introspection. The question, however, is whether mental activity inaccessible through introspection is necessarily *unconscious*. It is true that, from the perspective of clinical psychology, these two modalities are operationally indistinguishable, since the clinicians' sole gauge of their patients' range of consciousness is the patients' own introspective reports. However, from a theoretical standpoint, it is conceivable that mental activity the ego cannot access through introspection could still be conscious, in the sense of being phenomenally experienced somewhere in the psyche. If so, this has significant implications for our understanding of the nature of consciousness—and of its relationship to brain function—in the fields of neuropsychology, neuroscience and philosophy of mind.

Indeed, although the conflation between lack of introspective access and lack of consciousness is operationally justifiable in a clinical setting, the widespread use of the qualifier 'unconscious' today suggests an intrinsic dichotomy in the

[2] Throughout this dissertation, I use the word 'mental' as a synonym of 'phenomenal'; *except in this chapter*. Because this chapter was originally published as an article in a psychology journal, here the word 'mental' is associated with cognitive activity, instead of qualia. According to this definition, mental processes aren't necessarily conscious, for cognition can conceivably take place unconsciously. And if they are conscious, mental processes then entail the acquisition of knowledge and/or understanding, which implies more than just the presence of phenomenality.

nature of mental processes: some supposedly *aren't* experienced whilst others, somehow, are. This implies that consciousness is not fundamental to mentation, but a property that emerges from particular arrangements or configurations of neurons. Primed and driven by this assumption, significant resources are spent in neuropsychology and neuroscience today in an effort to figure out what these arrangements or configurations are. Hypotheses currently under investigation vary from vast topologies of information integration across neurons (Tononi 2004) to microscopic quantum processes within neural microtubules (Hameroff 2006).

The present paper, on the other hand, elaborates on the possibility that these efforts are misguided, for introspectively-inaccessible mental processes may still be conscious: they may be phenomenally experienced in a manner—or in a segment of the psyche—that escapes egoic introspection. This way, the notion of an unconscious, despite the broad use and influence of the term in today's psychology, may at root be a linguistic inaccuracy originating from mere operational convenience. If so, then consciousness may not be the product of specific arrangements or configurations of neural activity, but a fundamental property of *all* mentation. The implications of this possibility for neuropsychology, neuroscience and philosophy of mind are hard to overestimate.

5.3 Defining and gauging consciousness

Before we can meaningfully discuss unconsciousness—the alleged lack of consciousness—we must, of course, have clarity regarding the meaning of the word 'consciousness.' What does it mean to say that a mental process is conscious? In this paper, I shall use a rigorous definition well-accepted in neuropsychology, neuroscience and philosophy of mind: mental activity is *conscious* if, and only if, there is something—*anything*—it is like to have such mental activity in and of itself (Nagel 1974, Chalmers 2003). (A less rigorous but more easily understandable formulation of this definition is this: mental activity is conscious if there is something it *feels* like to have such mental activity in and of itself. The verb 'to feel,' however, is too ambiguous to be used in a rigorous definition, so philosophers of mind have reached consensus around the formulation I originally proposed above.) This way, if mental activity is *un*conscious, then there is nothing it is like to have such activity in and of itself, even if it, in turn, causes or influences conscious activity. Notice that this definition of consciousness honors our intuitive understanding of the word: you only consider yourself conscious right now because there is something it is like to be you while you read this paper. Otherwise, you would necessarily be unconscious.

To remain consistent with our intuitive understanding of words, I shall also say that mental activity corresponds to *experience* if, and only if, it is conscious. You experience reading this paper because you are conscious of it right now. If you were not, what sense would there be in saying that you experience it?

According to these definitions, higher-order thought (as defined in Schooler 2002: 340) is unnecessary for there to be consciousness. The presence of the mere qualities of raw experience—which philosophers of mind call *qualia*—is already sufficient for a mental process to be considered conscious. In this context, the categorization proposed by Schooler is helpful: he distinguishes between "non-conscious (unexperienced), conscious (experienced), and meta-conscious (re-represented)" mental processes (2002: 339). Only the latter entails higher-order thought.

Now notice that *direct* insight into one's conscious inner life is limited to those experiences one's ego can access through introspection and then report to self or others. In the words of Klein, "It is *only* in virtue of knowledge by acquaintance that we know our mental states. … Accordingly, the use of introspective reports as a reliable and informative source of information about mental states has seen a resurgence over the past few decades" (2015: 361, original emphasis). For this reason, the study of the Neural Correlates of Consciousness (NCCs) still largely consists in correlating objective measurements of neural activity with introspective assessments (Koch 2004): patterns of neural activity accompanied by reported experience are considered NCCs. Indeed, as Newell and Shanks recently wrote, "Whereas issues about how to define and measure awareness were once highly prominent and controversial, it now seems to be generally accepted that awareness should be operationally defined as reportable knowledge" (2014: 15).

The problem is that, as I shall shortly elaborate upon, for the subject's ego to access and report an experience there must be: (a) an associative link between the ego and the experience; and (b) a meta-conscious re-representation of the experience. Therefore, while subjects can report non-dissociated meta-conscious processes, *they fundamentally cannot distinguish between truly unconscious processes and conscious processes that simply aren't meta-conscious*, for *both* types are equally unreportable to self and others. This is an alarming conclusion, for much of the work indicating the presence of an unconscious is based on (the lack of) introspective reports of experience. The next two sections expand on all this.

In what follows, I shall assume that introspective reports are as good as "reliable, relevant, immediate, and sensitive" (Newell and Shanks 2014: 3). This is charitable towards the hypothesis of an unconscious, for—as Newell and Shanks argued (2014)—much of the evidence behind this hypothesis can be attributed to methodological artifacts: delayed introspective assessments leading to impaired recall, experimenters not providing sufficient opportunity for subjects to report the introspective insights they actually have, cross-task confusion, etc. My goal is to show that, *even if* the research underpinning the existence of an unconscious were free of methodological artifacts, there would *still* be compelling reasons to posit that mental processes unaccompanied by introspective reports of experience can be conscious nonetheless.

5.4 Non-self-reflective experiences

To gain introspective access to an experience it is not enough to merely have the experience; we must also consciously know *that* we have it. After all, what introspective insight could we gain about an experience of which we are not explicitly aware? Schooler elaborates:

> Critical to both the centrality of the conscious/non-conscious distinction, and its equation with reportability, is the assumption that people are explicitly aware of their conscious experiences. However, this assumption is challenged when subjective experience is dissociated from the explicit awareness of that experience. Such dissociations demonstrate the importance of distinguishing between consciousness and 'meta-consciousness.' (2002: 339.)

The conscious knowledge *of* the experience—which comes *in addition to* the experience itself—is what Schooler calls a "re-representation":

> Periodically attention is directed towards explicitly assessing the contents of experience. The resulting meta-consciousness involves an explicit *re-representation* of consciousness in which one interprets, describes, or otherwise characterizes the state of one's mind. (2002: 339-340, emphasis added.)

Although re-representation is necessary for introspection, it is largely absent, for instance, in dreams (Windt and Metzinger 2007). This demonstrates compellingly that mental activity does *not* need to be re-represented in order to be experienced—after all, who can seriously doubt that dreams are experienced?—but only to be introspectively accessed. During ordinary dreams we simply experience, without consciously knowing *that* we experience.

More formally, suppose that one has an experience *X*. To gain introspective access to *X* one must have conscious knowledge *N* of *X*. But *N*—the "re-representation"—is a separate experience in its own right. One experiences the *knowing of X* as a quality closely related to, but distinct from, *X* itself. *N* is not encompassed, entailed or implied by *X*. Indeed, Schooler highlights the fact that re-representations can even *mis*represent the original experiences:

> Once meta-consciousness is triggered, translation dissociations can occur if the re-representation process misrepresents the original experience. Such dissociations are particularly likely when one *verbally* reflects on non-verbal experiences or attempts to take stock of ambiguous or subtle perceptual experiences. (2002: 340, emphasis added.)

To make these abstract considerations more concrete, consider your breathing right now: the sensation of air flowing through your nostrils, the movements of your diaphragm, the inflation and deflation of your lungs, etc. Were you *not* experiencing these sensations a moment ago, before I directed your attention

to them?[3] Or were you just unaware *that* you were experiencing them all along? By directing your attention to these sensations, did I make them *conscious* or did I simply cause you to experience the *extra* quality of knowing *that* the sensations were conscious? Clearly, even waking experiences can occur without re-representation.

Re-representations are the product of a self-reflective *configuration of* consciousness, whereby the latter turns in upon itself so to objectify its own contents (Kastrup 2014: 104-110). In humans, this usually occurs through the use of "semiotic mediation" (Valsiner 1998), which is our ability to re-represent our experiences by *naming* them explicitly or implicitly. Gillespie gives an example: "In order to obtain dinner one must first name … one's hunger … This naming, which is a moment of self-reflection, is the first step in beginning to construct, semiotically, a path of action that will lead to dinner" (2007: 678).

Naturally, nothing prevents experiences from occurring outside the field of self-reflection—that is, occurring without being explicitly or implicitly named. Nixon, for instance, calls these "unconscious experiences" (2010: 216), which in my view is an oxymoron but illustrates the subtlety of the point. He lists several examples: blindsight (Stoerig and Cowey 1997), prosopagnosia (Sacks 1985), sleepwalking, post-hypnotic suggestion, etc. Indeed, the emergence of so-called "no-report paradigms" in contemporary neuroscience attests to the abundant presence of waking experiences that are unreportable because they fall outside the field of self-reflection (Tsuchiya *et al*. 2015, Vandenbroucke *et al*. 2014).

Moreover, *the neural activity patterns of the NCCs themselves* suggest circumstantially—yet compellingly—that many NCCs correspond merely to a self-reflective configuration of consciousness. To see this, notice first that the conscious knowledge *N* of an experience *X* is triggered by the occurrence of *X*. For instance, it is the occurrence of a sense perception that triggers the realization that one is perceiving something. *N*, in turn, evokes *X* by directing attention back to it: the realization that one is perceiving something naturally shifts one's mental focus back to the original perception. So we end up with a back-and-forth cycle of evocations whereby *X* triggers *N*, which in turn evokes *X*, which again triggers *N*, and so forth. See Figure 5.1 for an illustration.

[3] Notice that *attention* is required to explicitly assess an experience at a metacognitive—that is, self-reflective—level.

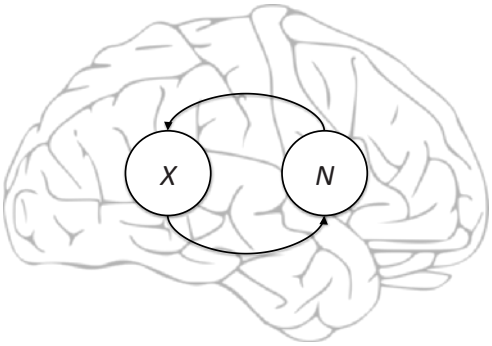

Figure 5.1: Illustrative caricature of oscillatory evocations between an experience (*X*) and the meta-conscious knowledge of the experience (*N*).

As it turns out, recent characterizations of the NCCs show precisely this pattern of reverberating back-and-forth communications between different brain regions (Dehaene and Changeux 2011, Boly *et al*. 2011, van Gaal *et al*. 2011). When damage to the primary visual cortex presumably interrupts this reverberation, patients display blindsight (Paller and Suzuki 2014: 387)—that is, the ability to correctly discriminate moving objects despite the *reported* inability to see them. This is precisely what one would expect if the reverberation in question were the oscillations between *X* and *N*: the objects *are* consciously perceived—therefore explaining how the patients can discriminate them—but the patients do not know *that* they consciously perceive the objects.

I thus submit that many NCCs are, in fact, the correlates only of a potentially very small subset of consciousness—namely, meta-consciousness or self-reflection—instead of consciousness proper. The introspectively inaccessible character of experience that isn't re-represented constitutes the first mechanism through which seemingly unconscious mental activity may, in fact, be conscious. There is yet another mechanism, which will be explored in the next section.

5.5 Dissociated experiences

Dissociative states are well recognized in psychiatry today, featuring prominently in the DSM-5 (American Psychiatric Association 2013). Their hallmark is "a disruption of and/or discontinuity in the normal integration of consciousness, memory, identity, emotion, perception, body representation, motor control, and behavior" (Black and Grant 2014: 191). In other words, dissociation entails fragmentation of the contents of consciousness.

There are different forms of dissociation. Klein (2015), for instance, discusses a form in which the subject's ego loses the sense of ownership of some of the subject's own mental states. This occurs when consciousness can no longer

"relate to its object in a particular, self-referential way" (Klein 2015: 362). He lists several examples, such as the case of a man who, after an accident, could accurately report the content of his memories but "was unable to experience that content as his own" (Klein 2015: 368). Notice, however, that the man's ego could *still* access the content; just not identify with it.

In what follows, I shall focus on a strong form of dissociation in which the ego *cannot even access* certain contents of consciousness. In its pathological variations, this is known as Dissociative Identity Disorder (DID). A person suffering from DID exhibits multiple, disjoint centers of consciousness called alters. Each alter experiences the world as a distinct personality (Braude 1995).

Although there has been debate about the authenticity of DID as a psychiatric condition—after all, it is conceivable that patients could fake it—research has confirmed DID's legitimacy (Kelly *et al*. 2009: 167-174 & 348-352). Two recent studies are particularly interesting to highlight. In 2015, doctors reported on the case of a German woman who exhibited a variety of alters (Strasburger and Waldvogel). Peculiarly, some of her alters claimed to be blind while others could see normally. Through EEGs, the doctors were able to ascertain that the brain activity normally associated with sight wasn't present while a blind alter was in control of the woman's body, even though her eyes were open. When a sighted alter assumed executive control, the usual brain activity returned. This is a sobering result that shows the literally *blinding* power of dissociation. In another study (Schlumpf *et al*. 2014), investigators performed functional magnetic resonance imaging (fMRI) brain scans on both DID patients and actors simulating DID. The scans of the actual patients displayed clear and significant differences when compared to those of the actors. Undoubtedly, thus, DID is real.

Normally, only one of the alters has executive control of the body at any given moment. The important question for the purposes of the present paper is then this: Can the *other* alters, who are *not* in control of the body, remain conscious or do they simply fade into unconsciousness? If they can remain conscious, the implication is that a person can have multiple *concurrent* but dissociated centers of *consciousness*, as originally hypothesized by Frederic Myers and Pierre Janet (Kelly *et al*. 2009: 305-317). Presumably, then, each center has its own private, parallel stream of experiences.

Occasionally, however, the dissociation isn't bilateral: a first alter is able to gain partial access to the experiences of a second, without the second alter being able to access the experiences of the first. This rare kind of unilateral dissociation provides tantalizing indications that alters can remain conscious even when not in control of the body. In Morton Prince's well-known study of the 'Miss Beauchamp case' of DID, one of the alters—called Sally—"was a co-conscious personality in a deeper sense. When she was not interacting with the world, she did not become dormant, but persisted and was active" (Kelly *et al*. 2009: 318). Sally maintained that she knew

> everything Miss Beauchamp … does at the time she does it,—knows what she thinks, hears what she says, reads what she writes, and sees

> what she does; that she knows all this as a separate co-self, and that her knowledge does not come to her afterwards … in the form of a memory. (Prince, as quoted in Kelly *et al*. 2009: 318)

Stephen Braude's more recent work reinforces the view that alters can be co-conscious "discrete centers of self-awareness" (1995: 67). He points—as evidence for this hypothesis—at the struggle of different alters for executive control of the body and the fact that alters "might intervene in the lives of others [i.e. other alters], intentionally interfering with their interests and activities, or at least playing mischief on them" (Braude 1995: 68). It thus appears that alters can not only be concurrently conscious, but that they can also vie for dominance with each other.

Strong dissociation is not restricted to DID—its extreme form—or to pathology, for that matter. Indeed, the foundational hypothesis of depth psychology entails a form of natural dissociation between the conscious ego and the so-called 'unconscious.' As such, it is plausible—in fact, there is overwhelming clinical evidence for it in the annals of depth psychology—that we all have at least one dissociated mental subsystem that we cannot access through introspection. Ernest Hilgard (1977) conceived of these dissociated subsystems as conscious, much as Myers, Janet and Braude did.

Thus, the possibility that presents itself to us is that we may all have one or more *conscious* 'others' within ourselves, dissociated from our ego. If this is so, then (a) our ego ordinarily has no introspective access to the experiences of these 'others'; and, consequently, (b) the study of the NCCs is largely blind to the potentially idiosyncratic patterns of neural activity corresponding to such dissociated experiences. This is the second mechanism through which apparently unconscious mental activity may, after all, be conscious.

5.6 A model of dissociation

Wegner (2002) proposes an analogy for explaining alters: different operating systems running on the same hardware. This way, the transfer of executive control from one alter to another would be analogous to shutting down Windows and rebooting the computer with Linux. This, of course, only accounts for strictly alternating personalities and thus fails to explain much of the clinical data cited above. Nonetheless, it still suggests a starting point for a plausible model of dissociation.

If we define an *experiential frame* as the set of all qualities we experience at a given moment—encompassing our conscious perceptions, thoughts, emotions, bodily sensations, imagination, etc.—conscious life can be modeled as a chain of experiential frames. This is graphically illustrated in Figure 5.2, wherein experiential frames F1 to F*n* are shown. Each frame is evoked by the previous frame through cognitive associations, in the sense that e.g. our particular thoughts in the present moment largely determine which emotions we experience in the next moment; or that our emotions in the present moment largely determine our actions—and therefore perceptions—in the next

moment; and so on. These cognitive associations are represented by the arrows linking frames together in Figure 5.2.

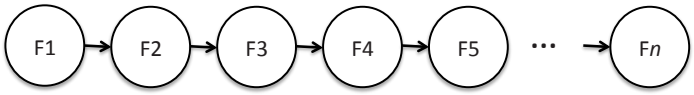

Figure 5.2: Conscious life as a chain of experiential frames connected through cognitive associations.

Wegner's suggestion can then be visualized as in Figure 5.3. The chain of experiential frames—denoted F—corresponding to a first alter is interrupted by experiential frames—denoted F'—corresponding to a second alter. The key point is that, once executive control is assumed by the experiential frames F' of the second alter, the corresponding experiential frames F of the first alter *cease to exist*. There is no parallelism of experience: either the mental contents of the first alter are experienced or those of the second alter; never those of both concurrently. As such, this is a *sequential model of dissociation* and, as we've seen, it isn't sufficient to explain the clinical data cited.

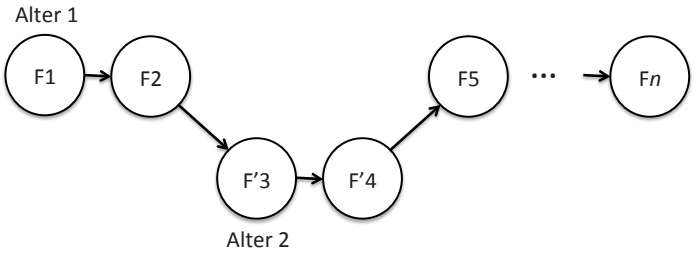

Figure 5.3: The sequential model of dissociation in the context of DID.

Alternatively, we can hypothesize that the chains of experiential frames of *both* alters are *always* present, concurrently and in parallel. Executive control of the body simply switches between the two parallel chains, as shown in Figure 5.4. Experiential frames drawn in grey represent those without executive control, *but still conscious*. This is thus a *parallel model of dissociation*, which illustrates the hypothesis of "co-consciousness" (a term originally coined by Morton Prince, as discussed by Kelly *et al*. 2009: 317).

