Panpsychism and Mental Monism: Comparison and Evaluation

Peter B. Lloyd
School of Computing, University of Kent, Canterbury, England
peter@peterblloyd.com

Abstract

In the contemporary debate on the Cartesian mind-body problem, the theories of panpsychism and mental monism are receiving growing attention as candidate accounts of the conscious mind. Panpsychism asserts that consciousness pervades a spatially extended physical universe, while mental monism asserts that consciousness exists in a purely mental domain, and denies the reality of physical substance and space. In this paper, I compare the two theories and evaluate whether each is true. They seem fundamentally contradictory, but I propose that they are closer than they seem. I argue that their apparent difference rests on a category-mistake (in Ryle's term), a misplaced reification of the physical. My claim is that physics is topic neutral (in Foster's use of the term) and that consequently situating consciousness in a purely notional physical universe adds nothing to our concept of reality beyond what mental monism gives us. I argue that: First, when we see physics correctly through the lens of topic neutrality, the physical part of panpsychism disappears and the theory reduces to mental monism. Second, that mental monism is true as a general thesis. That is, reality ultimately consists of nothing but conscious minds. Third, that panpsychism is false and that a Berkeleyan flavor of mental monism resolves the Cartesian mind-body problem, which in recent years has received attention in the form of Chalmers' Hard Problem. Besides providing a resolution of this philosophical problem, mental monism also potentially offers a platform on which to build a naturalistic account of psi phenomena.

Contents

1. Introduction 3
2. Comparison: Panpsychism *v* Mental Monism 6
2.1 Topic Neutrality 6
2.2 Private-Language Argument 7
2.3 Topic Neutrality of Physics 9
2.4 A Normalized Formulation of Panpsychism 10
2.5 Generalized Panpsychism 13
2.6 Non-individuation of subjects in mental monism 14
2.7 A Naïve Formulation of Mental Monism 16
2.8 A Normalized Formulation of Mental Monism 18
3. Evaluation 1: Mental Monism is True 19
3.1 First limb: Consciousness is not Grounded in Physics 21
3.2 Constructs: A Prerequisite for the Second limb 23
3.3 Second Limb: A Mind-Independent World is not Referenceable 26
3.4 Conclusion for Evaluation 1: Mental Monism is True 29
3.5 Corollary 1: Mental Monism Entails the Metamind 29
3.6 Corollary 2: Personal Minds are Subsets of the Mental Universe 30
3.7 Corollary 3: Physical Events Cannot Cause Mental Events 32
3.8 Corollary 4: The Brain is a Virtual Interface 33
4. Evaluation 2: Panpsychism is False 33
4.1 Argument from Relativity I: Consciousness is Not Spatiotemporal 34
4.2 Argument from Relativity II: Physical v Mental Simultaneity and Sequence 36
4.3 Corollary 1: Non-Simultaneity v Non-Unified Metamind 38
4.4 Corollary 2: Hypothesis of Universal Mental Time 38
4.5 Argument from Excision 41
4.6 Spatially Separable Panpsychism is Untenable 43
5. General Conclusion 44
6. Postrequisite: Hierarchical Language-Games 44
7. Discussion: Potential Applications 48
8. References Cited 49

1. Introduction

The discoveries of physics over the past three centuries should have inoculated us against the notion that common sense is a sure-fire guide to understanding reality. We should have learned this from Galileo's two cannonballs (light and heavy falling together), or from Eddington's two tables (one a solid slab of wood, the other a vibrant cloud of microphysical particles), or from Einstein's relativity, or from quantum mechanics. Common sense tells me that I am a conscious mind resting inside a lump of matter organized as a brain, and that the wooden chair I am sitting on, despite being made of similar molecules, cannot conceivably be conscious. Careful reflection on the inadequacy of this common-sense picture, aided by a survey of the recent literature of consciousness studies (e.g. Chalmers 1996), should disabuse us of the naïve notion that reality is anything like this. In fact, whatever the correct picture turns out to be, we can safely assume it will fly in the face of common sense. Therefore, we should park our common sense and assess theories of consciousness on a careful logical examination of their merits. Let us pay no heed to the fact that panpsychism and mental monism obviously breach naïve common sense.