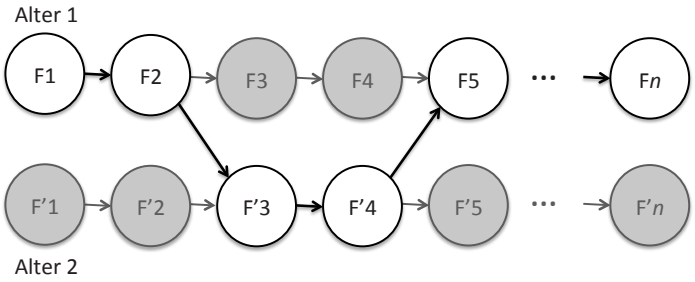

Figure 5.4: The parallel model of dissociation in the context of DID.

We have seen that DID is a pathological form of dissociation, but that we may all naturally have strongly dissociated mental subsystems that never—or very seldom—vie for executive control of the body. These would constitute the so-called 'unconscious' of depth psychology. Figure 5.5 illustrates how such strongly dissociated mental subsystems can be modeled under the proposed framework. For simplicity, only the ego and one dissociated subsystem are shown. The 'other' in this case—represented by the dissociated chain of experiential frames F'—is content to live its inner life in the background of egoic activity. It only manifests its presence through indirect, subtle influences on egoic experiences, as represented by the dashed arrows vertically linking the two chains. These subtle influences can take many forms, such as: dissociated emotions influencing our egoic thoughts and behaviors (Lynch and Kilmartin 2013: 100); dissociated beliefs and expectations influencing our egoic perceptions (Eagleman 2011: 20-54); dissociated drives manifesting themselves symbolically in the form of dreams (von Franz and Boa 1994, Jung 2002, Fonagy *et al*. 2012); etc.

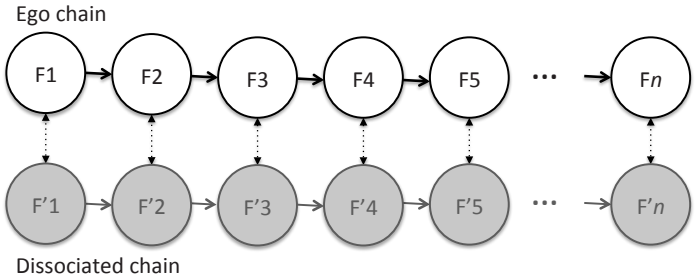

Figure 5.5: The parallel model of dissociation in a depth-psychological context.

Admittedly, limitations in our ability to gauge consciousness currently prevent us from asserting with certainty, on an empirical basis, that the parallel model of dissociation is correct. However, by the same token, we can also not assert that it isn't. The brain seems to have sufficient resources for this kind of

parallelism and, if anything, the clinical data is suggestive of its validity (again, Kelly *et al*. 2009: 305-322 and Braude 1995). The parallel model should, therefore, be considered not only plausible but perhaps even probable, in which case it further substantiates the notion that the 'unconscious' may be—well—conscious.

5.7 Discussion

I have elaborated on the hypothesis that there may be no such a thing as an unconscious mental process. All mental processes may be conscious, in the sense that there may be something it is like to have such mental processes in and of themselves. Our impression that some mental processes are unconscious may arise from (a) their consisting in non-self-reflective experiences not amenable to introspection or (b) their being strongly dissociated from the executive ego and, therefore, inaccessible to it.

Underlying this entire paper is the differentiation between consciousness proper and particular *configurations of* consciousness, such as self-reflection and dissociative states. It is rather disturbing how often these notions are conflated not only in general psychology, but also in neuroscience and philosophy of mind. For instance, a relatively recent article (Gabrielsen 2013) talks about the emergence of consciousness in human babies when what is discussed is—as per the argument developed in this paper—likely to be the emergence of meta-consciousness.[4]

Dijksterhuis and Nordgren also "define conscious thought as object-relevant or task-relevant cognitive or affective thought processes that occur while the object or task is *the focus of one's conscious attention*" (2006: 96, emphasis added). They insist, "it is very important to realize that *attention is the key* to distinguish [*sic*] between unconscious thought and conscious thought. *Conscious thought is thought with attention*" (Dijksterhuis and Nordgren 2006, emphasis added). In appealing to *attention*, as opposed to experience or *qualia*, they are implicitly associating consciousness with self-reflection or re-representation, as discussed in Section 5.4.

Even more strikingly, Cleeremans (2011) *explicitly defines* consciousness as self-reflection. He overtly conflates experience with meta-consciousness and reportability:

> *Awareness*, on the other hand, always seems to minimally entail the ability of knowing *that* one knows. This ability, after all, forms the basis for the verbal reports we take to be the most direct indication of awareness. And when we observe the absence of such ability to report on the knowledge involved in our decisions, we rightfully conclude that the decision was based on unconscious knowledge. Thus, it is when an agent

[4] For clarity, by "emergence of meta-consciousness" I mean here the early, or even precursor, stages of meta-consciousness.

> exhibits *knowledge* of the fact that he is sensitive to some state of affairs that we take this agent to be a conscious agent. This *second-order* knowledge, I argue, critically depends on *learned* systems of meta representations, and forms the basis for conscious experience. (Cleeremans 2011: 3)

This isn't a recent problem. When one reads the original texts of the founders of depth psychology whilst holding the distinction between consciousness and meta-consciousness in mind, one quickly realizes that, when they spoke of unconsciousness, the founders often meant a lack of *meta*-consciousness—not of experience proper. This is abundantly evident, for instance, in an essay written by Carl Jung in the 1920s or early 1930s, called "The Stages of Life" (Jung 2001: 97-116).

It could be argued that the distinction between experience and meta-consciousness is merely a semantic point. However, consider this: by conflating consciousness proper with *self-reflective* consciousness, we also indirectly equate non-self-reflective consciousness with unconsciousness; we absurdly imply that dreams—which largely lack self-reflection (Windt and Metzinger 2007)—aren't experienced. Instead of the three categories proposed by Schooler—namely, "non-conscious (unexperienced), conscious (experienced), and meta-conscious (re-represented)" (2002: 339)—we are left with only two: non-conscious and meta-conscious. Consequently, we are forced to collapse the conscious onto the non-conscious and, in the process, end up disregarding the extraordinary phenomenon of *qualities of experience*.[5] Clearly, this isn't merely semantic.

Most importantly, the philosophical implications of mistaking consciousness for *meta*-consciousness are significant. If some mental processes were truly unconscious while others are conscious, it would follow that consciousness is the product of some specific anatomical and/or functional arrangements of brain activity. In other words, consciousness would be derivative, as opposed to fundamental. Philosophically, this would corroborate the ontology of physicalism (Stoljar 2016) while contradicting alternatives like panpsychism (Strawson *et al*. 2006), cosmopsychism (Shani 2015) and idealism (Kastrup 2017b). It would leave us with no way to circumvent the arguably insoluble 'hard problem of consciousness' (Chalmers 2003).

On the other hand, if consciousness is inherent to all mental processes, then the specific anatomical and/or functional parameters of different processes correspond merely to different *contents and/or configurations of consciousness*—that is, to the particular qualities that are experienced—but do not determine the presence or absence of consciousness itself. This allows us to circumvent the 'hard problem of consciousness' altogether, by inferring that consciousness is primary. While it's not my intent in this paper to argue for or against any particular ontology of mind, it is significant that a lucid, critical

[5] That is, we end up sweeping the 'hard problem of consciousness' (Chalmers 2003) under the rug.

interpretation of the available empirical data leaves more avenues of philosophical inquiry open.

If we are true to the spirit of the words 'consciousness' and 'experience,' diligent in our interpretation of empirical observations—both experimental and clinical—and rigorous in our use of concepts, we are led not only to the conclusion that *all* mental processes may be conscious, but that consciousness itself may be fundamental.

6. Self-Transcendence Correlates with Brain Function Impairment

This paper first appeared in the *Journal of Cognition and Neuroethics*, ISSN: 2166-5087, Vol. 4, No. 3, pp. 33-42, in January 2017. A summary of this paper has also appeared in *Scientific American* on 29 March 2017.[1]

6.1 Abstract

A broad pattern of correlations between mechanisms of brain function impairment and self-transcendence is shown. The pattern includes such mechanisms as cerebral hypoxia, physiological stress, transcranial magnetic stimulation, trance-induced physiological effects, the action of psychoactive substances and even physical trauma to the brain. In all these cases, subjects report self-transcending experiences often described as "mystical" and "awareness-expanding," as well as self-transcending skills often described as "savant." The idea that these correlations could be rather trivially accounted for on the basis of disruptions to inhibitory neural processes is reviewed and shown to be implausible. Instead, this paper suggests that an as-of-yet unrecognized causal principle underlying the entire pattern might be at work, whose further elucidation through systematic research could hold great promise.

6.2 Introduction

In this paper, 'self-transcendence' is defined as the abrupt—thus *not* gradual—broadening of one's sense of self through a step-function enrichment of one's subjective inner life. This can happen, for instance, when one suddenly acquires (a) a feeling that one is no longer confined to the spatio-temporal locus of the physical body; (b) entirely new mental skills that one has never attempted to develop through learning or training; or (c) unfamiliar emotions, insights or inner imagery. This essay attempts to show that there is a consistent pattern of correlations between self-transcendence—so defined—and a broad variety of brain function impairment mechanisms. In other words, several types of brain function *impairment* are consistently accompanied by *richer* inner life. This is counterintuitive and suggests a common underlying causal principle yet to be understood in its full scope.

In the next sections, several mechanisms of brain function impairment and the resulting self-transcendence effects will be reviewed. The goal is to establish a broad pattern by highlighting the similarities of the mechanisms and their effects.

---

[1] At the time of this writing, the *Scientific American* essay was freely available online at: https://blogs.scientificamerican.com/guest-blog/transcending-the-brain/.

6.3 Cerebral hypoxia

Fainting or near-fainting caused by restrictions of oxygen supply to the brain is known to induce liberating feelings of self-transcendence. For instance, the potentially fatal 'choking game' played by teenagers worldwide (Macnab 2009) is an attempt to induce such feelings through partial strangulation (Neal 2008: 310-315). The psychotherapeutic technique of holotropic breathwork (Rhinewine and Williams 2007), as well as more traditional yogic breathing practices, use hyperventilation to achieve similar effects: by increasing blood alkalinity levels, they interfere with normal oxygen uptake in the brain and ultimately lead to what is described as an expansion of ordinary awareness (Taylor 1994). Even straightforward hyperventilation outside a therapeutic context can lead to self-transcending experiences, such as described in this anecdotal—though representative—report:

> One of us stood against a tree and breathed deeply for a while and then took a very deep breath. Another pushed down hard on his ribcage … This rendered the subject immediately unconscious … When I tried it, I didn't think it would work, but then suddenly I was in a meadow which glowed in yellow and red, everything was extremely beautiful and funny. This seemed to last for ages. I must say that I have never felt such bliss ever again. (Retz 2007)

Finally, pilots undergoing G-force induced Loss of Consciousness (G-LOC)—whereby blood is forced out of the brain, causing hypoxia—report "memorable dreams" phenomenologically similar to near-death experiences (Whinnery and Whinnery 1990), which are notoriously self-transcending in character.

6.4 Generalized physiological stress

Near-Death Experiences (NDEs) are the prime examples of self-transcendence associated with dramatically reduced brain function due to e.g. cardiac arrest (van Lommel 2001). They reportedly entail life-transforming phenomenality—encompassing insights, emotions and rich inner imagery—far surpassing the envelop of ordinary experiences (Kelly *et al.* 2009: 367-421), despite overwhelming disruption to the brain's ability to operate. A recent and well-publicized NDE, which occurred while the patient was under close supervision of medical staff, captures this self-transcendent dimension. In the patient's own words:

> I certainly don't feel reduced or smaller in any way. On the contrary, I haven't ever been this huge, this powerful, or this all-encompassing. … [I] felt greater and more intense and expansive than my physical being. (Moorjani 2012: 69)

In a related manner, traditional initiatory rituals in pre-literate cultures sought to reveal the true nature of self and world through physical ordeals (Eliade 2009). It is reasonable to imagine that these ordeals—such as long sessions in sweat lodges, exposure to the elements, extreme exertion and even poisoning—

physically compromised brain function through generalized physiological stress, thereby inducing self-transcending experiences.

6.5 Electromagnetic impairment

The use of transcranial magnetic stimulation can inhibit activity in localized areas of the brain by impairing the associated electromagnetic fields. As reported in a study (Blanke *et al*. 2002), when neural activity in the angular gyrus of a patient with epilepsy was inhibited in this way, self-transcending out-of-body experiences were induced.

6.6 Trance-induced impairment

During the practice of so-called 'psychography,' an alleged medium enters a trance state and writes down information allegedly originating from a transcendent source beyond the medium's ordinary self. A neuroimaging study (Peres 2012) revealed that experienced mediums displayed marked reduction of activity in key brain regions—such as the frontal lobes and hippocampus—when compared to regular, non-trance writing. Despite this, text written under trance scored consistently higher in a measure of complexity than material produced without trance. As an observant science journalist remarked, more complex writing

> typically would require more activity in the frontal and temporal lobes—but that's precisely the opposite of what was observed. To put this another way, the low level of activity in the experienced mediums' frontal lobes should have resulted in vague, unfocused, obtuse garble. Instead, it resulted in more complex writing samples than they were able to produce while not entranced. Why? No one's sure. (DiSalvo 2012)

6.7 Chemical impairment

Psychedelic substances have been known to induce powerful self-transcending experiences (Strassman 2001, Griffiths *et al*. 2006, Strassman *et al*. 2008). It had been assumed that they did so by exciting parts of the brain. Yet, recent neuroimaging studies have shown that psychedelics do largely the opposite (Carhart-Harris *et al*. 2012, Palhano-Fontes *et al*. 2015, Carhart-Harris *et al*. 2016).[2] In an article he wrote for *Scientific American Mind*, neuroscientist Christof Koch (2012b) expressed his surprise at these results. Carhart-Harris (2012: 2138), for instance, reported "only decreases in cerebral blood flow" under the influence of a psychedelic. Perhaps even more significantly, "the magnitude

[2] A later study performed at the University of Zürich has confirmed this further, showing that a psychedelic causes "significantly reduced absolute perfusion" (that is, blood flow) in just about every region of the brain, whilst leading to "profound subjective drug effects" (Lewis *et al*. 2017).

of this decrease [in brain activity] predicted the intensity of the subjective effects" of the psychedelic (*ibid*.). As such, the significant self-transcending experiences that follow psychedelic intake are—counterintuitively—accompanied by reductions of brain activity.

6.8 Physical damage

If the trend above is consistent, we should expect some types of physical brain damage to also correlate with self-transcending experiences. And indeed, this has been reported. In a recent study (Cristofori 2016), CT scans of more than one hundred Vietnam War veterans showed that damage to the frontal and parietal lobes increased the likelihood of self-transcending "mystical experiences." In a previous study (Urgesi *et al*. 2010), patients were evaluated before and after brain surgery for the removal of tumors, which caused collateral damage in surrounding tissue. Statistically significant increases in feelings of self-transcendence were reported after the surgery.

The self-transcending character of experiences that accompany certain types of brain injury has been evocatively described by neuroanatomist Jill Bolte Taylor, following a stroke that damaged her brain's left hemisphere:

> my perception of my physical boundaries was no longer limited to where my skin met air. I felt like a genie liberated from its bottle. The energy of my spirit seemed to flow like a great whale gliding through a sea of silent euphoria. (Taylor 2009: 67)

The similarity to Moorjani's experience quoted earlier (2012: 69) is striking, despite the latter having been caused by generalized physiological stress, not a left-hemisphere stroke.

Not only 'mystical experiences' correlate with brain damage, but also the emergence of new mental skills. The literature reports many cases of so-called 'acquired savant syndrome,' wherein an accident or disease leading to brain injury gives rise to genius-level abilities (Lythgoe *et al*. 2005, Treffert 2006, Treffert 2009: 1354, Piore 2013). There are examples of such abilities arising after meningitis, bullet wounds to the head and even with the progression of dementia (Miller *et al*. 1998, 2000).

6.9 Discussion

As we've seen, there is a broad pattern associating a variety of brain impairment mechanisms with self-transcending experiences. A potential explanation for this is that brain function impairment could disproportionally affect inhibitory neural processes, thereby generating or bringing into awareness other neural processes associated with self-transcending experiences. There are, however, problems with this explanation.

Under the physicalist assumption that experience is constituted or generated by brain activity, an increase in the richness of experience—as often entailed by

self-transcendence—must be accompanied by an increase in the metabolism associated with the neural correlates of experience (Kastrup 2016b). This is so because (a) there supposedly is nothing to experience but its neural correlates; and (b) richer experience spans a broader information space in awareness that only increased metabolism can create in the physical substrate of the brain. Any other alternative would decouple experience from the workings of the living brain information-wise, contradicting physicalism. As such, it is difficult to see why a reduction of oxygen supply to the brain *as a whole*—as in partial strangulation, hyperventilation, G-LOC, cardiac arrest, etc.—would selectively affect inhibitory neural processes, while maintaining enough oxygen supply to feed an increase in the neural correlates of experience.

Alternatively, one could speculate that self-transcending experiences occur only *after* normal brain function resumes, subsequent to e.g. restoration of oxygen supply. This, however, cannot account for several of the cases reviewed above. For instance, during the neuroimaging studies of the psychedelic state (e.g. Carhart-Harris *et al*. 2012) researchers collected subjective reports of self-transcendence while *concurrently* monitoring the subjects' reduced brain activity levels. The same holds for the neuroimaging study of psychography (Peres 2012). Similarly, in the case of acquired savant (e.g. Treffert 2006, Treffert 2009: 1354) new mental skills are also *concomitant* with the presence of physical damage in the brain. And even in the case of NDEs, there are arguments for why confabulation after resumption of normal brain function cannot account for some of the reported experiences (Kelly *et al*. 2009: 419-421).

Appeals to impairment of inhibitory processes to explain acquired savant syndrome are particularly difficult to defend. They necessarily entail that the savant skills are pre-developed in the brain but remain inhibited. Brain function impairment occasioned by e.g. trauma then supposedly unlocks these dormant skills by shutting down inhibitory processes. One must wonder, however, how the brain could have developed extraordinary skills, such as e.g. prodigious aptitude for calculations, without any training. And if these skills—many of which are advantageous for survival—were latent in us all, why would the brain have evolved to keep them inhibited in the first place?

It is conceivable that individual cases of self-transcendence could have their own idiosyncratic explanation, unrelated to the other cases, and that the overall pattern suggested in this paper is a red herring. For instance, one could tentatively explain (a) the euphoric effects of hypoxia by speculating that it e.g. somehow triggers the brain's reward system, while accounting for (b) the expansion of one's sense of identity beyond the physical body—as reported by Taylor (2009: 67)—through e.g. damage to the orientation association area of the left brain hemisphere. But given the sometimes-striking similarities in the phenomenality reported across the cases reviewed and the fact that *all* cases—despite their different mechanisms of action—entail impairment of brain function, the question is whether it is plausible that no common causal principle is at work.

The current data is at least suggestive of a single, yet-unrecognized causal principle underlying all cases. More systematic studies of the subjective effects of brain function impairment—leveraging e.g. psychedelic compounds and trans-cranial magnetic stimulation—in specific brain regions could help unveil this principle. Could one e.g. reliably trigger savant skills or mystical experiences by inhibiting neural activity in particular areas under controlled conditions? What would the implications of such a scenario be? Questions such as these hold not only great public interest, but also high significance for both neuroscience and neurophilosophy.