There has been a resurgence of interest in panpsychism, and even in subjective idealism (e.g. Blamauer 2011, Bruntrup & Jaskolla 2017, Goldschmidt & Pearce 2017, Chalmers 2017b), despite the entrenched disdain for both in mainstream science and philosophy. I will examine the relationship between these two theories, and assess their validity, but I will not examine *physical* monism here, as the literature already contains many critiques of it. In contrast, the comparative critique of panpsychism and mental monism has received little attention, and is timely in view of the resurgence of both theories.

Besides endeavoring to resolve the long-standing Cartesian mind-body problem, mental monism or panpsychism may offer a platform for a naturalistic account of psi phenomena (e.g. Lloyd 1999b, Kastrup 2014). There is, however, much work needed to get from philosophy to psi, and a consideration of this is reserved for a follow-up paper. Here, I examine only the philosophical foundations.

The starting point for both schools of thought is that conscious experience is a bona fide component of the real world, but the next point—where they diverge—is how they integrate consciousness with the everyday world. Panpsychism takes for granted a mash-up of physics and naïve realism. It takes physical space and substance

as real existents within which consciousness can exist. Conscious experience, on this view, is somehow spatially located, co-extensively with physical stuff. To account for how conscious experience is packaged into these lumps of matter, the topic neutrality of physics is wheeled in, like a Trojan horse, to allow conscious experience to penetrate physical substance. Panpsychism's story is that physics provides a detailed account of the extrinsic relations between physical things, but says zilch about any intrinsic qualities of physical stuff. Conscious awareness, on the other hand, is replete with intrinsic qualities but is silent on the location of those qualia in physical space. So, the theory marries the two up: consciousness fills out the physical stuff with intrinsic qualities, and conversely physics gives conscious experience a spatial home to sit in. What this account lacks, however, is the full unfolding of the ramifications of the topic neutrality of physics. The trouble is this: if physics is nothing but an assemblage of extrinsic relations, then the physical world is a castle in the air: it melts into nothing. Extrinsic relations need something real to relate, for otherwise they are purely notional. Thus, physics cannot provide the grounded home that consciousness is supposed to be looking for. The brain, instead of being a solid computing machine for the conscious mind to inhabit as if it were software enthroned in hardware, is itself just vaporware, less concrete than the mind itself. Panpsychism is playing bait and switch: it holds up the physical world as a home for consciousness, but then allows consciousness into the physical world only by evacuating that world of a substantive core. If physics consists in nothing but extrinsic relations between things whose intrinsic nature is conscious, then—if we may call a spade a spade—that is a species of mental monism. What we are really saying is that (a) reality is wholly constituted by conscious minds (that's the mental monism bit), and that (b) those conscious minds and their parts are structured in such a way that they can be described by the extrinsic relations of physics (that's the window-dressing of panpsychism).

That, in a hand-waving precis, is the burden of this paper's first part. In what follows, I shall unpack these ideas more rigorously. In the second part I shall argue that, on the one hand, consciousness is not reducible to physics and, on the other hand, a mind-independent reality is not something we could ever know or make reference to, and hence the only reality that we can coherently assert the existence of is a purely mental one. In the third part, I shall argue that a unified stream of

consciousness cannot be tied in with a spatially distributed physical system, therefore no form of panpsychism can hold. This leads to the conclusion that a Berkeleyan form of mental monism, rather than a panpsychist form, is correct.

Idealism has several varieties (e.g. subjective v absolute idealism, micro- v macro- v cosmo-idealism). Chalmers (2018) offers a broad taxonomy of idealism, but acknowledges that the neat categories are not watertight ("Berkeley looks like a macro-idealist, at least before God enters his picture") and he classifies Advaita Vedanta as cosmic idealism, even though it asserts that Brahman is non-different from each person's Atman, so Chalmers has to call it a macro/cosmic idealism, rather like the later Berkeleyan leanings. The present paper attempts no such overview, but addresses certain theories that are of interest, namely Berkeleyan subjective idealism and Śankaran Advaita Vedanta, which turn out to be conceptually related despite being historically and culturally remote.

Panpsychists' ambivalence towards idealism is illustrated by two interviews with Galen Strawson, a leading proponent of panpsychism. Strawson (2015) said, "This view [panpsychism] has absolutely nothing to do with Berkeley's idealism. The buildings are out there. … The building is not in any sense an idea in someone's mind as in Berkeley's scheme," but in Strawson (2018) we hear an opposite view – Q: "The pure panpsychism, the strong form, does that move in the direction of so-called idealism?" Strawson: "I have trouble with the word 'idealism' as I prefer to say it is a form of 'mentalism', but as the word 'idealism' is used, broadly speaking, yes it does." I submit that the source of this ambivalence is that, despite the contrary appearance, panpsychism logically entails mental monism. Chalmers also notes the ambivalence between Strawsonian panpsychism and idealism: "Strawson (2006) looks like a micro-idealist but has more sympathy with cosmic idealism."

Finally, a caveat. In discussing these topics, we are no longer in the salubrious aisles of the market place of ideas: we are in the badlands of philosophy. The enterprise of physical science despises panpsychism as a throwback to pre-modern animism, and is unimpressed by Chalmers' (2015) warning that "there is good reason to think that any view of consciousness must embrace some counterintuitive conclusions." However outré panpsychism may be, though, my observation since Lloyd (1999a,b) has been that almost everybody despises that dread philosophy that dare not speak its name: Berkeleyan subjective idealism—notwithstanding a few

recent positive signs, such as Chalmers' (2018) small mercy on subjective idealism, that "there is a non-negligible probability that it is true."

Terminology. Mental monism is the theory that reality consists of minds only. It has historically been called 'idealism', 'subjective idealism', and 'mentalism', and I will use those terms interchangeably when citing other authors. There is sometimes a debate about whether a microscopic particle's putative consciousness constitutes a 'mind' as it is so simple and cannot plausibly be expected to think. This is a sterile debate over how we wish to define our terms. For the purposes of this paper, I shall use 'mind' to denote any 2-tuple <C,S> where C is a time-varying system of phenomenal content and S is a fixed subject. The 'subject' is the agent of the acts of experience and volition; the 'phenomenal content' comprises the objects of those acts. These terms may be employed differently elsewhere, but this is how I will use them here. A 'system of phenomenal content', C, is a 3-tuple <E,V,A> comprising sets of experiences E, volitions V, and relations between them, A—for example, shifting attention leftwards from a red patch, and finding a blue patch. Experiences comprise qualitative content, whose minimal discernible elements are 'qualia'. This definition places no limit on scale: a 'mind' might be so tight as to contain a single, unchanging quale (micro-panpsychism), or it might encompass the whole of reality (cosmo-panpsychism).

2. Comparison: Panpsychism *v* Mental Monism

I will examine the relationship between panpsychism and mental monism in two theses: (a) The topic neutrality of physics entails that in panpsychism minds are not really 'in' space. My approach here will be to reduce both panpsychism and mental monism to 'normalized' formulations, which reveal their similarities more clearly. (b) In mental monism, the non-individuation of subjects entails that the mental counterpart of each object can be construed as a mind. This, I argue, reveals a structural conformity between panpsychism and mental monism. I will conclude that panpsychism is a special case of mental monism.

2.1 Topic Neutrality

Topic neutrality is a key to this approach to the mind-body problem. Chalmers (2013, §3) attributes the insight of physics' topic neutrality to Russell (1927), defining it thus: "physics reveals the relational structure of matter but not its intrinsic nature." Ryle (1954) introduced the term 'topic neutral' but misapplied it by saying that

propositions about consciousness are topic neutral, on grounds akin to Wittgenstein's claim that consciousness is, in effect, ineffable (1953, §243). Thirty years later, Foster (1982) got it right, stating that physics must be topic neutral, not propositions about consciousness.