7. Concluding Remarks

7.1 Matter as the outer appearance of inner experience

The ontology discussed in this dissertation has at its foundation an observation as simple as it is far-reaching: *matter is the outer appearance of inner experience.* This, *and only this*, is what matter is. Nature generously teaches us this lesson every time we look at a living organism's brain: the neural activity we discern is part of what the organism's inner life *looks like* when registered from a second-person perspective; that is, from across a dissociative boundary. The *matter* constituting those neurons is the extrinsic appearance of feeling, emotion, thought, imagination, etc. And since this is what matter *is*, the inanimate universe—also made of matter—must itself be the extrinsic appearance of universal inner life. After all, why would matter be one thing under one set of circumstances—namely, when constituting a living brain—and then something else under another set of circumstances—namely, when constituting the inanimate universe of rocks, clouds and stars? This indicates that the inanimate universe as a whole must be, in a certain sense, akin to a brain. And indeed, the network topology of the universe at its largest scales does resemble that of a brain (Krioukov *et al*. 2012); so much so that astrophysicist Franco Vazza and neuroscientist Alberto Feletti considered the similarity "truly remarkable" and "striking":

> It is truly a remarkable fact that the cosmic web is more similar to the human brain than it is to the interior of a galaxy; or that the neuronal network is more similar to the cosmic web than it is to the interior of a neuronal body. Despite extraordinary differences in substrate, physical mechanisms, and size, the human neuronal network and the cosmic web of galaxies, when considered with the tools of information theory, are strikingly similar. (2017)

Allow me to reiterate my point, in the hope that repetition helps reveal its full force: 'matter' is merely the name we give to the extrinsic appearance of conscious experience, as perceived from across a dissociative boundary. *There is nothing more to it.* This painfully simple insight, repeatedly intimated in nature, is all one needs to come to a categorical interpretation of natural phenomena that answers all fundamental questions and avoids all fundamental problems, such as the 'hard problem of consciousness' and the 'subject combination problem.'

7.2 Alternative formulations of dissociation-based idealism

The formulation of idealism developed in this dissertation rests on the notion of *dissociation*—a localized blockage in the excitatory dynamics of universal consciousness—as a primary causal phenomenon inherent to the possible behaviors of nature's sole ontological primitive. In other words, dissociation is thought to explain life and the world, as opposed to being explained by them.

But the formulation presented Chapter 3 is, in principle, not the only one that could be woven around this ground-level notion of dissociation. One could, for instance, conceive of an alternative formulation based on the observation that regular dream images are directly generated by our dreaming psyche—through self-excitation—*already in the very form that they are experienced*. In other words, dream images aren't coded phenomenal representations of some other phenomenal dynamics; they aren't extrinsic appearances of qualitatively different intrinsic views. Instead, they are a self-contained movie directly and autonomously generated by our dreaming psyche. So an alternative formulation of idealism could be this: instead of thinking of the inanimate universe as the extrinsic appearance of the states of a transpersonal phenomenal field, the corresponding images could be generated at a level of universal consciousness prior to, or underlying, dissociation *already in the form we experience them*.

To visualize this, imagine that alters of universal consciousness—that is, living organisms such as you and me—are analogous to the seemingly separate branches of a shrub, which ultimately come together at the hidden rhizome. This way, dissociation is the process that creates branches by seemingly separating segments of the shrub. But this process operates somewhat superficially, in that it doesn't affect the unitary rhizome. In this analogy, the inanimate universe we all seem to co-inhabit is a collective dream generated by the rhizome and then broadcast—after some perspectival filtering and adaptation—to all branches already in the form it is experienced. The Jungian notion of a 'collective unconscious' (Jung 1991) capable of producing archetypal dreams fits nicely with the hypothesis I am trying to describe here: according to it, the world is the waking dream generated by the 'collective unconscious' and then broadcast to each of our individual psyches.

This isn't as far-fetched as it may sound at first. Indeed, as discussed in Chapter 5, there is an empirically known form of dissociation according to which subjects lose the sense of ownership of their own phenomenal states (Klein 2015). In this context, it is not unreasonable to imagine that empirical reality is a collective stream of imagination that we lose our sense of ownership of, thereby mistakenly concluding that it corresponds to a world outside and independent of consciousness.

Nonetheless, such a seemingly elegant formulation of idealism fails because, whereas it can parsimoniously explain the inanimate universe, *it cannot explain the presence of other conscious organisms in it*. If the collective dream we call 'the world' were broadcast from the rhizome to the individual branches, why or how would one branch experience the presence of other branches—that is, other people and living organisms—in its dream? After all, two TV receivers tuned to the same channel can display the same movie, *but not images of each other within that movie*.

For the same reason that this alternative formulation of idealism does away with a transpersonal phenomenal field surrounding the alters, it must also do away with the conscious inner life of other living organisms. This effectively

reduces it to solipsism and renders redundant the very need to explain a shared world to begin with.

The key difference between this alternative formulation and the analytic idealism elaborated in Chapter 3 is this: the former entails that the images on our personal screen of perception are themselves irreducible. The latter, on the other hand, posits that personal perceptions are grounded in the phenomenal states of a transpersonal field, which in turn are *qualitatively different from personal perceptions*.

Bishop Berkeley's formulation of idealism is similar to the alternative discussed here, insofar as it also entails that personal perceptions are irreducible. As explained by Barfield, "Berkeley held that … the representations *as such* [that is, personal perceptions], are sustained by God in the absence of human beings" (2011: 36, original emphasis). Even some present-day academic philosophers continue to entertain the idea that the contents or qualities of personal perception are themselves irreducible: "In perception, our finite unities of consciousness come to literally overlap with the unity of consciousness that is reality" (Yetter-Chappell 2018). So the phenomenal states one experiences when one sees the world are supposedly *the same phenomenal states* encompassed by "the unit of consciousness that is reality," with which one overlaps during the act of perceiving.

I believe that all formulations of idealism entailing such irreducibility of personal perceptions fail, either because of the difficulties discussed in the previous section or at least because of the consequences of trying to circumvent these difficulties.

7.3 There are noumena, but they are experiential

By maintaining that personal perceptions are partially grounded in something outside the personal self—that is, outside alters of universal consciousness—I am positing something at least analogous to what Barfield called the "unrepresented" (2011) and Kant the "noumenal." Indeed, I am maintaining that *there is a shared reality beyond the alters*—namely, the experiential states beyond our respective dissociative boundaries—underlying our personal perceptions of the world. This shared reality would still exist even if we and all other living beings ceased to be. So the noumenal does exist, my point being simply that its essential nature is experiential; the noumenal itself consists of experiences, even though these experiences are qualitatively different from personal perceptions. In summary, according to the ontology defended in this dissertation, *there are noumena but they are experiential*.[1]

[1] Here I am deliberately avoiding the word 'phenomenal'—using the qualifier 'experiential' instead—to avoid confusion: Kant uses the term 'phenomenal' exclusively in connection with intentional content, whereas I use it, throughout this dissertation, in its modern analytic sense. According to Ned Block, for instance, "*Phenomenal* consciousness is *experience*; the *phenomenally* conscious aspect of a

Indeed, Kant and Barfield expressly did *not* specify the ontological character of the noumenal and the unrepresented, respectively, so neither is necessarily dichotomous with the experiential. The experiential states that make up the noumena are not perceptual states of alters, but constitute a transpersonal qualitative field surrounding the alters.

7.4 The conundrum of spacetime

Space and time are built into language: any statement about what nature is or how it works presupposes a spacetime scaffolding. Without extension in at least one dimension, the various states of nature would overlap and become indistinguishable from one another. Information about nature—as defined by Shannon (1948)—would thus vanish and there would remain literally nothing to be said about it.

My earlier analogy between experiences and vibrations of consciousness seems to also presuppose a spacetime scaffolding circumscribing consciousness. After all, vibrations entail some form of movement in space and time (think of a guitar string playing a musical note: it moves up and down as time passes). So it could be argued that analytic idealism, in addition to consciousness itself, assumes a spacetime scaffolding as extra ontological primitive, wherein consciousness can then 'move' so to have or produce particular experiences. But this would contradict my core claim that universal phenomenal consciousness is the sole ontological primitive.

For this core claim to obtain, both space and time must, instead, be nothing more than *qualities of experience*. Time must exist only insofar as what we call 'past' is an experiential quality characteristic of memory and 'future' an experiential quality characteristic of imagined possibilities or expectations. Space, in turn, must exist only insofar as it is the experiential quality of a certain *relationship* between perceived objects. This way, spacetime must be only an amalgamation of qualities—amenable to mathematical modeling—that themselves exist only *in* universal consciousness. This is only plausible if (a) physics remains viable without a fundamental spacetime scaffolding and (b) the felt sense of temporal flow associated with experience is an illusion.

Starting with (a), Einstein's relativity theory arguably implies a static 'block universe' wherein the passage of time is illusory. This realization has motivated attempts by physicists to recast the laws of physics without time (Barbour 1999, Rovelli 2018) or space (Smolin 2013). Now, modern quantum gravity theories posit that *both* space and time—spacetime—are emergent from more fundamental quantum processes (Crowther 2014). Clearly, thus, physics remains viable without the postulate that the spacetime scaffolding is fundamental; it may even *require* that it not be so.

---

state is *what it is like to be* in that state" (1995: 227, emphasis added). According to this modern definition—and unlike Kant's usage—even purely endogenous experiences, with no intentional content, are phenomenal states.

However, this may seem to pose a problem for analytic idealism, as argued by philosopher Susan Schneider:

> Suppose that our ordinary sense of duration is just an illusion, and reality is timeless. If this is the case, the point shouldn't be that the fundamental layer of reality is experiential. The point should be, instead, that fundamental reality is nonexperiential.[2]

The conclusion is derived from Schneider's intuition that,

> if there is no time ... how could there be experience? Conscious experience has a felt quality that involves flow; thoughts seem to be present in the "now," and they change from moment to moment. Timeless experience is an oxymoron.[3]

So if time is not fundamental, then neither can experience be—or so her argument goes.

To secure the plausibility of analytic idealism I must, therefore, show that Schneider's intuition here is flawed. Indeed, whereas *demonstrating* that spacetime is an illusion would likely require multiple doctoral dissertations across a variety of fields, I believe it is possible to show—rather easily—that, *insofar as Schneider is appealing to phenomenal introspection*, her assertion that "timeless experience is an oxymoron" is false.[4]

To see it, consider these questions: Where's the past? Is it anywhere 'out there'? Can you point at it? Clearly not. What makes you conceive of the idea of a past is the fact that you have episodic memories. But these memories can only be referenced insofar as they are experienced *now*, *as memories*. There has never been a point in your entire life in which the past has been anything more than memories experienced *now*.

The same applies to the future: Where's the future? Is it anywhere 'out there'? Can you point at it and say "there is the future"? Clearly not. Our idea of a future arises from expectations and imaginings experienced *now*, always now, *as expectations and imaginings*. There has never been a point in your life in which the future has been anything more than expectations and imaginings experienced *now*.

Therefore, as far as careful phenomenal introspection can reveal, experience lacks any true temporal flow. It only ever happens *now*. Timelessness seems to be precisely an intrinsic property of experience. Time, on the other hand, seems

---

[2] See Schneider's essay on *Scientific American*'s *Observations* blog at: https://blogs.scientificamerican.com/observations/spacetime-emergence-panpsychism-and-the-nature-of-consciousness/.

[3] *ibid.*

[4] I have published a slightly more elaborate version of the argument that follows also on *Scientific American's Observations* blog, which is available at: https://blogs.scientificamerican.com/observations/do-we-actually-experience-the-flow-of-time/.

to be merely *a particular phenomenal state* experienced timelessly; a cognitive construct or story we tell ourselves *now*, always now. As a matter of fact, there is compelling empirical evidence supporting the notion that time is indeed a cognitive construct (e.g. Buonomano 2018, Eagleman 2009).

To avoid confusion, allow me to explicitly relate these ideas to the ontological theses of presentism (the notion that only present things exists, not the past or the future) and eternalism (the notion that past, present and future things all exist): with presentism, my claim is that only the *now* can be known to exist; there is nothing other than the now insofar as we can introspectively access; past and future, *as ontological entities outside the now,* are merely theoretical abstractions. However, with eternalism, I also claim that the experiential contents we label as 'past' and 'future'—and which motivate the theoretical abstractions of a past and a future ontologically distinct from the now—*are as real as what we call the 'present,'* in that *all three exist solely as phenomenal states experienced in the now.* In this latter sense, past, present and future are ontologically equivalent. Indeed, the partitioning of the salient conceptual space between presentism and eternalism is not the most appropriate for the ideas I am attempting to convey here, and hence should be regarded with caution.

To sum it up, it seems to be the case that (a) physics remains entirely viable without a fundamental spacetime scaffolding and (b) the felt sense of temporal flow associated with experience is—at least insofar as we can assess through careful introspection—an illusion, a story we tell ourselves timelessly. The plausibility of analytic idealism is thus preserved: there is currently no refutation of the notion that spacetime is merely an amalgamation of qualities of experience. On the contrary: there are tantalizing signs that spacetime emerges as a cognitive construct within consciousness.

But in this case, I must somehow reconcile the hypothesis that spacetime isn't fundamental with the earlier analogy between experiences and vibrations of consciousness. This is what I now set out to do.

To begin with, the analogy must be regarded solely as such: as *an analogy*. This way, experiences are *like* vibrations of consciousness. The intent of the analogy is to help one visualize how various experiences can be distinct from each other without requiring that there be anything to them but consciousness itself. As a matter of fact, I defined experiences as *excitations*—as opposed to outright vibrations—of consciousness in the hope that the term 'excitation' wouldn't commit me as much to dimensional extension.

The problem is that, if one wants to *talk* about the nature of reality, one must presuppose a *metaphorical* spacetime scaffolding. This is an unavoidable *concession* to the limitations of language.[5] Nonetheless, acknowledging that

[5] It could be argued here that Kant wrote about the noumena, the things-in-themselves, without using spatiotemporal terms. Notice, however, that Kant merely postulated the *existence* of the noumena—merely *pointed* to them—whilst insisting, at the same time, that they are fundamentally unknowable and cannot be

this dimensional scaffolding is simply a kind of illusion inherently imposed by the structure of our cognition doesn't change the practical problem at hand: whatever reality precedes spacetime ontologically is unreachable by linguistic reasoning. At best, one can articulate *projections* of this otherwise ineffable reality onto the cognitive scaffolding of spacetime.

Here is an analogy to illustrate what I mean: one cannot read a letter written in a piece of paper that has been folded multiple times over into a small, nearly dimensionless crumple. The characters overlap and the information they contain becomes indiscernible. Only by unfolding the crumple—i.e. *extending* the piece of paper—can one make sense of the message it bears. Reality prior to spacetime is, in a sense, like the paper crumple: one needs to unfold it along the dimensions of space and time to render it amenable to linguistic articulation.

Does this mean that spacetime-bound language can never articulate valid and meaningful conclusions? No. All it means is that the conceptual systems built within the framework of spacetime cannot be *ultimately* true. After all, *ex hypothesi*, spacetime is merely a cognitive construction. However, those conceptual systems can still be *penultimately* true in the sense that they can accurately *correspond to* something ontologically prior to dimensionality. Valid spacetime-bound conclusions can thus be regarded as *projected images* of ultimate truths, adapted to the requirements and limitations of human reasoning by dimensional extension.

This way, to say that experiences are vibrations of universal phenomenal consciousness admittedly cannot be ultimately true, for consciousness—as sole ontological primitive—does not occupy a spacetime scaffolding prior to itself. But it can still be true in the penultimate sense that vibrations *correspond to* something true—though ineffable—about consciousness prior to dimensional extension; that vibrations are akin to an accurate *projected image* of what ultimately happens in consciousness when it experiences.

That one cannot *directly* say something coherent about an ultimate truth does not invalidate penultimate conceptual constructs. They can still tell one something *in*directly true about what reality is and how works.

---

characterized. He also did not approach the problem of explaining how perception arises from the noumena. If Kant had made any attempt to say something *about* the noumena, he would have immediately faced the linguistic limitations indicated here, and would have had to adopt spatiotemporal terms. As a matter of fact, in an incoherent move already criticized by Arthur Schopenhauer in *The World as Will and Representation* (1818), Kant talks of the noumena as *causes* of perception. Causality, of course, presupposes spacetime extension, so here we have Kant implicitly characterizing the noumena in spatiotemporal terms, even though they are supposed to exist outside spacetime.

7.5 Visualizing the ontological primitive

As a matter of fact, *space itself* can be coherently regarded as the quality of human experience that corresponds most closely to universal phenomenal consciousness, different segments of the latter corresponding to different regions of space. Indeed, that two living organisms—the extrinsic appearances of alters—never occupy the same volume of space reflects the notion that different alters are located in different segments of universal consciousness, as discussed in Chapter 3.

Moreover, that we think of empty space as a void, a nothing, reflects the notion that unexcited universal consciousness cannot, by definition, be experienced. Even the idea that unexcited universal consciousness must still have intrinsic properties—otherwise there would be nothing to eventually *get* excited—finds a correspondence in how we think of space at least since the early 20th century: empty space, too, is believed to be a void with intrinsic properties.

The correspondences here are clear: there is a strong sense in which, as far as human cognition is concerned, empty space *is* universal consciousness, the contents of space being excitations of universal consciousness. Moreover, since space is simply a facet of spacetime, I suggest that it is closer to the truth to think of spacetime *as* universal consciousness than as a scaffolding *occupied by* universal consciousness.

If one is to make and talk about philosophy, it is unavoidable to frame one's thoughts and discourse in terms of spacetime extension. For the true analytic idealist, this is admittedly a concession, for spacetime is not in the idealist's reduction base. The assertions made should thus not be regarded as ultimate. But they are still true and meaningful as far as they go.

7.6 Future work

This dissertation is perhaps a first step in reviving idealism in the context of contemporary analytic philosophy. As such, it opens up more avenues of inquiry than it closes. Here, I shall mention two important topics that I believe deserve more thorough analysis.

The first is the philosophical understanding of dissociative processes in the human psyche and the application of Dissociative Identity Disorder (DID) as an analogy for universal-level dissociation. An often-repeated criticism against analytic idealism is, for instance, that although alters of universal consciousness can clearly experience one another from a second-person perspective—the way a person can see and even shake hands with another—the same allegedly isn't the case for the alters of a DID patient. In other words, there allegedly isn't anything a first dissociated personality of a DID patient looks like from the perspective of a second dissociated personality of the same patient.

The most straightforward answer to this criticism is, of course, that DID is merely an *analogy*: the claim is not that universal consciousness literally suffers from DID, but that something *like* DID happens at a universal level. There does

not need to be a complete correspondence between human-level DID and universal-level dissociation for the argument to be valid.

Having said that, there are tantalizing indications that the correspondence is more far-reaching than hitherto suggested in this dissertation. For instance, research has shown that different alters of a DID patient can—and do—appear as characters in the dreams of the patient (Barrett 1994: 170-171). So there actually *is* something other dissociated personalities look like from the point of view of the host personality having the dream.

More significantly, the same research has also shown that *different alters of a DID patient can experience the same dream concurrently, each from its own subjective point of view within the dream*. This is so significant that one illustrative example deserves extensive quoting:

> The host personality, Sarah, remembered only that her dream from the previous night involved hearing a girl screaming for help. Alter Annie, age four, remembered a nightmare of being tied down naked and unable to cry out as a man began to cut her vagina. Ann, age nine, dreamed of watching this scene and screaming desperately for help (apparently the voice in the host's dream). Teenage Jo dreamed of coming upon this scene and clubbing the little girl's attacker over the head; in her dream he fell to the ground dead and she left. In the dreams of Ann and Annie, the teenager with the club appeared, struck the man to the ground but he arose and renewed his attack again. Four year old Sally dreamed of playing with her dolls happily and nothing else. Both Annie and Ann reported a little girl playing obliviously in the corner of the room in their dreams. Although there was no definite abuser-identified alter manifesting at this time, the presence at times of a hallucinated voice similar to Sarah's uncle suggested there might be yet another alter experiencing the dream from the attacker's vantage. (*ibid.*: 171)

Taking this at face value for the sake of argument, what it seems to show is that, while dreaming, a dissociated human psyche can manifest multiple, *concurrently conscious* alters that experience each other from a second-person perspective, just as a person sees and shakes hands with another in waking life. The alters' experiences are also mutually consistent, in the sense that the alters all seem to experience the same series of events, each from its own subjective perspective. The correspondence with what is argued to happen in the case of universal-level dissociation is uncanny.