Since the grounds for Ryle's 'inverted' topic neutrality are essentially those articulated most clearly by Wittgenstein, we will look at his philosophy.

Wittgenstein (1953, §243) introduced the concept (but not the term) of 'effective ineffability' with his thought-experiment of the beetle in the box, which marked his notorious private-language argument against the possibility of saying anything meaningful about conscious experience. Essentially this is the conundrum that since I can't know what you actually see when you look at something red (maybe red looks blue to you), so Wittgenstein says that "red" and "blue" cannot refer to anything; and more generally that all the terms that seem to denote conscious experience can't actually refer to anything at all.

It is crucial to get topic neutrality the right way round, and to do that we must first defeat Wittgenstein's private-language argument, before analysing mental monism and panpsychism.

2.2 Private-Language Argument

Outside the philosophical community, Wittgenstein's conclusion against private languages seems egregiously counter-intuitive, but among philosophers it still commands much support. He illustrated it with his beetle-in-the-box thought-experiment, which is easy to grasp, but so opaque that its shadow still darkens the study of consciousness today.

Its experimental set-up gives each person a small box with an indeterminate object, the 'beetle', inside it. Wittgenstein rigs his thought-experiment to lead us by the nose to nod enthusiastically that the beetle cannot be referred to in any language because it is knowable only by its owner, and effectively ineffable. His constraint is (C_B): "No one can look into anyone else's box, and everyone says he knows what a beetle is only by looking at his beetle" (§293). Under this constraint, Wittgenstein argues we cannot agree on any terms referring to parts or properties of a beetle because that would require shared observations in which to ground consensual definitions, but constraint C_B prevents shared observations. Without agreed terms, we

can communicate nothing about a beetle-in-the-box. Likewise, C_M says that nobody can see inside your mind, hence we can communicate nothing about a mind-in-the-brain.

This argument is both stronger and weaker than Wittgenstein requires, for it undermines not just private language but all language. For example, if a beetle crawls across the pavement, I can shout, "That beetle is green!" and thereby communicate about the beetle. The logic of Wittgenstein's argument, however, casts doubt on whether people see the same beetle (maybe a Cartesian demon feeds false images into each person's mind), and whether other people hear my words (the demon makes them hear something else), and hence implies that public language is impossible just as private language is. This nihilistic conclusion—that we should stop pretending to communicate—reveals that Wittgenstein's private-language argument is stronger than he wants.

As long as there are no Cartesian demons at play, linguistic communication can work, and this is enough to defeat Wittgenstein's argument that such communication can never function. Furthermore, the actual existence of a Cartesian demon is such a wildly implausible hypothesis as to be dismissed as a practical possibility. Indeed the vast weight of circumstantial evidence indicates that public and private communication both work and that our universe is at least largely free of Cartesian demons. In addition, the uniformity of human sense organs suggests a uniformity of perception. Hence, when I say "I see red," I almost certainly know what I mean, and you probably do too. At most, Wittgenstein could argue that we risk mistakes in private languages; but we also risk mistakes in public languages. In other words: communication is fallible. Thus, Wittgenstein's private-language argument proves to be weaker than he wants.

Therefore, statements about experiences (such as, "I see red") are not an anemic formal system, as Wittgenstein implied, but a full-blooded, albeit fallible, means of communication. This knocks out Ryle's topsy-turvy idea of topic-neutrality. Having thus established the *meaningfulness* of stating that we have unmediated acquaintance with the phenomenal content of our conscious minds, whereas we never have such acquaintance with physical things, now we may address the *correctness* of that statement. Thus, we return—from our Wittgensteinian detour— to the question of whether physics is topic neutral.

2.3 Topic Neutrality of Physics

As I mentioned above, the notion of topic neutrality stems from the ideas of intrinsic and extrinsic properties. By the 'extrinsic properties' of a physical thing we mean those that it possesses by virtue of its relations to other physical things. An obvious example is velocity: a body can travel at x ms^{-1} only in relation to some other body. Every other physical property—mass, energy, charge, field strength, and so on—is extrinsic in the same sense. A gravitational field is defined entirely by its effect on massive bodies: there is nothing else to the field over and above its extrinsic relation to other physical things. If you were to imagine a purely intrinsic physical property, call it 'hylasity', then it would *ipso facto* not interact with other physical things, since a disposition to act upon other physical things would, by definition, be an extrinsic relation. Pure hylasity would have leave no physical trace and could never be measured. It would be indistinguishable from a fiction. That is why all the properties that physics actually does handle are extrinsic relations, as it is only through extrinsic interactions that observation becomes a possibility, and physics constitutionally has nothing to say about things it cannot observe.

Physics is built from formulae stating the structure and dynamics of entities that are defined analytically through the structure and dynamics of smaller entities, and ultimately in undefined fundamentals (such as mass and charge). Physics formulae have no terms for the intrinsic quality, the 'quiddity', of these entities, nor any theoretical apparatus for handling qualitative properties, nor any third-person experimental procedures for measuring them. Therefore, physics is wholly about extrinsics and is silent about any putative intrinsics, which is to say it is topic neutral. Consciousness, in contrast, is constituted by an assemblage of intrinsic qualities, and propositions about consciousness are manifestly not topic neutral. Several philosophers have made this point in consciousness studies, for example Rosenberg (2004).

Physicists mostly hold a pre-philosophical view of the 'quiddity' of physical entities, and have a vague notion that they are studying a mind-independent reality. They insist that they are measuring 'real' things in the laboratory, and that therefore physics cannot be just a closed, formal system. Take any physics textbook, though, and peruse its equations: you will find no term for quiddity, because physics has no need of it. Physics as it is practiced—as opposed to how it is fantasized—is topic

neutral.

Some philosophers have tried to insert wiggle room between the qualityless world of physics and the qualityful world of the mind. Thus, 'panprotopsychism' says that physical things have 'protophenomenal' properties that are not phenomenal but give rise to the phenomenal (Chalmers 2013, §1). Similarly, 'neutral monism' claims that basic things possess proto-physical and proto-phenomenal properties that yield physical and phenomenal properties, respectively. Any such proto-phenomenal properties are irrelevant to the present philosophical discussion. Either we have direct acquaintance with them (that is to say, they are included in the phenomenal content of our conscious minds), or not. In the latter case, all terms are analytical and ultimately resolve into undefined fundamental properties such as mass, charge, and whatever novel proto-phenomenal properties are conjured up. This paper's reasoning treats any topic-neutral thing in the same way, and would therefore also apply to proto-phenomenal things. For clarity, I shall consider only physical things below under the rubric of the topic neutral. In the unlikely event of anyone actually discovering protophenomenal things in the future, the same arguments would apply.

2.4 A Normalized Formulation of Panpsychism

Panpsychism's core idea, that reality is made up of basic units having a phenomenal interior and a physical exterior, can be articulated in two ways that ultimately convey the same sense.

Chalmers (2015) laid down a common formulation: "Panpsychism [is] the view that fundamental physical entities have conscious experiences."[1] Nagasawa and Wager (2016) similarly wrote that the physical world "is mentally—and in particular 'phenomenally'—propertied." This presumes that physical entities are real things that can possess qualitative mental properties. It can be contrasted with the thesis that Chalmers calls micro-idealism, which presumes that minds are real things that can possess micro-physical properties. Thus:

(Pφ) (Micro-)panpsychism: The basic units of reality are fundamental physical entities, and some or all of these possess 'elemental' minds (not composed of smaller minds).

(Pψ) Micro-idealism: The basic units of reality are elemental minds, and some of these possess the extrinsic relations of fundamental physical entities.

To express the basic intuition without saying whether the physical or mental units are fundamental, we can use this:

(Pυ) Neutral panpsychism:

- There is a one-to-one mapping between fundamental physical entities and elemental minds.
- The physical properties of those physical entities are constituted by extrinsic relations between such entities.
- The intrinsic properties of minds are phenomenal mental qualities.