One may nonetheless object to this correspondence by pointing out that the alters of a DID patient can only experience one another from a second-person perspective if the host is in a particular state of consciousness—namely, dreaming. But notice that, since universal consciousness is, *ex hypothesi*, all there is, there is no 'outside world' at its level. So the only state of consciousness conceivably available to it is one entailing self-generated experiences analogous to dreaming. Might we then be alters of universal consciousness experiencing one another within a universal 'dream,' just as the alters of a DID patient experience one another during the patient's dream?

It is important to point out, however, that, despite decades of clinical work, research on DID is still in its infancy. Indeed, only in the past few years has the existence of DID been scientifically established. Our understanding of the psychic processes underlying dissociation is still fairly limited; particularly when it comes to the dream-life of DID patients. As research on human-level dissociation advances, we should be in a better position to understand the true extent to which DID serves as an analogy for universal-level dissociation.

The second topic of future work that deserves attention is our understanding of meta-consciousness and its modeling under the excitation metaphor of experience. I've described this metaphor in Chapter 3: experiences can be regarded as self-excitations of universal phenomenal consciousness. I've also claimed in Chapter 5 that meta-consciousness reflects a particular *configuration of* consciousness. Bringing these two metaphors together is certainly feasible: if we imagine universal phenomenal consciousness as a membrane—just as M-theorists imagine their 'branes'—experiences can be regarded as vibrations of this membrane and meta-consciousness as a particular topological configuration of the membrane. More specifically, meta-consciousness can be visualized as the membrane folding in on itself, so that the patterns of vibration in a first segment of the membrane induce corresponding vibrations in a second segment folded on top of the first. The patterns of vibration in the first segment constitute a raw experience, whereas the patterns of vibration in the second segment constitute the meta-conscious realization that one is having the raw experience.

However, since there are many nuances inherent to the notion of meta-consciousness that were not explored in this dissertation, more work is required here. For instance, does meta-consciousness presuppose, entail or imply a Heideggerian sense of a separate self inhabiting the world? If so, how could this sense be accommodated by the topological analogy discussed above? What are the qualitative differences between a raw experience and the meta-conscious realization that one is having the raw experience? How do these differences arise and how can the corresponding mental processes be modeled according to the topological metaphor? And so forth.

It remains my hope that this dissertation motivates analytic philosophers to reconsider idealism not only as a viable, but perhaps even the most promising, avenue for circumventing the fundamental problems faced today in ontology. I hope the ideas presented here are but the seeds for a wave of new developments in the coming years and decades, which will pursue with more depth and detail the many avenues of reasoning opened up by this dissertation.

Appendix A. Not Its Own Meaning: A Hermeneutic of the World

This paper first appeared in *Humanities*, ISSN: 2076-0787, Vol. 6, No. 3, Article No. 55, in August 2017.

A.1 Abstract

The contemporary cultural mindset posits that the world has no intrinsic semantic value. The meaning we see in it is supposedly projected onto the world by ourselves. Underpinning this view is the mainstream physicalist ontology, according to which mind is an emergent property or epiphenomenon of brains. As such, since the world beyond brains isn't mental, it cannot *a priori* evoke anything beyond itself. But a consistent series of recent experimental results suggests strongly that the world may in fact be mental in nature, a hypothesis openly discussed in the field of foundations of physics. In this essay, these experimental results are reviewed and their hermeneutic implications discussed. If the world is mental, it points to something beyond its face-value appearances and is amenable to interpretation, just as ordinary dreams. In this case, the project of a Hermeneutic of Everything is metaphysically justifiable.

A.2 Introduction

To be amenable to interpretation, things and phenomena must point beyond themselves, thereby embodying semantic value or sense. For instance, these squiggles of ink on paper—which we call written words—mean more than just squiggles of ink on paper: they point to something beyond themselves. Similarly, the inner imagery we experience in dreams points to something beyond their face-value appearances, which has motivated depth psychologists to develop extensive hermeneutics of dreams (e.g. Ackroyd 1993, von Franz and Boa 1994, Jung 2002, Fonagy *et al*. 2012). Finally, the symbolisms of religious myths point to something that transcends the face-value appearances of the symbols themselves and engages people at an emotional level (Kastrup 2016a).

Influenced by twentieth century positivism and existentialism, the contemporary cultural mindset posits that things and phenomena only have semantic value insofar as we project this value onto them. Summarizing the essence of this mindset, Sartre wrote: "there exist concretely alarm clocks … But … then I discover myself suddenly as the one who gives its meaning to the alarm clock … the one who finally makes the values exist" (1992: 77). Analogously, squiggles of ink mean more than squiggles of ink only insofar as we stipulate by convention that they do so. To the extent that alarm clocks and written words are inventions of human beings, it is reasonable to assert that their meaning consists in what we project onto them.

However, the contemporary cultural mindset extends this notion of projected meaning to nature itself. Fire only represents "the inseminating fury of sex and the ardor of the ascetic" (Ronnenberg and Martin 2010: 84) insofar as we project passion onto it. Stones only represent eternity (*ibid*.: 106) insofar as we project timelessness onto them. Without our projections, stones mean just stones; fire means just fire. In and of itself, the world supposedly is its own meaning. It does not inherently point to anything beyond its own appearance on the screen of perception. Whatever sense we may see in a fact of the world is supposedly a confabulation of human cognition, not intrinsic to the fact itself. "In this case," as Zemach put it, "one may say either that this fact has no sense, or that the only sense it has is provided by its form" (2006: 363). In other words, "The sense of the world is identical with its form" (*ibid*.: 367). Ortiz-Osés put it perhaps most simply: "When taken 'existentially,' existence seems to lack sense, whereas sense taken 'essentially' would appear to lack existence" (2008: 65).

As a result, our culture believes that the semantic value of the world is simply an artifact of human minds. The world doesn't have a story to tell, a suggestion to make or an insight to convey. It isn't saying anything. There is nothing meaningful to be gleaned from the world, just utilitarian predictions to be made about its behavior. Under such ethos, projects such as Ortiz-Osés'—meant to formulate a symbolic hermeneutic of the world premised on the notion that "the whole of existence contains an almost secret essence" (2008: 1)—become metaphysically precarious, which Ortiz-Osés himself seems to have acknowledged (*ibid*.: 65).

At the root of this state of affairs is the split between mind and world that characterizes our present worldview. Indeed, according to the mainstream physicalist ontology, the fundamental building blocks of reality are physical elements that exist independently of mind (Stoljar 2016). The latter, in turn, is supposedly constituted or generated by particular local arrangements of these physical elements, such as brains inside skulls. Consequently, mind is insulated from the external world surrounding it beyond the skull.

The problem, of course, is that only mind can host intrinsic semantic value, for the latter consists of cognitive associations: the intrinsic meaning of an experience is the emotions, insights and inner imagery it evokes. For instance, the feeling of hunger may evoke inner imagery related to food because there is a cognitive association between the feeling and the imagery. A memory from childhood may evoke the emotion of happiness because there is a cognitive association between the memory and the emotion. These associative links are an exclusive feature of mentation.

So if semantic value is essentially mental and mind is insulated from the world beyond the skull, then semantic value cannot exist in the world. A non-mental world can *be* evoked, but it cannot intrinsically *evoke* anything. Such separation between meaning and world is what motivates our contemporary culture to consider the world semantically mute. "The human mind has abstracted from the whole all … meaning, and claimed [it] exclusively for itself," wrote Tarnas (2010: 432).

Within mind, cognitive associations can go on indefinitely, as endless chains of evocations: a daydream may lead to a thought, which may evoke an emotion, which may trigger a memory, which may lead to another thought, and so on (Karunamuni 2015: 2-3). But once we leave the inner space of mentation by evoking an external fact in the world, the chain must end. The world is the chain's final destination, for it cannot *a priori* evoke anything else in turn. This semantic end point is what we call a 'literal fact.' Everything prior to it is sign, simile or allegory—roundabout, indirect ways to arrive at the destination. According to our contemporary cultural mindset, the value of these indirections is entirely conditioned upon their ability to ultimately point at literal facts. Anything short of it is considered delusion, for it allegedly can't be anchored in truth.

But does our current scientific understanding of reality truly corroborate this split between mind and world, inside and outside? Are we justified in taking for granted that the world 'out there' is fundamentally distinct or separate from the mind 'in here'? If not, could the world carry intrinsic semantic value and be amenable to interpretation, just as dreams are? Could there be a valid hermeneutic of the world, a vision of it as symbolic, suggestive of something beyond its own face-value appearances on the screen of perception? What would the implications of this possibility be for the way we relate to the world? These are the questions addressed in this essay.

In Section A.3, the latest experimental results emerging from the field of quantum physics will be briefly reviewed. They empirically indicate that mind and world aren't, after all, fundamentally distinct or separate. Section A.4 will show how this continuity between mind and world can explain why the axioms of rational thought describe and model the world so uncannily accurately. In Section A.5, the hermeneutic implications of the mental world hypothesis will be discussed. Section A.6 then compares the analysis in Section A.5 with what some of the world's philosophical and spiritual traditions have to say about the nature and meaning of the world. Finally, Section A.7 concludes this essay with a brief discussion.

A.3 The ontological status of the world

The mainstream physicalist notion that the world is outside and independent of mind is an abstract explanatory model constructed in thought, not an empirical observation. After all, what we call 'the world' is available to us solely as 'images'—defined here broadly, so to include any sensory modality—on the screen of perception, which is itself mental. We *interpret* the contents of perception as coming from a world outside mind because this seems to explain the fact that we all share the same world beyond the boundary of our skin, as well as the fact that the laws that govern this world do not depend on our personal volition. Stanford physicist Prof. Andrei Linde, well known for his theories of cosmological inflation, summarized it thus:

> Let us remember that our knowledge of the world begins not with matter but with perceptions. I know for sure that my pain exists, my "green" exists, and my "sweet" exists. I do not need any proof of their existence, because these events are a part of me; everything else is a theory. Later we find out that our perceptions obey some laws, which can be most conveniently formulated if we assume that there is some underlying reality beyond our perceptions. This model of material world obeying laws of physics is so successful that soon we forget about our starting point and say that matter is the only reality, and perceptions are only helpful for its description. This assumption is almost as natural (and maybe as false) as our previous assumption that space is only a mathematical tool for the description of matter. But in fact we are substituting reality of our feelings by a successfully working theory of an independently existing material world. And the theory is so successful that we almost never think about its limitations until we must address some really deep issues, which do not fit into our model of reality. (1998: 12)

This model of reality has intuitive implications amenable to confirmation—or refutation—through subtle experimental arrangements, which Linde alluded to when he spoke of "some really deep issues." Indeed, the properties of a physicalist world should exist and have definite values even when this world is not being observed: the moon should exist and have whatever weight, shape, size and color it has even if nobody is looking at it. Moreover, a mere act of observation should not change the values of these properties: the weight, shape, size and color of the moon should not become different simply because someone happened to look at it.

Operationally, these intuitive tenets of physicalism are translated into the notion of 'non-contextuality': the outcome of an observation should not depend on the way other, separate but simultaneous observations are performed. After all, the properties being observed are supposed to be independent of observation. What I perceive when I look at the night sky should not depend on the way other people look at the night sky along with me, for the properties of the night sky uncovered by my observation should not depend on theirs. Clearly—and in line with physicalism—non-contextuality implies that the world is independent of perception, insofar as perception constitutes observation. My perceptions should simply *reveal* what the properties of the world are in and of themselves.

The problem is that, according to quantum theory, the outcome of an observation *can* depend on the way another, separate but simultaneous observation is performed. For instance, if two particles A and B are prepared in a special way, the properties of particle A as seen by a first observer—say, Alice—are predicted to correlate with the way another observer—say, Bob—simultaneously looks at particle B. This is so even when A and B—and, therefore, Alice and Bob—are separated by arbitrarily long distances. For instance, what Alice sees when she looks at particle A in, say, London, depends on the way Bob concurrently looks at particle B in, say, Sydney. If the properties

of the world were outside and independent of Alice's and Bob's minds—that is, outside and independent of their perceptions—this clearly shouldn't be the case; unless there is some observation-independent hidden property, covertly shared by A and B *and entirely missed by quantum theory*, which could account for the correlations. This was Einstein's point when he (in)famously suggested that quantum theory was incomplete (Einstein, Podolsky and Rosen 1935). However, as mathematically proven by John Bell (1964), the correlations predicted by quantum theory cannot be accounted for by these kinds of observation-independent hidden properties.

Consequently, quantum theory appears to contradict non-contextuality and render physicalism untenable. A conceivable way to avoid this conclusion while accepting quantum theory would be to posit that particles A and B, or Alice and Bob themselves, somehow 'tip each other off' during observation, *instantaneously and at a distance*, so to coordinate their actions and produce the predicted correlations. This, however, would require faster-than-light communication and fly in the face of the overwhelmingly confirmed theory of special relativity.

Alternatively, a physicalist could attempt to salvage non-contextuality and the notion of a world outside and independent of mind by rejecting quantum theory itself. Yet, as it turns out, since Alain Aspect's seminal experiments (Aspect, Grangier and Roger 1981, Aspect, Dalibard and Roger 1982, Aspect, Grangier and Roger 1982) the predictions of quantum theory in this regard have been repeatedly confirmed, with ever-increasing rigor. For instance, in an experiment performed in Geneva, Switzerland, in 1998 (Tittel *et al.*), the particles A and B were separated by more than 10 km—as opposed to the 12 meters of Aspect's original experiment (1981)—reducing the already low likelihood that they could be creating the correlations predicted by quantum theory through some kind of signal exchange. Despite this greater separation, the predictions of quantum theory were again confirmed.

Then, still in 1998 but this time in Innsbruck, Austria, another experiment (Weihs *et al.*) was done to eliminate another far-fetched possibility: that, *in advance* of the preparation of particles A and B, 'Alice,' 'Bob' and the system responsible for the preparation could somehow be 'pre-agreeing' on a hidden plan of action, so to later create the correlations without need for faster-than-light communication ('Alice' and 'Bob,' in this case, were automated measurement apparatuses). To close this unlikely 'conspiracy' loophole, the behaviors of 'Alice' and 'Bob' were programmed randomly and only *after* particles A and B had already been prepared. Nonetheless, the correlations predicted by quantum theory were yet again confirmed.

Critics continued to speculate about other far-fetched loopholes in these experiments. In an effort to address and close all conceivable loopholes, Dutch researchers have recently performed an even more tightly controlled test, which—unsurprisingly by now—echoed the earlier results (Hensen *et al.* 2015). This latter effort was considered by the periodical *Nature* the "toughest test yet"

(Merali 2015). Given all this, it seems now untenable to argue against the veracity of quantum theory.

The only alternative left for physicalists is to try to circumvent the need for faster-than-light signal exchanges by imagining and postulating some form of non-locality: nature must have—or so they speculate—observation-independent hidden properties that are *not* confined to particular regions of spacetime, such as particles A and B. In other words, the argument is that the observation-independent hidden properties allegedly missed by quantum theory are 'smeared out' across space and time. It is this omnipresent, invisible but objective background that supposedly orchestrates the correlations predicted by quantum mechanics. Non-contextuality and physicalism can thus be salvaged; or can they?

The problem, of course, is that non-local hidden properties are arbitrary: they produce no predictions beyond those already made by standard quantum theory. As such, it could be argued that they represent an effort "to modify quantum mechanics to make it consistent with [one's] view of the world," so to avoid the need "to modify [one's] view of the world to make it consistent with quantum mechanics" (Rovelli 2008: 16).

Be it as it may, it turns out that certain specific correlations predicted by quantum theory are incompatible with non-contextuality *even for large classes of non-local hidden properties* (Leggett 2003). Studies have now experimentally confirmed these correlations (Gröblacher *et al*. 2007, Romero *et al*. 2010), thus putting non-contextuality in even more serious jeopardy. To reconcile these results with physicalism would require a profoundly counterintuitive redefinition of what we call 'objectivity.' And since our contemporary cultural mindset has come to associate objectivity with reality itself, the science press felt compelled to report on some of these results by pronouncing, "Quantum physics says goodbye to reality" (Cartwright 2007).

More recent experiments have again contradicted non-contextuality and confirmed that, unlike what one would expect if the world were separate or distinct from mind, the observed properties of the world indeed cannot be said to exist prior to being observed (Lapkiewicz *et al*. 2011, Manning *et al*. 2015). For all intents and purposes, the world we perceive is *a product of observation*. Commenting on this, physicist Anton Zeilinger has been quoted as saying that "there is no sense in assuming that what we do not measure [that is, observe] about a system has [an independent] reality" (Ananthaswamy 2011).

So the question now is: Can some form of physicalism survive the failure of non-contextuality? We have seen earlier that the intuitive tenets of physicalism are: (a) there exists a world outside mind; and (b) mere observation doesn't change this independently existing world. The failure of non-contextuality clearly rules out (b). Can (a) still make any sense in the absence of (b)? If it can, then the world outside mind must somehow *physically change, instantaneously*, every time it is observed. The plausibility of this notion aside, notice that one never gets to see the observation-independent world, for it supposedly changes instantly, in an *observation-dependent* manner, the moment one looks at it.

Clearly, the only motivation to entertain this notion is to try to salvage some rather artificial and counterintuitive form of physicalism. And even if such an attempt were to succeed, the world we actually experience would *still* be conditioned by mind, insofar as it would be an outcome of conscious perception. For the purposes of this paper, therefore, the result would be indistinguishable from a truly mental world.

Already in 2005, Johns Hopkins physicist and astronomer Prof. Richard Conn Henry had seen enough. In an essay he penned for *Nature*, he claimed, "The universe is entirely mental. ... There have been serious [theoretical] attempts to preserve a material world—but they produce no new physics, and serve only to preserve an illusion" (Henry 2005: 29). The illusion he was referring to was, of course, that of a world outside mind.

Naturally, Conn Henry's position is controversial and debate around it continues to unfold. Nonetheless, the experiments do show that the idea of a mental world must be taken seriously, if nothing else for the sheer power of the empirical evidence now accumulated. Moreover, philosophers have recently proposed coherent ontologies that can, at least in principle, make sense of reality without the need to postulate anything distinct from mind itself (Kastrup 2017e, Nagasawa and Wager 2016, Shani 2015). These ontologies provide coherent frameworks in which the experimental results can be placed and interpreted.

Finally, notice that, although the argument in this section has been based on quantum mechanical experiments carried out on microscopic particles under laboratory conditions, we know that the implications of quantum theory apply to our macroscopic world of tables and chairs as well. Indeed, quantum effects have been experimentally demonstrated for macroscopic objects at room temperature (Lee *et al*. 2011, Klimov *et al*. 2015). As such, the failure of non-contextuality indicates that the seemingly objective world we live in is a result of mental process at work and, as such, akin to a transpersonal dream: the tables, chairs, stars and galaxies we perceive within it do not have an existence independent of our minds.

A.4 The continuity of mind and world

In a famous paper titled "The Unreasonable Effectiveness of Mathematics in the Natural Sciences," physicist Eugene Wigner (1960) discussed "the miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics." Indeed, abstract methods and results developed purely in thought have, again and again, succeeded in precisely describing concrete phenomena. That axiomatic intuitions turn out to correctly predict and model the structure and dynamics of the world at large is difficult to make sense of under physicalism, this probably being the reason why Wigner used the word 'miracle' twelve times in his paper. After all, lest we incur the fallacy of circular reasoning, under physicalism we cannot logically argue for the validity of logic

beyond our own minds, so the world could very well be absurd (Albert 1985). That it is not is Wigner's "miracle."