This captures the substantive intention of panpsychism but remains neutral about the ontology, presuming primacy of neither physical nor mental things. Most panpsychists have no preference for Pφ or Pψ, and so Pυ adequately expresses their opinion. For example, the Wikipedia entry on panpsychism uses Pφ and Pψ interchangeably: "In philosophy, panpsychism is the view that consciousness, mind, or soul (psyche) is a universal and primordial feature of all things. Panpsychists see themselves as minds in a world of mind."

On the other hand, some philosophers specifically advocate Pφ, and insist that the basic units of reality are situated in physical space, and hence that conscious minds are likewise situated in space (e.g. Strawson 2013).

Here, I shall argue that mental locality is impossible. In summary, the argument is as follows. As minds do not partake of the extrinsic relations that make up physics, they cannot share the kind of spatial relations that physical bodies have. On the other hand, as physics is topic-neutral, the only space that it knows about is the one comprising those extrinsic spatial relations. There is no substantive space described by physics, over and above those relations. Therefore, there is simply no physical space that minds are capable of inhabiting.

That consciousness is non-spatial has been somewhat neglected in the literature, as McGinn (1995) observed: "Even those who recognize that consciousness poses problems for materialism in virtue of its phenomenal character seldom acknowledge that its non-spatiality is also a major stumbling-block for materialism—despite the fact that Descartes took it (and not qualia) to be critical." McGinn offers a blanket of intuitions against the spatiality of consciousness, but it seems to me that the crucial argument is the non-co-location of conscious experience with spatial properties.

The physical properties of an object are co-located in space by virtue of being third-person-observable in the same place. Those properties are found to affect observational instruments (be they sensory organs or artificial devices) in, or on, the space that is occupied by the object.

For example, consider a doughnut. A third-person observer can measure the shape and size of the doughnut, its mass, charge, color, texture, chemical composition, and such-like—all of which are co-located in the same bounded volume of space. Its temperature is found by sticking a thermometer in it; its weight by placing it on a weighing scale; its volume by dropping it in a graduated jug of water. The phenomenal content of the doughnut, on the other hand, cannot be measured by a third-person observer. Neither by sensory organs nor by devices can an observer discover the supposed conscious quiddity of the doughnut. Therefore, there is no operational procedure for finding the phenomenal content in the same volume of space as the familiar physical properties. *Ergo* the mental quiddity has no spatial location.

Phenomenal contents' being associated with a spatially located entity does not entail the phenomenal contents' being so located themselves. To be spatially located is to be discoverable in a volume of space by a third-person observation. That cannot be done with someone else's mind, for e.g. any experience you have whilst surgically inspecting my brain is in your mind, not in mine. Thus, the supposed position of a mind is unobservable because the mind cannot partake of physical processes. Assertions that your mind is 'in' your head, or is hovering six inches above it, or is even on the Moon, are therefore equally untestable, unfalsifiable, and frankly meaningless.

Wittgenstein (1958, pp 7-8) and Ayer (1963, p 103) dismissed the meaningfulness of literally ascribing a spatial position to a mental experience, and Wittgenstein noted that any ascription would be a figurative usage (as, for example, one can have a pain 'in' a phantom limb in the sense that it is projected into the body image of the limb).

The spatial relation between physical entities is one more extrinsic relation. Just as particles might have different electrical charges, different masses, different kinetic energies, so they also have different spatial coordinates and velocities in relation to each other, and to other physical particles. To ascribe any of these physical observables to a mind that is associated with that particle is to commit a category-

mistake: a mind is not something that can have a charge, or a mass, or a spatial location or velocity.

For this reason, I reject Strawson's insistence on Pφ. Excluding mental localization from Pφ makes Pφ versus Pψ a distinction without a difference, so I take Pυ as expressing what we mean by panpsychism.