If the world is mental, however, the correspondence between the intuitive foundations of rational thought and the way the world works is perfectly natural. That we take the basic tenets of logic and mathematics to be self-evident truths betrays their *archetypal nature* in the Jungian sense: they reflect deeply ingrained mental templates according to which thought unfolds (Jung 1991). As a matter of fact, psychologist Marie-Louise von Franz went as far as to argue that the natural numbers themselves are archetypal (1974). Then—and here is the key point—*the fact that these archetypes extend into the world clearly indicates that the world itself is mental and continuous with our minds*. If there is no intrinsic separation between our minds and the objects of perception, naturally these objects should comport themselves in a way consistent with mental archetypes. Perceptual objects should be an expression of archetypal patterns in just the same way that thoughts are, so the world should be consistent—as it is—with our logic and mathematics. The apparent eeriness of Wigner's "miracle" melts away.

To visualize all this consider the following analogy: if mind is like a guitar string, then particular conscious experiences are like particular notes or patterns of vibration of the string. In this case, the mental archetypes discussed above are analogous to the elasticity, mass and length of the string, which determine its normal modes of vibration. Some of the archetypically-defined normal modes of mind thus correspond to the laws of nature, which we discern as regularities on the screen of perception: they reflect some of the 'notes' in which mind naturally 'plays' in the world at large.

Wigner's "miracle" is not only explainable by, but also constitutes further evidence for, the mental world hypothesis. As such, it is high time we considered the implications of this hypothesis for how best to live our lives.

A.5 The implications of a mental world

Strong empirical evidence pointing to the conclusion that the world we experience is a result of transpersonal mental processes at work has now been reviewed. There is no fundamental separation between mind 'in here' and world 'out there,' which explains why the archetypes of rational thought describe nature so well. Yet, the latter point is not the sole implication of a mental world: if our minds are continuous with the environment we inhabit, nothing prevents the world from *intrinsically* evoking mental contents beyond perception, such as insights and emotions.

Indeed, according to analytical psychology, our nightly dreams carry intrinsic semantic value because they are manifestations of deeply ingrained psychological archetypes seeking to express themselves (Jung 1991). By interpreting the archetypal messages our dreams present to us in symbolic form we can, therefore, achieve meaningful insights that escape the reach of ordinary waking introspection (Ackroyd 1993, von Franz and Boa 1994, Jung 2002). Now,

if the world is akin to a collective dream also produced by mental archetypes, as discussed in the previous section, *then the same rationale should apply to our waking lives*. The meanings we think to discern in the world may not, after all, be mere personal projections, but actual properties *of the world*. All empirical facts may be archetypal symbols: extrinsic appearances of immanent mental dynamics. The entire cosmic narrative may be hinting at something prior to, and beyond, itself.

In a mental world, the images we perceive on the screen of perception aren't essentially different from our own imagination, except in that the former are shared across observers. This collective 'world dream' symbolically points to underlying *trans*personal mental dynamics, just as regular dreams symbolically point to underlying *personal* mental dynamics. As such, the world is amenable to hermeneutics: it means something; it points to something beyond its face-value appearances; it evokes something *a priori*; it is not its own meaning.

A.6 What the world's traditions have to say

Curiously, despite empirical evidence for the mental world hypothesis having become available only in relatively recent times, philosophical and spiritual traditions have been hinting at the intrinsic semantic value of the world for millennia. For instance, based on his in-depth study of ancient Islamic mysticism, Henry Corbin suggested that the purpose of life is to interpret the world as a metaphor of transcendent meaning. He wrote:

> To come into this world ... means ... to pass into the plane of existence which in relation to [Paradise] is merely a metaphoric existence. ... Thus coming into this world has meaning only with a view to *leading that which is metaphoric back to true being*. (As quoted in Cheetham 2012: 59, emphasis added.)

That the world isn't literal but metaphorical implies that it isn't the end of a chain of cognitive associations. Instead, its very purpose is to evoke, to point to cognition beyond its face-value appearances.

Analogously, in a clear suggestion that the things and phenomena of the world are symbols of transpersonal mental patterns, Hong Zicheng wrote in the sixteenth century:

> The chirping of birds and twittering of insects are all *murmurings of the mind*. The brilliance of flowers and colors of grasses are none other than the patterns of the Dao. (2006: 105, emphasis added.)

Still along similar lines, the Hermetic tradition suggests that the world is a mental creation in a transpersonal mind:

> That Light, He said, am I, thy God, *Mind* ... *Mind is Father-God*. ... He [God] *thinketh* all things manifest ... [and] manifests through all things and in all. (Mead 2010: 3, 23, emphasis added.)

It then proceeds to suggest that the world is the symbolic image of these immanent, transpersonal mental processes:

> Holy art Thou, O God … *of whom All-nature hath been made an image*. (Mead 2010: 11, emphasis added.)

In the West, the inception of these notions goes, of course, all the way back to Plato and his 'Theory of Ideas,' according to which the ontological ground of reality is archetypal thoughts in a transpersonal mind (Ross 1951). The visible world around us is supposedly modeled after the patterns of these archetypal thoughts, which it thus symbolically points to.

Echoing all this, Nisargadatta Maharaj, a twentieth-century exponent of the Advaita Vedanta tradition in India, said:

> When you see the world you see God. There is no seeing God apart from the world. Beyond the world to see God is to be God. (1973: 58)

Thus, our only access to God is through the images on the screen of perception that we call the world. These images are the extrinsic appearance of God's conscious inner life. Beyond them, the only way to know God is to gain *direct* access to God's inner life—that is, to *be* God.

I will mention just one more example, since an exhaustive review of how these ideas are represented in the world's traditions is beyond the scope of this brief essay. Christian mystic and scientist Emanuel Swedenborg wrote extensively of the "correspondences" between the natural and spiritual worlds (2007: 63). These correspondences imply that the things and phenomena of the natural world are symbolic images of deeper, transcendent truths. The "correspondences" were Swedenborg's attempt to formulate a hermeneutic of the world.[1]

A.7 Discussion

Physicalism has served important practical purposes over the past couple of centuries. It has provided scientists and engineers with an effective—if simplistic and ultimately wrong—picture of the world, conducive to the development of technology. By thinking of objects and natural phenomena as having standalone reality independent of their own minds, practitioners could achieve the degree of detachment and objectivity necessary for describing the world without bias. The predictive models of nature's behavior that resulted from this effort now lie at the foundation of our technological civilization.

[1] Here it would have been interesting to mention the vast literature of medieval scholasticism in Europe that resonates directly with the ideas presented in this dissertation. For instance, in his analysis of the thought of medieval scholars, Owen Barfield, based mostly on the writings of Thomas Aquinas, says that, to them, "the world is the thought of God" (2011: 95). Allowing for some leeway regarding the use of the word 'God,' this is precisely a conclusion of Chapter 3.

But whilst valuable in a utilitarian sense, this focus on nature's *behavior*—as opposed to nature's *meaning*—is extraordinarily limiting to the human spirit. We are meaning-seeking animals (Frankl 1991, Tillich 1952). A long and productive life enabled by continuous advances in technology is ultimately vacuous and sterile if devoid of meaning. And the same worldview that facilitates the advancement of technology precludes us from finding and appreciating the meaning of life in the world. This, in essence, is perhaps the greatest dilemma of the contemporary zeitgeist.

In such a context, the alternative notion that the world points to something beyond its face-value appearances offers enriching new perspectives. After all, the world we inhabit now carries intrinsic semantic value; a message. Like the Voynich manuscript (Reddy and Knight 2011), it is akin to a book written in a yet-undeciphered language, clamoring for a suitable hermeneutic. Ortiz-Osés' project (2008) turns out to rest on solid metaphysical foundations after all. Each of us, as individuals, can now give ourselves permission to dedicate our lives to *finding meaning in the world*, reassured by the knowledge that this meaning *is really there* even if we can't immediately apprehend it. And whereas the world's meaning won't disappear if we refuse to look for it, the point is that the *option* to look is given legitimacy.

Because of its preoccupation with measurement and predictive modeling, contemporary culture is forgetting to read the letter for the sake of describing the envelope. The physical universe we can measure is merely the carrier of something implied. Exaggerated focus on the predictive models of science, crucial as they are for the development of technology, may distract us from fulfilling what may be our natural and innate telos. In the words of Ortega y Gasset, "Scientific truth is an exact truth, but incomplete and penultimate, that is forcedly integrated in another kind of truth, ultimate and complete yet inexact" (as quoted in Ortiz-Osés 2008: 30).

Looking upon the world interpretatively, as a scholar looks upon an ancient text while trying to decipher its meaning, is not only metaphysically and teleologically sound, it can also make life more wholesome. Psychotherapist Thomas Moore offers us an example: by looking upon our family members as characters and our family stories as episodes of a great saga, meant to subtly evoke something above and beyond its pedestrian literal appearances, we open ourselves up to the deeper archetypal sense they express (2012: 32). By extrapolating this powerful idea further, we can look upon our entire life as a small but crucial element of an unfathomable, symbolic cosmic drama. The experiences we go through are no longer literal and pedestrian, but carry deeper, hidden significance. Indeed, in a mental world it is as unreasonable to interpret life literally as it is to interpret dreams literally. Whoever thinks that a dream is exactly what it appears to be at face value? Most people's instinct upon having an intense dream is to immediately ask themselves: *What does it mean?* Looking upon life in the same way—and asking oneself the same question—can bestow on it a much more spacious, open and wholesome outlook.

With its focus on closed, literal explanations, the physicalist ontology that informs the contemporary zeitgeist decrees that the world has no intrinsic meaning. Instead of an open book waiting to be deciphered and grasped, the world becomes just pixels to be measured; an endless string of quantifiable parameters carrying no message. Instead of the starting point of an open, epic journey along endless cognitive associations, wherein the meanings evoked constitute and ultimately reveal the uncanny reflection of the observer in the observed, the world becomes the end point of a botched quest that never even gets started. By doing this, the physicalist ontology gives us permission to procrastinate in semantic nihilism and an engineered sense of closure. It stops us from pursuing what the Islamic mystics studied by Corbin thought to be the purpose of life. For the ultimate meaning of it all may not be discernible in any particular end point or conclusion, but only in the cognitive gestalt entailed by a *circumambulation*—to use a handy Jungian term—of associative threads. It may be discernible only in a "galaxy" of semantic fields that "are intimately connected, and their significations influence one another, so that the most important sense is found diffuse in its whole" (Ortega y Gasset, as quoted in Ortiz-Osés 2008: 28).

Historically speaking, the denial of the intrinsic symbolic meaning of the world is a recent aberration (Tarnas 2010). The antidote for this aberration is an extension of the application of hermeneutics beyond all discernible boundaries. What we need is a hermeneutic of the entire cosmos; a *Hermeneutic of Everything*.

Appendix B. The Physicalist Worldview as Neurotic Ego-Defense Mechanism

This paper first appeared in *SAGE Open*, ISSN: 2158-2440, Vol. 6, No. 4, doi: 10.1177/2158244016674515, in October 2016.

B.1 Abstract

The physicalist worldview is often portrayed as a dispassionate interpretation of reality motivated purely by observable facts. In this article, ideas of both depth and social psychology are used to show that this portrayal may not be accurate. Physicalism—whether it ultimately turns out to be philosophically correct or not[1]—is hypothesized to be partly motivated by the neurotic endeavor to project onto the world attributes that help one avoid confronting unacknowledged aspects of one's own inner life. Moreover, contrary to what most people assume, physicalism creates an opportunity for the intellectual elites who develop and promote it to maintain a sense of meaning in their own lives through fluid compensation. However, because this compensatory strategy does not apply to a large segment of society, it creates a schism—with corresponding tensions—that may help explain the contemporary conflict between neo-atheism and religious belief.

B.2 Introduction

A worldview is a narrative in terms of which we relate to ourselves and reality at large. It is a kind of cultural operating system that gives us tentative answers to foundational questions such as 'What are we?' 'What is the nature of reality?' 'What is the purpose of life?' and so on (Kastrup 2014). Although many different worldviews vie for dominance today, the academically endorsed physicalist narrative defines the mainstream, despite its many difficulties (Kastrup 2014, 2015, Nagel 2012). This reigning worldview posits that physical entities outside consciousness are the building blocks of reality. Consciousness, in turn, is supposedly an epiphenomenon or emergent property of certain complex arrangements of these entities. As such, under physicalism, consciousness must be reducible to physical arrangements outside and independent of experience (Stoljar 2016).

Physicalism is often portrayed as a worldview that, in contrast to, for example, religion or spirituality, is based solely on objective facts. The present article, however, hypothesizes that the formative principles and motivations

[1] As I have extensively articulated earlier in this dissertation, my position is that physicalism is demonstrably inferior to idealism on both logical and empirical grounds. Yet, since the scope of this particular paper is restricted to psychology, my tone had to be neutral regarding philosophical matters.

underpinning the physicalist narrative—whether it ultimately turns out to be philosophically correct or not—are partly subjective, reflecting neurotic ego-defense maneuvers meant, as described by Vaillant, to "protect the individual from painful emotions, ideas, and drives" (1992: 3). This becomes clear when one lifts core concepts of depth psychology to the social and cultural spheres. However, as a mostly clinical approach, depth psychology requires some elaboration before being applied at a theoretical level.

The modern understanding of depth psychology can be traced back to the late 19th and early 20th centuries, in the works of Frederic Myers, Pierre Janet, William James, Sigmund Freud and Carl Jung (Kelly *et al.* 2009). Its foundational inference is that the human psyche comprises two main subdivisions: a conscious and an 'unconscious' segment. The conscious segment of the psyche comprises experiences a person has introspective access to and can report. According to the analytical school of depth psychology, the "ego" is defined as the experiential center of this segment (von Franz 1964: 161), and it is in this specific sense that I use the word 'ego' throughout the present article. In contrast, the so-called 'unconscious' segment of the psyche comprises mental contents the person has no introspective access to and cannot report. Nonetheless, depth psychologists assert that 'unconscious' mental contents can, and do, influence the person's manifest thoughts, feelings and behaviors.

Because the ability to report an experience is a metacognitive capacity on top of the experience itself (Schooler 2002), a more rigorous articulation of the difference between the conscious and 'unconscious' segments of the psyche is this: conscious mental contents are those a person *both* experiences *and* knows that he or she experiences them. 'Unconscious' mental contents, on the other hand, are those the person either does not experience or does not know *that* he or she experiences them (Kastrup 2014: 104-110). In other words, conscious mental contents fall within the field of egoic self-reflection and, therefore, can be reported, whereas 'unconscious' mental contents escape this field and, therefore, cannot be reported. Indeed, the existence of mental contents that are experienced but cannot be reported—even to oneself—is now well established in neuroscience, which has prompted the emergence of so-called "no-report paradigms" (Tsuchiya *et al.* 2015).

However, as clinical psychologists can only gauge consciousness based on what their patients report, anything outside the field of self-reflection is indistinguishable from true unconsciousness. This explains the somewhat inaccurate terminology choice of the founders of depth psychology.[2]

Some critics have questioned the existence of an 'unconscious' segment of the psyche on philosophical grounds (Stannard 1980: 51-81). However, recent

[2] See Chapter 5 for a much more extensive elaboration on the nature of the 'unconscious,' including the role of dissociation, which I have not discussed in this particular article. In a nutshell, my position is that there is no actual unconscious, but simply *conscious* mental processes inaccessible to egoic introspection because they (a) escape the field of self-reflection or (b) are strongly dissociated from the ego.

empirical results in neuroscience show the presence of broad cognitive activity that individuals cannot report, but which nonetheless causally conditions the individuals' manifest thoughts, feelings or behaviors (Augusto 2010, Eagleman 2011, Westen 1999). Recent neuroimaging studies of the psychedelic state have also corroborated the depth-psychological view that ego suppression—in the form of reduction of neural activity in the brain's default mode network—brings otherwise 'unconscious' mental contents into awareness (Carhart-Harris *et al*. 2012, Carhart-Harris *et al*. 2016, Palhano-Fontes *et al*. 2015).

On the basis of these empirical results, the core idea of depth psychology—that is, that a segment of the psyche that escapes self-reflective introspection can causally condition our thoughts, feelings and behaviors—cannot be dismissed. And because cultural narratives are the compound result of an aggregation of the thoughts, feelings and behaviors of individuals, depth-psychological insights are valid starting points for an analysis of the psychological underpinnings of our culture's mainstream worldview.

In Sections B.3 and B.4, I review ways in which the physicalist narrative can give us permission to avoid confronting unwanted affects in the 'unconscious' segment of our psyche. In Section B.5, I elaborate on how physicalism can conceivably even nurture its proponents' sense of meaning in life. This latter section is based on theories of social psychology, rather than depth psychology, but it still leverages the notion of an 'unconscious': in hypothesizing that physicalism is an expression of fluid compensation, it presupposes that cognitive processes outside the field of self-reflection influence the feelings, thoughts and opinions subjects express. Finally, Section B.6 briefly sums up the key ideas defended in this article.

B.3 Ego protection through projection

According to depth psychology, a neurosis is the expression of an inner psychic conflict caused by the ego's refusal to acknowledge, confront and ultimately integrate unwanted affects rising from the 'unconscious' (Jung 2014: 137). To keep these affects at bay, the ego uses a variety of defense mechanisms, among which denial, distortion, dissociation, repression and so on (Vaillant 1992). A particularly common defense mechanism is *projection* (*ibid*.), whereby one circumvents the need to confront ego-threatening forces within oneself by ascribing the corresponding attributes to the outer environment. As such, projections can be said to partly hijack and manipulate one's worldview in an attempt to prevent short-term suffering. My hypothesis is that, through projection, the physicalist worldview gives us permission to avoid confronting some of what we find disagreeable within ourselves. This can be achieved in a variety of subtle ways.

For instance, we all have a sense of our own existence and identity. Lucid introspection reveals that the root of this sense is our consciousness—our capacity to be subjects of experience. After all, if we were not conscious, what could we know of ourselves? How could we even assert our own existence?

Being conscious is what it means to *be* us. In an important sense—perhaps even the *only* important sense—we are first and foremost consciousness itself, the rest of our self-image arising afterward, as thoughts and images constructed *in* consciousness.

From this perspective, the physicalist narrative's attempt to reduce consciousness to physical entities outside subjectivity is counterintuitive, for it divorces the alleged nature of consciousness from our felt sense of identity. We do not *feel* as though we were a bunch of physical particles bouncing around inside our skull. Instead, we feel that we are the subjective 'space' wherein our experiences unfold, including our ideas about physical particles. Hence, there is a sense in which the physicalist narrative can be said to *project* the felt essence of ourselves onto something distinctly other. According to it, we are not really 'here,' grounded in our subjective sense of being, but somewhere 'over there,' in an abstract world fundamentally beyond the felt concreteness of our inner lives. As such, the physicalist narrative entails an *emptying out* of what it means to be us; a kind of secular kenosis. "I am no ghost, just a shell," laments the art character Annlee (Huyghe and Parreno 2003: 35), whose predicament is that of many of us in contemporary society.

The kenosis entailed by the physicalist narrative can exonerate its proponents from responsibility for their choices and actions. Consider this passage by Sam Harris: "Did I consciously choose coffee over tea? No. *The choice was made for me by events in my brain* that I … could not inspect or influence" (2012: 7-8, emphasis added). The projection of responsibility here is clear and the corresponding release described by Harris himself: "Losing a belief in free will has not made me fatalistic—in fact, *it has increased my feelings of freedom*. My hopes, fears, and neuroses seem *less personal*" (2012: 46, emphasis added). Indeed, under the ethos of such a worldview, there is no concrete reason for guilt or regret, for we allegedly are not what we experience ourselves to be. We are not responsible for what happens *here* because we are not—and have never been—really *here*. We are not ghosts in the machine but ghosts *conjured up by* the machine. In a significant sense, we do not really exist.

As a matter of fact, some proponents of the physicalist narrative go as far as to deny that consciousness exists. "Consciousness doesn't happen. It's a mistaken construct." These words of neuroscientist Michael Graziano (2016) should give anyone pause for thought. Here we have consciousness—whatever it may intrinsically be—denying that consciousness exists. Philosopher Daniel Dennett (1991) also claimed that consciousness is an illusion, a claim that seems to immediately contradict itself. After all, where do illusions occur if not in consciousness? [3] By appealing to metaphysical abstractions fundamentally

[3] In the words of David Bentley Hart, "The entire notion of consciousness as an illusion is, of course, rather silly. Dennett has been making the argument for most of his career, and it is just abrasively counterintuitive enough to create the strong suspicion in many that it must be more philosophically cogent than it seems, because surely no one would say such a thing if there were not some subtle and penetrating truth hidden behind its apparent absurdity. But there is none. The simple truth of the

beyond experience, such denials of our felt selves achieve a form of deliverance somewhat analogous to religious absolution. Surprisingly, as we will later see, they even help restore a sense of meaningfulness in life, following what I will call 'ontological trauma.'

The structure of these denials is fairly clear: first, consciousness weaves the conceptual notion that certain aspects of its own dynamics somehow exist outside itself; then, it projects its own essence onto these aspects. The corresponding dislocation of identity is apparent—and its neurotic character easy to grasp—with an analogy: imagine a painter who, having painted a self-portrait, points at it and declares himself to *be* the portrait. This, in essence, is what physicalists do, whether it is philosophically justifiable or not.[4] Their consciousness conceptualizes self-portraits within itself. Sometimes these self-portraits take the form of electrical impulses and neurotransmitter releases in the brain (Koch 2004). Other times, they take the shape of quantum transitions or potentials (Tarlaci and Pregnolato 2016). Whatever the case, their consciousness always points to a conceptual entity it creates within itself and then declares itself to *be* this entity. It dismisses its own primary, first-person point of view in favor of an abstract third-person perspective. Consider Dennett's words: "The way to answer these 'first-person point of view' stumpers is *to ignore the first-person point of view* and examine what can be learned from the third-person point of view" (1991: 336, emphasis added). The contempt for the *subject* of experience—the primary datum of existence and one's own felt identity—is palpable here; the kenosis nearly total.

The physicalist narrative may also give us permission to carve out and dismiss—again through the kenosis of projection—the most difficult aspects of our inner lives: our felt emotions. According to it, the feeling of an emotion is the internal perception of an "action program" triggered by certain stimuli (Damasio 2011). Although the action program itself is important insofar as it helps us survive and reproduce, the accompanying feeling of emotion is, in a sense, a mere side effect of the program's execution. For instance, the sight of another human being facing a predicament is a stimulus that triggers actions meant to help the victim and, consequently, increase the social cachet of the action taker. The *feeling* of compassion, in turn, is supposedly nothing but the inner perception of this evolutionarily useful reactive schema (Immordino-Yang *et al*. 2009); it allegedly has no primary or fundamental significance. Under such a narrative, it is easier to go into denial about our emotional lives when the going gets tough. We feel justified to dismiss or repress our traumas and demons, avoiding the often-painful work of psychological integration. The

matter is that Dennett is a fanatic: He believes so fiercely in the unique authority and absolutely comprehensive competency of the third-person scientific perspective that he is willing to deny not only the analytic authority, but also the actual existence, of the first-person vantage" (2017).

[4] For clarity, and at the cost of repeating myself, my position is that this is *not* philosophically justifiable, as I have extensively argued earlier in this dissertation.

physicalist narrative provides a foundation for rationalizing the choice of living an unexamined, superficial life. To a person desperate to avoid the specter of immediate and pungent suffering, the benefits of this stance may seem to far outweigh its potential long-term implications.

Surprisingly, the physicalist narrative can even offer us reassurance about death. According to it, there is literally nothing to fear about death itself, because it is allegedly the end of all experience, including the experiences of fear and pain. All of our problems and suffering are guaranteed to end at that point. The great and scary *unknown* of the experiential realm beyond physical existence vanishes in one fell swoop; the greatest angst of humankind is conquered. The psychological allure of this idea is powerful, yet most people do not seem to ever stop to consider it. We have come to take for granted the comforts that our mainstream worldview grants us.

To sum it up, by denying our felt sense of existence and identity, the physicalist narrative creates an opportunity to clear the ego of ultimate responsibility. By denying the fundamental reality of emotions, it creates an opportunity to protect the ego from a confrontation with far more powerful forces. And by projecting our ontological essence onto ephemeral arrangements of matter, it creates an opportunity to protect the ego from what has historically been the greatest angst of humankind: the experiential unknown of the after-death state.

B.4 Egoic control

It has been shown that religiosity can reflect a form of compensatory control (Kay *et al*. 2010): by believing that transcendent forces aligned with one's convictions govern the world, the ego avoids the anxiety associated with its own inability to overcome uncertainty. This way, religiosity creates an opportunity for control *by proxy*: although the ego cannot determine the course of nature, an external agency far superior to it is believed to do so in a way consistent with the ego's preferences. The ego's need to avoid anxiety by exerting control is thus *indirectly* fulfilled.

Going beyond religiosity, the physicalist narrative enables a sense of *direct* egoic control over nature. Indeed, a recent empirical study has shown that "believing that science is or will prospectively grant … mastery of nature imbues individuals with the belief that they are in control of their lives" (Stavrova, Ehlebracht and Fetchenhauer 2016: 234). Of course, by associating itself with science—in a philosophically questionable move that is nonetheless widely accepted—the physicalist narrative has become the enabler and ontological foundation of this belief. And because *direct* control—the notion that one can personally steer or at least predict what is going to happen—is known to be a key contributor to mental well-being (Langer and Rodin 1976, Luck *et al*. 1999), it stands to reason that the allure of physicalism in this regard could potentially be even stronger than that of religious control-by-proxy.

The opportunity for direct control offered by the physicalist narrative goes as far as conquering death itself: if consciousness is just an epiphenomenon or

emergent property of physical arrangements outside experience, it becomes conceivable that, through smart engineering, we could create means to upload our consciousness into more durable substrates such as silicon computers (Kurzweil 2005). Some physicalists even offer detailed roadmaps for achieving this (Sandberg and Boström 2008). The possibility of eternal life thus seems to open up, provided that consciousness can be instantiated in a computer by programming the computer with the patterns of information flow found in a person's brain.

This, however, is premised on the notion that a simulation of a mental phenomenon is equivalent, *in essence*, to the phenomenon itself. There are many compelling arguments against this notion in philosophy of mind, the most well known of which is perhaps John Searle's (2004). To gain some intuition about what these arguments generally entail, consider this: Do we have any reason to believe that, by performing a perfectly accurate simulation of kidney function in a computer, the computer will begin urinating on its desk? Clearly not. There is an essential difference between a computer simulation and the phenomenon it simulates; they are not the same thing, no matter how accurate the simulation. Yet, those hoping to 'upload consciousness' under the physicalist narrative seem to become so engrossed in abstraction that they lose touch with basic intuitions of plausibility. Their neurosis is, in this sense, comparable with religious dogmatism.

Although both the religious and physicalist narratives create an opportunity for conquering death, the Promethean door to immortality opened by physicalism invests the *ego*—not deities—with the power to control transcendence through technology. This is seductively more direct, its only weakness—from a psychological standpoint—being that it is promissory: at present, nobody has ever managed to upload consciousness. Yet, some popular physicalist authors argue that consciousness uploading may be achievable *still in our own lifetime* (Kurzweil 2005, Sandberg and Boström 2008), which actualizes the potential allure of their worldview.

As seen in Section B.3 and this section, the implications of the physicalist narrative consistently help protect and invest the ego with authority. This is not to say that physicalism is entirely motivated by neurotic ego-defense maneuvers, for there is a philosophical argument behind it that cannot be dismissed. Nonetheless, the question is whether it is plausible that physicalism's significant ego-defense potential has *not* been, to some degree, an unexamined motivation for its development, promotion and adoption.

B.5 The question of meaning

Meaning—in the sense of significance and purpose—is probably the greatest asset any human being can possess. Psychotherapist Victor Frankl (1991), who practiced and led groups while detained in a concentration camp during World War II, asserted that the *will-to-meaning* is the most dominant human drive, in contrast to Nietzsche's will-to-power and Freud's will-to-pleasure. Meaning is

so powerful that, as Jung remarked, it "makes a great many things endurable—perhaps everything" (1995: 373). Philip K. Dick's alter ego Horselover Fat, in the novel *Valis*, embodies the essence of this drive: "Fat had no concept of enjoyment; he understood only meaning," wrote Dick (2001: 92). Like Fat, many of us see meaning as a higher value than power or pleasure. Our motivation to live rests in there being meaning in our lives. Today, we need meaning more than ever, for as Paul Tillich (1952) lucidly observed, the greatest anxieties of our culture are precisely those of *doubt* and *meaninglessness*.

And here is where an argument is often made for the impartiality of physicalism: as a worldview that, by turning the universe into a mechanical contraption fueled by mere chance, drains the meaning out of life, it cannot possibly be a neurotic ego-defense mechanism—or so the argument goes. Instead, the physicalist narrative must represent a courageous admission by "tough people who face the bleak facts" (Watts 1989: 65). It must embody an objective assessment of reality, not an emotional, irrational wish-fulfillment maneuver akin to religion. Compelling as it may seem at first, this argument fails careful scrutiny, for its premise is false.

Indeed, according to the Meaning Maintenance Model (MMM) of social psychology (Heine, Proulx and Vohs 2006)—which is perhaps better seen in the context of a broader theory of psychological defense (Hart 2013)—we can derive a sense of meaning from four different sources: self-esteem, closure, belonging, and symbolic immortality. In other words, we can find meaning in life through (a) cultivating a feeling of personal worth, (b) resolving doubts and ambiguities, (c) being part of something bigger and longer-lasting than ourselves, and (d) leaving something of significance behind—such as professional achievements—in the form of which we can 'live on' after physical death. A society's mainstream cultural narrative conditions how meaning can be derived from each of these four sources.

The key idea behind the MMM is that of *fluid compensation* as an ego-defense mechanism: If one of the four sources of meaning is threatened, an individual will tend to automatically compensate by seeking extra meaning from the other three sources. For instance, threats to self-esteem may cause the individual to reaffirm his or her model of reality, thereby bolstering closure.

As van Tongeren and Green (2010) have shown, a transcendent source of meaning, such as religion, plays the same role in fluid compensation as the other four sources. For instance, individuals tend to reaffirm their religious beliefs following disruption to their meaning system, in an effort to protect the latter. Van Tongeren's and Green's experiments have not only empirically substantiated the MMM, they have also shown that even *subliminal* threats to meaning trigger fluid compensation, strongly indicating that the 'unconscious' is integral to the process.

With this as background, my suggestion is that the physicalist narrative, in addition to being a rational hypothesis for making sense of the world,[5] may be an expression of fluid compensation by intellectual elites. In other words, instead of a threat to meaning, the physicalist narrative may actually reflect an attempt by these elites to protect and restore their sense of meaning through bolstering closure, self-esteem and symbolic immortality. The disruption that may have originally led to this compensatory move occurred around the mid- to late-19th century.

Indeed, it was at that time that we lost our ability to spontaneously relate to religious myths without linear intellectual scrutiny. "With Descartes and Kant, the philosophical relation between Christian belief and human rationality had grown ever more attenuated. By the late nineteenth century, with few exceptions, that relation was effectively absent," wrote Tarnas (2010: 311). The myths that had hitherto offered us meaning through the promise of *literal* immortality and metaphysical teleology became untenable. Taylor, who richly chronicled this historical transition, characterized the corresponding loss of meaning rather broadly and generally as "a wide sense of malaise at the disenchanted world, a sense of it as flat, empty" (2007: 302). He even hinted at fluid compensation when speaking of "a multiform search for something within, or beyond [the world], which could compensate for the meaning lost with transcendence" (*ibid.*).

While acknowledging that this generalized malaise was the matrix of what followed, I submit that a more specific, forceful and *personal* threat to meaning was necessary to mobilize the extraordinary level of academic and intellectual endorsement amassed by physicalism. After all—as Taylor himself described through what he called "the nova effect"—the malaise, in and by itself, fostered not only physicalism but also an explosion of myriad other worldviews.

I hypothesize that a profound and disturbing change in the intellectual elites' understanding of the nature *of their own being*—that is, an ontological trauma—was the specific, forceful and personal trigger that helped congeal the physicalist narrative. Having lost religion, the elites were left with the prospect of physical deterioration without the path to transcendence previously offered by an immortal soul. Hence, they were forced to face the inexorability of their own approaching death. And as we know from Terror Management Theory, mortality salience is a formidable threat to meaning (Pyszczynski, Greenberg and Solomon 1997) empirically shown to motivate investment in palliative worldviews (Burke, Martens and Faucher 2010). Ontological trauma may have thus triggered fluid compensation and ultimately led to the intellectual elites' championing of the physicalist narrative.

Indeed, many studies have shown that mortality salience leads to a heightened need for *closure* (Landau *et al.* 2004). This is fluid compensation in action.

[5] Again, here I am making a charitable concession to physicalism because the limited scope of this particular article—focused, as it is, on psychology—prevented me from arguing against it philosophically.

Notice also that the physicalist narrative is humanity's most significant attempt yet to achieve *closure* in our worldview. As multibillion-dollar experiments like the Large Hadron Collider—whose primary purpose is to 'close' the Standard Model of particle physics, with no immediate practical applications—illustrate, physicalism embodies an unprecedented effort to produce a causally complete, unambiguous model of reality. Nothing else in millennia of preceding history has come anywhere near it. I suggest that this is not coincidental: the physicalist narrative may reflect the elites' ego's attempt to regain, through heightened *closure*, the meaning it lost along with religion. Moreover, other modalities of fluid compensation may be at play here as well: by distinguishing themselves as a segment of society uniquely capable of understanding facts and concepts beyond the cognitive capacity of others, the scientists and academics who promote the physicalist narrative stand to gain in self-esteem. The cosmological scope of the scientific work they produce and leave behind upon their deaths can also be seen as a boost to symbolic immortality. Finally, recall Tillich's observation: *doubt* and *meaninglessness* anxiety dominate our culture's mindset. Is it humanly plausible that our mainstream narrative would have evolved to tackle only doubt and leave meaninglessness anxiety unaddressed?

All in all, the physicalist narrative does not necessarily represent a net loss of meaning for the intellectual elites who produced and continue to promote it. The transcendent meaning lost along with religion may be compensated for by an increase in closure, self-esteem and symbolic immortality. Unfortunately, however, this compensatory strategy cannot work for most ordinary people: the men and women on the streets do not have enough grasp of contemporary scientific theories to experience an increase in their sense of closure. Neither do they gain in self-esteem, because they are not part of the distinguished elites. Finally, insofar as ordinary people do not produce scientific work of their own, no particular gain in symbolic immortality is to be expected either.

In conclusion, the physicalist narrative may serve the egoic meaning needs of the intellectual elites who develop and promote it, but constitutes a significant threat to the sense of meaning of the average person on the streets. Perhaps for this reason, a large segment of society seeks meaning through alternative ontologies considered outdated and untenable by the intellectual elites, such as religious dualism (Heflick *et al*. 2015). This creates a schism—with corresponding tensions—between different segments of society, which may help explain the contemporary conflict between neo-atheism and religious belief.

B.6 Conclusion

The physicalist narrative, in contrast to the way it is normally portrayed, may not be dispassionate. It may be partly driven by the neurotic endeavor to project onto the world attributes that help us avoid confronting unacknowledged aspects of our own inner lives. Moreover, contrary to what most people assume, physicalism creates an opportunity for the intellectual elites who develop and promote it to maintain a sense of meaning in their own

lives through fluid compensation. However, because this compensatory strategy does not apply to a large segment of society, it creates a schism—with corresponding tensions—that may help explain the contemporary conflict between neo-atheism and religious belief.

Bibliography

Ackroyd, E. (1993). *A Dictionary of Dream Symbols*. London, UK: Cassell Illustrated.

Adelson, EH (1995). Checker shadow illusion. [Online]. Available from: http://persci.mit.edu/gallery/checkershadow [Accessed 12 April 2017].

Albert, H. (1985). *Treatise on Critical Reason*. Princeton, NJ: Princeton University Press.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC: American Psychiatric Publishing.

Ananthaswamy, A. (2011). Quantum magic trick shows reality is what you make it. *New Scientist*, June 22. [Online]. Available from: https://www.newscientist.com/article/dn20600-quantum-magic-trick-shows-reality-is-what-you-make-it/ [Accessed 14 June 2016].

Aspect, A., Grangier, P. and Roger, G. (1981). Experimental tests of realistic local theories via Bell's theorem. *Physical Review Letters*, 47 (7): 460-463.

Aspect, A., Dalibard, J. and Roger, G. (1982). Experimental test of Bell's inequalities using time-varying analyzers. *Physical Review Letters*, 49 (25): 1804-1807.

Aspect, A., Grangier, P. and Roger, G. (1982). Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: A new violation of Bell's inequalities. *Physical Review Letters*, 49 (2): 91-94.

Augusto, LM (2010). Unconscious knowledge: A survey. *Advances in Cognitive Psychology*, 6: 116-141.

Barbour, J. (1999). *The End of Time: The Next Revolution in Physics*. Oxford, UK: Oxford University Press.

Barfield, O. (2011). *Saving the Appearances: A Study in Idolatry*. Oxford, UK: Barfield Press.

Barrett, D. (1994). Dreams in dissociative disorders. *Dreaming*, 4 (3): 165-175.

Bell, J. (1964). On the Einstein Podolsky Rosen Paradox. *Physics*, 1 (3): 195-200.

Black, DW and Grant, JE (2014). *The Essential Companion to the Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC: American Psychiatric Publishing.

Blanke, O. *et al*. (2002). Stimulating illusory own-body perceptions: The part of the brain that can induce out-of-body experiences has been located. *Nature*, 419: 269-270.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Science,* 18: 227-287.

Boly, M. *et al*. (2011). Preserved feedforward but impaired top-down processes in the vegetative state. *Science*, 332 (6031): 858-862.

Boswell, J. (1820). *The Life of Samuel Johnson, LL. D.* (Volume 1). London, UK: J. Davis, Military Chronicle and Military Classics Office.

Braude, S. (1995). *First Person Plural: Multiple Personality and the Philosophy of Mind*. New York, NY: Routledge.

Brenner, ED *et al*. (2006). Plant neurobiology: An integrated view of plant signaling. *Trends in Plant Science*, 11 (8): 413-419.

Buonomano, D. (2018). *Your Brain Is a Time Machine: The Neuroscience and Physics of Time*. New York, NY: W. W. Norton & Company.

Burke, BL, Martens, A. and Faucher, EH (2010). Two decades of terror management theory: A meta-analysis of mortality salience research. *Personality and Social Psychology Review*, 14 (2): 155-195.

Carhart-Harris, RL *et al*. (2012). Neural correlates of the psychedelic state as determined by fMRI studies with psilocybin. *Proceedings of the National Academy of Sciences of the United States of America*, 109 (6): 2138-2143.

Carhart-Harris, RL *et al*. (2016). Neural correlates of the LSD experience revealed by multimodal neuroimaging. *Proceedings of the National Academy of Sciences of the United States of America* (PNAS Early Edition), doi: 10.1073/pnas.1518377113.

Carruthers, P. and Schechter, E. (2006). Can panpsychism bridge the explanatory gap? *Journal of Consciousness Studies*, 13 (10-11): 32-39.

Cartwright, J. (2007). Quantum physics says goodbye to reality. *IOP Physics World*, April 20. [Online]. Available from: http://physicsworld.com/cws/article/news/2007/apr/20/quantum-physics-says-goodbye-to-reality [Accessed 14 June 2016].

Catani, D. (2013). *Evil: A History in Modern French Literature and Thought*. London, UK: Bloomsbury.

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

Chalmers, D. (2003). Consciousness and its place in nature. In: Stich, S. & Warfield, T. (eds.). *Blackwell Guide to the Philosophy of Mind*. Malden, MA: Blackwell.

Chalmers, D. (2006). Strong and weak emergence. In: Davies, P. and Clayton, P. (eds.). *The Re-Emergence of Emergence*. Oxford, UK: Oxford University Press.

Chalmers, D. (2016a). The combination problem for panpsychism. In: Brüntrup, G. & Jaskolla, L. (eds.). *Panpsychism*. Oxford, UK: Oxford University Press.

Chalmers, D. (2016b). Panpsychism and panprotopsychism. In: Brüntrup, G. & Jaskolla, L. (eds.). *Panpsychism*. Oxford, UK: Oxford University Press.

Chalmers, D. (2018). Idealism and the mind-body problem. In: Seager, W. (ed.). *The Routledge Handbook of Panpsychism*. London, UK: Routledge.

Cheetham, T. (2012). *All the World an Icon: Henry Corbin and the Angelic Function of Beings*. Berkeley, CA: North Atlantic Books.

Cleeremans, A. (2011). The radical plasticity thesis: How the brain learns to be conscious. *Frontiers in Psychology*, 2, article 86.

Coleman, S. (2014). The real combination problem: Panpsychism, micro-subjects, and emergence. *Erkenntnis*, 79 (1): 19-44.

Cristofori, I. *et al*. (2016). Neural correlates of mystical experience. *Neuropsychologia*, 80: 212-220.

Crowther, K. (2014). *Appearing Out of Nowhere: The Emergence of Spacetime in Quantum Gravity*. Doctoral Dissertation, The University of Sidney.

Damasio, A. (2011). Neural basis of emotions. *Scholarpedia*, 6 (3): 1804.

Dehaene, S. and Changeux, J-P (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, 70 (2): 200-227.

Dennett, D. (1991). *Consciousness Explained*. London, UK: Penguin Books.

Dick, PK (2001). *Valis*. London, UK: Gollancz.

Dijksterhuis, A. and Nordgren, LF (2006). A theory of unconscious thought. *Perspectives on Psychological Science*, 1 (2): 95-109.

DiSalvo, D. (2012). When you inject spirit mediums' brains with radioactive chemicals, strange things happen. *Forbes*, November 18. [Online]. Available from: http://www.forbes.com/sites/daviddisalvo/2012/11/18/when-you-inject-spirit-mediums-brains-with-radioactive-chemicals-some-really-strange-things-happen/ [Accessed 25 February 2017].

Eagleman, DM (2009). Brain time. In: Brockman, M. (ed.). *What's Next? Dispatches on the Future of Science*. New York, NY: Vintage.

Eagleman, DM (2011). *Incognito: The Secret Lives of the Brain*. New York, NY: Canongate.

Eddington, AS (1928). *The Nature of the Physical World*. New York: The MacMillan Company.

Einstein, A., Podolsky, B. and Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47: 777-780.

Eliade, M. (2009). *Rites and Symbols of Initiation: The Mysteries of Birth and Rebirth*. New York, NY: Spring Publications.

Fields, C. *et al*. (2017). Eigenforms, interfaces and holographic encoding: Toward an evolutionary account of objects and spacetime. *Constructivist Foundations*, 12 (3): 265-291.

Floridi, L. (2008). Trends in the philosophy of information. In: Adriaans, P. and Benthem, J. van (eds.). *Handbook of the Philosophy of Science, Volume 8:*

Philosophy of Information. Amsterdam, The Netherlands: Elsevier, pp. 113-131.

Fonagy, P. *et al*. (eds.) (2012). *The Significance of Dreams*. London, UK: Karnac Books.

Ford, BJ (2010). The secret power of the single cell. *New Scientist*, 206 (2757): 26-27.

Fraassen, BC van (1980). *The Scientific Image*. Oxford, UK: Oxford University Press.

Fraassen, BC van (1990). *Laws and Symmetry*. Oxford, UK: Oxford University Press.

Frankl, VE (1991). *The Will to Meaning* (Expanded ed.). New York, NY: Meridian.

Franklin, SP (1997). *Artificial Minds*. Cambridge, MA: MIT Press.

Franz, M-L von (1964). The process of individuation. In: Jung, CG (ed.). *Man and His Symbols*. New York, NY: Anchor Press.

Franz, M-L von (1974). *Number and Time: Reflections Leading Toward a Unification of Depth Psychology and Physics*. Evanston, IL: Northwestern University Press.

Franz, M-L von and Boa, F. (1994). *The Way of the Dream*. Boston, MA: Shambhala Publications.

Fredkin, E. (2003). An introduction to digital philosophy. *International Journal of Theoretical Physics*, 42 (2): 189-247.

Fredkin, E. (n.a.). *On the Soul (draft)*. [Online]. http://www.digitalphilosophy.org//wp-content/uploads/2015/07/on_the_soul.pdf [Accessed 2 July 2016].

Friston, K. (2013). Life as we know it. *Journal of the Royal Society Interface*, 10 (86): 20130475.

Friston, K., Sengupta, B. and Auletta, G. (2014). Cognitive dynamics: From attractors to active inference. *Proceedings of the IEEE*, 102 (4): 427-445.

Gaal, S. van *et al*. (2011). Dissociable brain mechanisms underlying the conscious and unconscious control of behavior. *Journal of Cognitive Neuroscience*, 23 (1): 91-105.

Gabrielsen, P. (2013). When does your baby become conscious? *Science*, April 18. [Online]. Available from: http://www.sciencemag.org/news/2013/04/when-does-your-baby-become-conscious [Accessed 10 September 2017].

Gillespie, A. (2007). The social basis of self-reflection. In: Valsiner, J. and Rosa, A. (eds.). *The Cambridge Handbook of Sociocultural Psychology*. New York, NY: Cambridge University Press, pp. 678-691.

Glasersfeld, E. von (1987). An introduction to radical constructivism. In: Watzlawick, P. (ed.). *The Invented Reality*. New York, NY: W. W. Norton & Company.

Godfrey-Smith, P. (2014). *Philosophy of Biology*. Princeton, NJ: Princeton University Press.

Goff, P. (2006). Experiences don't sum. *Journal of Consciousness Studies*, 13 (10-11): 53-61.

Goff, P. (2009). Why Panpsychism doesn't help us explain consciousness. *Dialectica*, 63 (3): 289-311.

Graziano, M. (2016). Consciousness is not mysterious. *The Atlantic*, January 12. [Online]. Available from: http://www.theatlantic.com/science/archive/2016/01/consciousness-color-brain/423522/ [Accessed 26 February 2017].

Greene, B. (2003). *The Elegant Universe: Superstrings, Hidden Dimensions, and the Quest for the Ultimate Theory*. New York, NY: WW Norton & Company.

Griffin, D. (1998). *Unsnarling the World-Knot*. Eugene, OR: Wipf & Stock.

Griffiths, RR *et al*. (2006). Psilocybin can occasion mystical-type experiences having substantial and sustained personal meaning and spiritual significance. *Psychopharmacology*, 187 (3): 268-283.

Gröblacher, S. *et al*. (2007). An experimental test of non-local realism. *Nature*, 446: 871-875.

Hameroff, S. (2006). Consciousness, neurobiology and quantum mechanics: The case for a connection. In: Tuszynski, J. (ed.). *The Emerging Physics of Consciousness*. Berlin, Germany: Springer Verlag.

Harris, S. (2012). *Free Will*. New York, NY: Free Press.

Harris, S. (2016). You are more than your brain! *Big Think*, September 4. [Online]. https://www.facebook.com/BigThinkdotcom/videos/10153879575418527/ [Accessed 1 November 2016].

Hart, J. (2013). Toward an integrative theory of psychological defense. *Perspectives on Psychological Science*, 9 (1): 19-39.

Hart, DB (2017). The illusionist. *The New Atlantis*, Summer/Fall issue: 109-121.

Hassin, R., Uleman, J. and Bargh, J. (eds.) (2005). *The New Unconscious*. New York, NY: Oxford University Press.

Hassin, RR (2013). Yes it can: On the functional abilities of the human unconscious. *Perspectives on Psychological Science*, 8 (2): 195-207.

Heflick, NA *et al*. (2015). Death awareness and body-self dualism: A why and how of afterlife belief. *European Journal of Social Psychology*, 45 (2): 267-275.

Heine, SJ, Proulx, T. and Vohs, KD (2006). The meaning maintenance model: On the coherence of social motivations. *Personality and Social Psychology Review*, 10 (2): 88-110.

Henry, RC (2005). The mental universe. *Nature*, 436: 29.

Hensen, B. *et al*. (2015). Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526: 682-686.

Hilgard, E. (1977). *Divided Consciousness*. New York, NY: John Wiley & Sons.

Hoffman, DD (2009). The interface theory of perception: Natural selection drives true perception to swift extinction. In: Dickinson, S. *et al*. (eds.). *Object Categorization: Computer and Human Vision Perspectives*. Cambridge, UK: Cambridge University Press.

Hoffman, DD and Singh, M. (2012). Computational evolutionary perception. *Perception*, 41 (9): 1073-1091.

Horgan, T. and Potrč, M. (2000). Blobjectivism and indirect correspondence. *Facta Philosophica*, 2 (2): 249-270.

Husserl, E. (1970). *The Crisis of European Sciences and Transcendental Phenomenology: An Introduction to Phenomenological Philosophy*. Evanston, IL: Northwestern University Press.

Huyghe, P. and Parreno, P. (2003). *No Ghost Just a Shell*. Cologne, Germany: Verlag der Buchhandlung Walther König.

Immordino-Yang, MH *et al*. (2009). Neural correlates of admiration and compassion. *Proceedings of the National Academy of Sciences of the United States of America*, 106 (19): 8021-8026.

Jaskolla, LJ and Buck, AJ (2012). Does panexperiential holism solve the combination problem? *Journal of Consciousness Studies*, 19 (9-10): 190-199.

Jung, CG (1991). *The Archetypes and the Collective Unconscious* (2nd ed.). London, UK: Routledge.

Jung, CG (1995). *Memories, Dreams, Reflections*. London, UK: Fontana Press.

Jung, CG (2001). *Modern Man in Search of a Soul*. New York, NY: Routledge.

Jung, CG (2002). *Dreams*. London, UK: Routledge.

Jung, CG (2014). *Analytical Psychology: Its Theory and Practice* (2nd ed.). London, UK: Routledge.

Karunamuni, ND (2015). The five-aggregate model of the mind. *SAGE Open*, 5 (2), doi:10.1177/2158244015583860.

Kastrup, B. (2014). *Why Materialism Is Baloney*. Winchester, UK: Iff Books.

Kastrup, B. (2015). *Brief Peeks Beyond*. Winchester, UK: Iff Books.

Kastrup, B. (2016a). *More Than Allegory*. Winchester, UK: Iff Books.

Kastrup, B. (2016b). What neuroimaging of the psychedelic state tells us about the mind-body problem. *Journal of Cognition and Neuroethics*, 4 (2): 1-9.

Kastrup, B. (2017a). Self-transcendence correlates with brain function impairment. *Journal of Cognition and Neuroethics*, 4 (3): 33-42.

Kastrup, B. (2017b). An ontological solution to the mind-body problem. *Philosophies*, 2 (2), doi:10.3390/philosophies2020010.

Kastrup, B. (2017c). On the plausibility of idealism: Refuting criticisms. *Disputatio*, 9 (44): 13-34.

Kastrup, B. (2017d). There is an 'unconscious,' but it may well be conscious. *Europe's Journal of Psychology*, 13 (3): 559-572.

Kastrup, B. (2017e). Making sense of the mental universe. *Philosophy and Cosmology*, 19: 33-49.

Kastrup, B. (2018a). The universe in consciousness. *Journal of Consciousness Studies*, 25 (5-6): 125-155.

Kastrup, B. (2018b). Conflating abstraction with empirical observation: The false mind-matter dichotomy. *Constructivist Foundations*, 13 (3): 341-361.

Kay, AC *et al*. (2010). Religious belief as compensatory control. *Personality and Social Psychology Review*, 14 (1): 37-48.

Kelly, EF *et al*. (2009). *Irreducible Mind: Toward a Psychology for the 21st Century*. Lanham, MD: Rowman & Littlefield.

Kihlstrom, JF (1997). Consciousness and me-ness. In: Cohen, J. and Schooler, JW (eds.). *Scientific Approaches to Consciousness*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 451-468.

Kihlstrom, J. and Cork, R. (2007). Anesthesia. In: Velmans, M. & Schneider, S. (eds.). *The Blackwell Companion to Consciousness*. Oxford, UK: Blackwell.

Kim, Y-H *et al*. (2000). A delayed 'choice' quantum eraser. *Physical Review Letters*, 84: 1-5.

Klein, SB (2015). The feeling of personal ownership of one's mental states: A conceptual argument and empirical evidence for an essential, but underappreciated, mechanism of mind. *Psychology of Consciousness: Theory, Research, and Practice*, 2 (4): 355-376.

Klimov, PV *et al*. (2015). Quantum entanglement at ambient conditions in a macroscopic solid-state spin ensemble. *Science Advances*, 1 (10), e1501015.

Koch, C. (2004). *The Quest for Consciousness: A Neurobiological Approach*. Englewood, CO: Roberts & Company Publishers.

Koch, C. (2012a). *Consciousness: Confessions of a Romantic Reductionist*. Cambridge, MA: MIT Press.

Koch, C. (2012b). This is your brain on drugs. *Scientific American Mind*, 1 May. [Online]. Available from: http://www.scientificamerican.com/article/this-is-your-brain-on-drugs/ [Accessed 9 August 2016].

Krioukov, D. *et al*. (2012). Network cosmology. *Scientific Reports*, 2, doi:10.1038/srep00793.

Kurzweil, R. (2005). *The Singularity Is Near*. New York, NY: Viking.

Landau, MJ *et al*. (2004). A Function of form: Terror management and structuring the social world. *Journal of Personality and Social Psychology*, 87 (2): 190-210.

Langer, E. and Rodin, J. (1976). The effects of choice and enhanced personal responsibility for the aged: a field experiment in an institutional setting. *Journal of Personality and Social Psychology*, 34 (2): 191-198.

Lapkiewicz, R. *et al*. (2011). Experimental non-classicality of an indivisible quantum system. *Nature*, 474: 490-493.

Lee, KC *et al*. (2011). Entangling macroscopic diamonds at room temperature. *Science*, 334 (6060): 1253-1256.

Leggett, AN (2003). Nonlocal hidden-variable theories and quantum mechanics: An incompatibility theorem. *Foundations of Physics*, 33 (10): 1469-1493.

Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, 64: 354-361.

Lewis, CR *et al*. (2017). Two dose investigation of the 5-HT-agonist psilocybin on relative and global cerebral blood flow. *NeuroImage*, July, doi:10.1016/j.neuroimage.2017.07.020.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8: 529-566.

Linde, A. (1998). *Universe, Life, Consciousness*. A paper delivered at the Physics and Cosmology Group of the "Science and Spiritual Quest" program of the Center for Theology and the Natural Sciences (CTNS), Berkeley, California. [Online]. Available from: web.stanford.edu/~alinde/SpirQuest.doc [Accessed 14 June 2016].

Lommel, P. van *et al*. (2001). Near-death experience in survivors of cardiac arrest: a prospective study in the Netherlands. *The Lancet*, 358 (9298): 2039-2045.

Luck, A. *et al*. (1999). Effects of video information on precolonoscopy anxiety and knowledge: a randomised trial. *The Lancet*, 354 (9195): 2032-2035.

Lynch, J. and Kilmartin, C. (2013). *Overcoming Masculine Depression: The Pain Behind the Mask*. New York, NY: Routledge.

Lythgoe, M. *et al*. (2005). Obsessive, prolific artistic output following subarachnoid hemorrhage. *Neurology*, 64 (2): 397-398.

Ma, X-S *et al*. (2013). Quantum erasure with causally disconnected choice. *Proceedings of the National Academy of Sciences of the Unites States of America*, 110 (4): 1221-1226.

Macnab, AJ *et al*. (2009). Asphyxial games or "the choking game": A potentially fatal risk behaviour. *Injury Prevention*, 15 (1): 45-49.

Maharaj, N. (1973). *I Am That*. Mumbai, India: Chetana Publishing.

Manning, AG *et al*. (2015). Wheeler's delayed-choice gedanken experiment with a single atom. *Nature Physics*, 11: 539-542.

Mathews, F. (2011). Panpsychism as paradigm. In: Blamauer, M. (ed.). *The Mental as Fundamental*. Frankfurt, Germany: Ontos Verlag.

Mead, GRS (translator) (2010). *The Corpus Hermeticum*. Whitefish, MT: Kessinger Publishing LLC.

Merali, Z. (2015). Quantum 'spookiness' passes toughest test yet. *Nature*, News, August 27. [Online]. Available from: http://www.nature.com/news/quantum-spookiness-passes-toughest-test-yet-1.18255 [Accessed 30 August 2015].

Merleau-Ponty, M. (1964). *The Primacy of Perception: And Other Essays on Phenomenological Psychology, the Philosophy of Art, History and Politics*. Evanston, IL: Northwestern University Press.

Miller, B. *et al*. (1998). Emergence of artistic talent in frontotemporal dementia. *Neurology*, 51 (4): 978-982.

Miller, B. *et al*. (2000). Functional correlates of musical and visual ability in frontotemporal dementia. *The British Journal of Psychiatry*, 176: 458-463.

Moore, T. (2012). *Care of the Soul: An Inspirational Programme to Add Depth and Meaning to Your Everyday Life*. London, UK: Piatkus Books.

Moorjani, A. (2012). *Dying To Be Me: My Journey from Cancer, to Near Death, to True Healing*. Carlsbad, CA: Hay House.

Nagasawa, Y. and Wager, K. (2016). Panpsychism and priority cosmopsychism. In: Brüntrup, G. and Jaskolla, L. (eds.). *Panpsychism*. Oxford, UK: Oxford University Press.

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, 83 (4): 435-450.

Nagel, T. (2012). *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False*. Oxford, UK: Oxford University Press.

Neal, RM (2008). *The Path to Addiction: "And Other Troubles We Are Born to Know."* Bloomington, IN: AuthorHouse.

Neumann, J. von. (1996). *Mathematical Foundations of Quantum Mechanics*. Princeton, NJ: Princeton University Press.

Newell, BR and Shanks, DR (2014). Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences*, 37 (1): 1-19.

Nixon, GM (2010). From panexperientialism to conscious experience: The continuum of experience. *Journal of Consciousness Exploration and Research*, 1 (3): 215-233.

Okasha, S. (2002). *Philosophy of Science: A Very Short Introduction*. Oxford, UK: Oxford University Press.

Ortiz-Osés, A. (2008). *The Sense of the World*. Aurora, CO: The Davies Group Publishers.

Palhano-Fontes, F. *et al*. (2015). The psychedelic state induced by ayahuasca modulates the activity and connectivity of the default mode network. *PLoS ONE*, 10 (2): e0118143.

Paller, KA and Suzuki, S. (2014). The source of consciousness. *Trends in Cognitive Sciences*, 18 (8): 387-389.

Partington, CF (ed.) (1837). *The British Cyclopædia of Natural History: A Scientific Classification of Animals, Plants, and Minerals; With a Popular View of Their Habits, Economy, and Structure*, Vol. 3. London, UK: WS Orr & Co., Amen Corner, Paternoster-Row.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA: Morgan Kaufmann.

Peres, J. *et al*. (2012). Neuroimaging during trance state: A contribution to the study of dissociation. *PLoS ONE*, 7 (11): e49360.

Piccinini, G. (2015). Computation in physical systems. In: Zalta, EN (ed.). *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition). [Online]. Available from: http://plato.stanford.edu/archives/sum2015/entries/computation-physicalsystems [Accessed 30 June 2016].

Piore, A. (2013). The genius within. *Popular Science*, March: 46-53.

Pyszczynski, T., Greenberg, J. and Solomon, S. (1997). Why do we need what we need? A terror management perspective on the roots of human social motivation. *Psychological Inquiry*, 8 (1): 1-20.

Reddy, S. and Knight, K. (2011). What we know about the Voynich manuscript. In: *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*. Stroudsburg, PA: Association for Computational Linguistics, pp. 78-86.

Retz (2007). Tripping without drugs: Experience with hyperventilation (ID 14651). *Erowid.org*. [Online]. Available from: http://www.erowid.org/exp/14651 [Accessed 25 February 2017].

Rhinewine, JP and Williams, OJ (2007). Holotropic breathwork: The potential role of a prolonged, voluntary hyperventilation procedure as an adjunct to psychotherapy. *The Journal of Alternative and Complementary Medicine*, 13 (7): 771-776.

Robinson, H. (2016). Dualism. In: Zalta, EN (ed.). *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition). [Online]. Available from: http://plato.stanford.edu/archives/spr2016/entries/dualism [Accessed 17 June 2016].

Romero, J. *et al*. (2010). Violation of Leggett inequalities in orbital angular momentum subspaces. *New Journal of Physics*, 12: 123007. [Online]. Available from: http://iopscience.iop.org/article/10.1088/1367-2630/12/12/123007 [Accessed 14 June 2016].

Ronnenberg, A. and Martin, K. (2010). *The Book of Symbols*. Cologne, Germany: Taschen.

Rosenberg, G. (2004). *A Place for Consciousness*. New York, NY: Oxford University Press.

Rosner, RI, Lyddon, WJ and Freeman, A. (eds.) (2004). *Cognitive Therapy and Dreams*. New York, NY: Springer Publishing Company.

Ross, WD (1951). *Plato's Theory of Ideas*. Oxford, UK: Oxford University Press.

Rovelli, C. (2008). Relational quantum mechanics. *arXiv:quant-ph/9609002v2*. [Online]. Available from: https://arxiv.org/abs/quant-ph/9609002v2 [Accessed 1 August 2017].

Rovelli, C. (2018). *The Order of Time*. London, UK: Penguin.

Russell, B. (2009). *Human Knowledge: Its Scope and Limits*. London, UK: Routledge Classics.

Ryle, G. (2009). *The Concept of Mind*. London, UK: Routledge.

Sacks, O. (1985). *The Man Who Mistook His Wife for a Hat*. New York, NY: Harper & Row.

Sandberg, A. and Boström, N. (2008). *Whole Brain Emulation: A Roadmap* (Technical Report #2008-3). Oxford, UK: Future of Humanity Institute, Oxford University. [Online]. Available from: http://www.fhi.ox.ac.uk/brain-emulation-roadmap-report.pdf [Accessed 26 February 2017].

Sartre, J-P (1992). *Being and Nothingness: A Phenomenological Essay on Ontology*. New York, NY: Washington Square Press.

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

Schlumpf, Y. *et al*. (2014). Dissociative part-dependent resting-state activity in Dissociative Identity Disorder: A controlled fMRI perfusion study. *PloS ONE*, 9, doi:10.1371/journal.pone.0098795.

Schooler, J. (2002). Re-representing consciousness: dissociations between experience and meta-consciousness. *Trends in Cognitive Science*, 6 (8): 339-344.

Seager, W. (2010). Panpsychism, aggregation and combinatorial infusion. *Mind and Matter*, 8 (2): 167-184.

Searle, JR (2004). *Mind: A Brief Introduction*. Oxford, UK: Oxford University Press.

Shani, I. (2015). Cosmopsychism: A holistic approach to the metaphysics of experience. *Philosophical Papers*, 44 (3): 389-437.

Shannon, CE (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27: 379-423 & 623-656.

Shermer, M. (2011). What is pseudoscience? *Scientific American*, September 1. [Online]. Available from: http://www.scientificamerican.com/article.cfm?id=what-is-pseudoscience [Accessed 7 August 2016].

Siegel, E. (2016). Ask Ethan: Is the universe itself alive? *Forbes*, January 23. [Online]. Available from: http://www.forbes.com/sites/startswithabang/2016/01/23/ask-ethan-is-the-universe-itself-alive [Accessed 25 February 2017].

Skrbina, D. (2007). *Panpsychism in the West*. Cambridge, MA: MIT Press.

Smolin, L. (2007). *The Trouble With Physics: The Rise of String Theory, the Fall of a Science, and What Comes Next*. New York, NY: Mariner Books.

Smolin, L. (2013). *Time Reborn: From the Crisis in Physics to the Future of the Universe*. Boston, MA: Houghton Mifflin Harcourt.

Stannard, DE (1980). *Shrinking History: On Freud and the Failure of Psychohistory*. Oxford, UK: Oxford University Press.

Stavrova, O., Ehlebracht, D. and Fetchenhauer, D. (2016). Belief in scientific–technological progress and life satisfaction: The role of personal control. *Personality and Individual Differences*, 96: 227-236.

Stoerig, P. and Cowey, A. (1997). Blindsight in man and monkey. *Brain*, 120 (3): 535-559.

Stoljar, D. (2016). Physicalism. In: Zalta, EN (ed.). *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition). [Online]. Available from: http://plato.stanford.edu/archives/spr2016/entries/physicalism/ [Accessed 26 February 2017].

Strasburger, H. and Waldvogel, B. (2015). Sight and blindness in the same person: Gating in the visual system. *PsyCh Journal*, 4 (4): 178-185.

Strassman, R. (2001). *DMT: The Spirit Molecule*. Rochester, VT: Park Street Press.

Strassman, R. *et al*. (2008). *Inner Paths to Outer Space*. Rochester, VT: Park Street Press.

Strawson, G. *et al*. (2006). *Consciousness and Its Place in Nature*. Exeter, UK: Imprint Academic.

Swedenborg, E. (2007). *Heaven and its Wonders and Hell*. Charleston, SC: BiblioBazaar.

Tarlaci, S. and Pregnolato, M. (2016). Quantum neurophysics: From non-living matter to quantum neurobiology and psychopathology. *International Journal of Psychophysiology*, 103: 161-173.

Tarnas, R. (2010). *The Passion of the Western Mind*. London, UK: Pimlico.

Taylor, JB (2009). *My Stroke of Insight: A Brain Scientist's Personal Journey*. New York, NY: Viking.

Taylor, K. (1994). *The Breathwork Experience: Exploration and Healing in Nonordinary States of Consciousness*. Santa Cruz, CA: Hanford Mead.

Taylor, C. (2007). *A Secular Age*. Cambridge, MA: Harvard University Press.

Tegmark, M. (2014). *Our Mathematical Universe: My Quest for the Ultimate Nature of Reality*. New York, NY: Vintage Books.

Tillich, P. (1952). *The Courage To Be*. New Haven, CT: Yale University Press.

Tittel, W. *et al*. (1998). Violation of Bell inequalities by photons more than 10 km apart. *Physical Review Letters*, 81 (17): 3563-3566.

Tongeren, DR van and Green, JD (2010). Combating meaninglessness: On the automatic defense of meaning. *Personality and Social Psychology Bulletin*, 36 (10): 1372-1384.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, 5 (42), doi: 10.1186/1471-2202-5-42.

Treffert, D. (2006). *Extraordinary People: Understanding Savant Syndrome*. Omaha, NE: iUniverse, Inc.

Treffert, D. (2009). The savant syndrome: An extraordinary condition. A synopsis: Past, present, future. *Philosophical Transactions of the Royal Society B*, 364 (1522): 1351-1357.

Tsuchiya, N. *et al*. (2015). No-report paradigms: Extracting the true neural correlates of consciousness. *Trends in Cognitive Science*, 19 (12): 757-770.

Urgesi, C. *et al*. (2010). The spiritual brain: Selective cortical lesions modulate human self-transcendence. *Neuron*, 65: 309-319.

Vaillant, GE (1992). *Ego Mechanisms of Defense: A Guide for Clinicians and Researchers*. Washington, DC: American Psychiatric Press.

Valsiner, J. (1998). *The Guided Mind*. Cambridge, MA: Harvard University Press.

Vandenbroucke, A. *et al*. (2014). Seeing without knowing: Neural signatures of perceptual inference in the absence of report. *Journal of Cognitive Neuroscience*, 26 (5): 955-969.

Varela, FJ, Thompson, E. and Rosch, E. (1993). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Vazza, F. and Feletti, A. (2017). The strange similarity of neuron and galaxy networks: Your life's memories could, in principle, be stored in the universe's structure. *Nautilus*, July 20. [Online]. Available from:

http://nautil.us/issue/50/emergence/the-strange-similarity-of-neuron-and-galaxy-networks [Accessed 21 July 2017].

Walls, LD (2003). *Emerson's Life in Science: The Culture of Truth*. Ithaca, NY: Cornell University Press.

Watts, A. (1989). *The Book: On the Taboo Against Knowing Who You Are*. New York, NY: Vintage Books.

Webster, R. (1995). *Why Freud Was Wrong: Sin, Science, and Psychoanalysis*. New York, NY: Basic Books.

Wegner, DM (2002). *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.

Weihs, G. *et al*. (1998). Violation of Bell's inequality under strict Einstein locality conditions. *Physical Review Letters*, 81 (23): 5039-5043.

Westen, D. (1999). The scientific status of unconscious processes: Is Freud really dead? *Journal of the American Psychoanalytic Association*, 47 (4): 1061-1106.

Whinnery, J. and Whinnery, A. (1990). Acceleration-induced loss of consciousness: A review of 500 episodes. *Archives of Neurology*, 47 (7): 764-776.

Whitehead, AN (1947). *Essays in Science and Philosophy*. New York, NY: Philosophical Library.

Wigner, E. (1960). The unreasonable effectiveness of mathematics in the natural sciences. *Communications on Pure and Applied Mathematics*, 13 (1): 1-14.

Windt, JM and Metzinger, T. (2007). The philosophy of dreaming and self-consciousness: what happens to the experiential subject during the dream state? In: Barrett, D. and McNamara, P. (eds.). *The New Science of Dreaming*. Westport, CT: Praeger, pp. 193-247.

Windt, J., Nielsen, T. and Thompson, E. (2016). Does consciousness disappear in dreamless sleep? *Trends in Cognitive Sciences*, 20 (12): 871-882.

Yetter-Chappell, H. (2018). Idealism without God. In: Goldschmidt, T. and Pearce, K. (eds.). *Idealism: New Essays in Metaphysics*. Oxford, UK: Oxford University Press.

Zemach, E. (2006). Wittgenstein's philosophy of the mystical. In: Copi, IM and Beard, RW (eds.). *Essays on Wittgenstein's Tractatus*. London, UK: Routledge, pp. 359-376.

Zicheng, H. (2006). *Vegetable Roots Discourse: Wisdom from Ming China on Life and Living: Caigentan*. Berkeley, CA: Shoemaker & Hoard.

Summary

This dissertation elaborates on a modern, analytic version of the ontology of idealism, according to which (a) phenomenal consciousness, as an ontological category, is fundamental; and (b) everything else in nature can ultimately be reduced to, or grounded in, patterns of excitation of phenomenal consciousness. It posits a reduction base consisting of a single element: spatially unbound, universal phenomenal consciousness. Its key challenge is then to explain how the seemingly distinct phenomenal inner lives of different subjects of experience can arise within this fundamentally unitary phenomenal field. This is sometimes called the "decomposition problem" in the literature and it is the core problem this dissertation attempts to tackle. Along the way, a variety of other challenges are addressed, such as: how we can reconcile idealism with the fact that we all inhabit a common external world; why this world unfolds independently of our personal volition or imagination; why there are such tight correlations between measured patterns of brain activity and reports of experience; etc.

The core of this dissertation consists of five papers published in academic journals. They are reproduced here, in chapters 2 through 6, without any change of substance.

Chapter 2 discusses what is perhaps the root of key unresolved problems in contemporary analytic philosophy: the tendency to try to make sense of nature by replacing concrete observations with theoretical abstractions. Such attempts often consist of mere word games, played in thought with a rich and shifting phantasmagoria of concepts. Chapter 2 attempts to make these word games explicit. It also suggests more epistemically reliable lines of reasoning that avoid unnecessary conceptual abstractions.

By pursuing these more reliable lines of reasoning, Chapter 3—the core of this dissertation—elaborates on an analytic formulation of idealism. It can be summarized thus: there is only universal phenomenal consciousness. We, as well as all other living organisms, are *dissociated alters* of this universal consciousness, analogously to how a person with Dissociative Identity Disorder (DID) manifests multiple disjoint centers of subjectivity also called 'alters.' We, and all other living organisms, are surrounded by the *trans*personal phenomenal activity of universal consciousness, which unfolds beyond the dissociative boundary of our respective alter. The inanimate world we perceive around us is the 'extrinsic appearance'—i.e. the phenomenal image imprinted from *across* our dissociative boundary—of this activity. The living organisms we share the world with are the extrinsic appearances of other alters.

Chapter 4 lists possible objections against such a consciousness-only ontology and tackles them one by one. It attempts to show that they are often based on logical fallacies such as question-begging, unexamined assumptions, misunderstandings of the implications of analytic idealism, etc.

One objection is exceptional because it poses some legitimate difficulties: a necessary implication of the ontology proposed in Chapter 3 is that an

organism's metabolism—*all* of it—is the extrinsic appearance of the organism's conscious inner life. This is reasonable enough for certain patterns of brain activity known to correlate with experiences accessible through introspection, but what about so-called 'unconscious' mental processes and metabolism beyond the brain, such as e.g. liver and kidney function?

Chapter 5 argues that, despite appearances to the contrary, there is no clear reason to believe that *any* mental process is truly unconscious. Instead, it attempts to show that there are, in fact, very good reasons to think that what we regard as unconscious mental processes correspond merely to an *illusion* of unconsciousness, which results from dissociative states or lack of metacognition. And once these two mechanisms—dissociative states and lack of metacognition—are identified, they can explain why experiences corresponding to areas of the living body beyond the nervous system can't be accessed through introspection.

Finally, Chapter 6 compiles and discusses a broad list of instances of brain function *impairment* that are accompanied by *enrichment* of conscious inner life and an *expansion* of one's sense of identity. The list includes cases as varied as asphyxiation, physical trauma to the head, the consumption of psychoactive substances that dampen brain activity, etc. Such correlations between impaired brain function and enriched conscious inner life are at least counterintuitive under the mainstream physicalist notion that conscious inner life is constituted or generated by brain activity. Under analytic idealism, on the other hand, they are to be expected: if normal brain function is part of the extrinsic appearance of a *dissociated* alter of universal consciousness, then *some* forms of reduction or impairment of normal brain function should be the extrinsic appearance of a reduction or impairment of the *dissociation*. And, of course, from a first-person perspective a reduction of dissociation must be experienced as an enrichment of conscious inner life: reintegrated memories, the recovery of a broader sense identity, renewed access to previously dissociated insights and emotions, reintegration of previously dissociated skills, etc. Contrary to physicalism, analytic idealism can thus not only accommodate, but also make sense of, the evidence discussed in Chapter 6.

Samenvatting

Dit proefschrift stelt voor een moderne, analytische versie van de ontologie van idealisme. Volgens deze zo genoemde 'analytische idealisme,' (a) fenomenaal bewustzijn, als een ontologische categorie, is fundamenteel; en (b) elk waarneembaar aspect van de natuur kan uiteindelijk worden gereduceerd tot, of geaard in, patronen van opwinding van fenomenaal bewustzijn. Het poneert een reductiebasis bestaande uit een enkel element: ruimtelijk ongebonden, universeel fenomenaal bewustzijn. De belangrijkste uitdaging is dan om uit te leggen hoe de schijnbare gescheiden fenomenale innerlijke levens van verschillende subjecten zich kunnen voordoen binnen dit fundamenteel eenvormige fenomenale veld. Dit wordt in de literatuur soms het 'decompositieprobleem' genoemd en het is het kernprobleem dat dit proefschrift probeert aan te pakken. Onderweg komen er een aantal andere uitdagingen aan de orde, zoals: hoe we analytische idealisme kunnen verzoenen met het feit dat we allemaal een gemeenschappelijke externe wereld bewonen; waarom deze wereld zich ontvouwt onafhankelijk van onze persoonlijke wil of verbeeldingskracht; waarom er zulke strakke correlaties zijn tussen gemeten patronen van hersenactiviteit en ervaringsrapporten; enz.

De kern van dit proefschrift bestaat uit vijf papers gepubliceerd in academische tijdschriften. Ze worden hier weergegeven, in hoofdstukken 2 tot en met 6, zonder enige verandering van inhoud.

Hoofdstuk 2 bespreekt wat misschien de oorzaak is van de belangrijkste onopgeloste problemen in de hedendaagse analytische filosofie: de neiging om te proberen de natuur toe te lichten door concrete waarnemingen te vervangen door theoretische abstracties. Zulke pogingen bestaan vaak uit louter woordspelletjes, gespeeld in gedachten met een rijke en veranderende fantasmagorie van concepten. Hoofdstuk 2 probeert deze woordspelletjes expliciet te maken. Het suggereert ook meer epistemische betrouwbare lijnen van redenatie die onnodige conceptuele abstracties vermijden.

Door deze meer betrouwbare redenaties na te streven, gaat Hoofdstuk 3—de kern van dit proefschrift—in op een analytische formulering van idealisme. Het kan als volgt worden samengevat: er is alleen universeel fenomenaal bewustzijn. Wij, evenals alle andere levende organismen, zijn gedissocieerde 'alters' van dit universeel bewustzijn, analoog aan hoe een persoon met Dissociatieve Identiteitsstoornis meerdere disjuncte centra van subjectiviteit manifesteert, ook wel 'alters' genoemd. Wij, en alle andere levende organismen, worden omringd door de transpersoonlijke fenomenale activiteit van universeel bewustzijn, die zich ontvouwt voorbij de dissociatieve grens van ons respectieve alter. De levenloze wereld die we om ons heen waarnemen is de 'extrinsieke verschijning'—d.w.z. het fenomenale beeld dat is afgedrukt vanuit onze dissociatieve grens—van deze activiteit. De levende wezens waarmee we de wereld delen zijn de extrinsieke verschijningen van andere alters.

Hoofdstuk 4 somt mogelijke bezwaren op tegen zo een ontologie met alleen bewustzijn en pakt ze een voor een aan. Het probeert te laten zien dat de bezwaren vaak gebaseerd zijn op logische drogredenen zoals cirkelredenering,

foutieve impliciete veronderstellingen, misverstanden over de implicaties van analytische idealisme, enz.

Eén bezwaar is uitzonderlijk omdat het enkele legitieme moeilijkheden oplevert: een noodzakelijke implicatie van de in Hoofdstuk 3 voorgestelde ontologie is dat de hele metabolisme van een organisme de extrinsieke verschijning is van de bewuste innerlijke leven van het organisme. Dit is redelijk voor bepaalde patronen van hersenactiviteit waarvan bekend is dat ze correleren met ervaringen die toegankelijk zijn via introspectie, maar hoe zit het met de zogenaamde 'onbewuste' mentale processen en het metabolisme buiten de hersenen, zoals b.v. lever- en nierfunctie?

Hoofdstuk 5 stelt dat er, ondanks het schijn van het tegendeel, geen duidelijke reden is om te geloven dat een mentaal proces ooit echt onbewust is. In plaats daarvan probeert het te laten zien dat er in feite zeer goede redenen zijn om te denken dat wat we beschouwen als onbewuste mentale processen alleen maar overeenkomen met een illusie van onbewustzijn, die het gevolg is van dissociatieve toestanden of gebrek aan metacognitie. En zodra deze twee mechanismen—dissociatieve toestanden en gebrek aan metacognitie—worden geïdentificeerd, kunnen ze verklaren waarom ervaringen die corresponderen met gebieden van het levende lichaam voorbij het zenuwstelsel niet toegankelijk zijn via introspectie.