Now, Chalmers (2018, p 5) defines micro-idealism as "the thesis that concrete reality is wholly grounded in micro-level mentality: that is, in mentality associated with fundamental microscopic entities (such as quarks and photons)"[2], and then asserts that "micro-idealism entails panpsychism". There is, however, an ambiguity in this languaging of physics-oriented forms of idealism: for example, the phrase "mentality associated with fundamental microscopic entities" could be read as idealism or as panpsychism. Saying that elemental mentality is 'associated' with microphysical entities does not entail that those mental states are states of the microphysical entities. Micro-Berkeleyanism is a counter-example, in which God's mind contains an 'archetype' for each microphysical entity, and God uses the states of the assembled archetypes to render the perceptions to human subjects. (Berkeley asserts a comprehensive system of archetypes that correspond to physical things, but does not say whether they are micro or macro.)

Micro-idealism, therefore, is conceptually distinct from panpsychism, but the difference is superficial, not substantial. I shall argue below that panpsychism entails idealism, and in fact that micro-panpsychism and micro-idealism are equivalent (and, later, that they are both wrong).

2.5 Generalized Panpsychism

Panpsychism's literature lacks a principled account of 'fundamental physical entities' that might define these 'ultimates' as Rosenberg (2004) and Goff (2009) called them. Do they include molecules, atoms, elementary particles, fields, or volumes of space? As a proton comprises three quarks, does it have a mind of its own, or do the individual quarks have minds? If an electron jumps to a lower orbit and emits a photon, was the electron in two minds before the jump—that of a low-energy electron and a photon—or is the photon's mind created in the moment of emission? If an elementary particle is in a superposition of two states, what is the ontological status of the associated phenomenal content? – does that particle's mind also go into superposition, as Lockwood (1989) suggested? If two electrons become entangled

and travel apart, do they have one mind or two? In an atom, does a cloud of indistinguishable electrons have one mind in total, or one mind per electron? Could big things have elemental minds? Macromolecules, microtubules, cells, organs, or organisms—in the absence of a principled account, it is anyone's guess. Chalmers (2018) raises the question of whether the concept of fundamental physical entities is valid anyway, in which case the bottom falls out of the micro-panpsychist worldview.

In view of this uncertainty, we might consider what I shall label 'generalized neutral panpsychism' (GPυ):

- There is a one-to-one mapping between some physical entities, not necessarily fundamental, and elemental minds.
- The physical properties of those physical entities are constituted by extrinsic relations between such entities.
- The intrinsic properties of those minds are phenomenal mental qualities.

As we shall see, however, nothing is gained by this move, as the reasoning still applies to any kind of panpsychism.

2.6 Non-individuation of subjects in mental monism

There are strong arguments, from Hume onwards, against the existence of the subject. Nevertheless, as Chalmers (2017a) wrote, "Wholesale eliminativism about subjects is not easy to stomach, especially for someone who is serious about phenomenal properties. These properties are defined as those characterizing what it is like to be a subject." I will make the working assumption that a subject exists. Nothing hangs on this, though. If it turns out that this is a mistake and there is no subject, my main conclusions are unaffected. The following will show that there is at most one subject, which is all we need here.

Our naïve presumption is that each mind has its own distinct subject. Certainly, Berkeley, in his earlier writings (1710, 1713), presumed this. Śankara (8th Century), in contrast, asserted an inter-mental unity of subject. As we shall now argue, that unity of subject follows logically from mental monism (see Lloyd 1999b §2.13.3).

Physicalism individuates subjects in two ways. First, each subject is deemed to be embodied in a brain, and inherits the spatiality of the brain. Hence one person's subject must be different from another's, because their subjects are in different places. Second, the subject is interlinked with the mind, and hence that mind's memories,

because the subject and the mind are supposed to reduce to the brain.

Mental monism, however, sees space as part of the virtual world that the community of minds constructs. The mind and its subject are therefore not located in space; *ergo* in mental monism subjects cannot be individuated by location. Minds, and hence subjects, exist in a non-spatial domain.

Likewise, in mental monism, the mind and its subject are not reducible to, or emergent from, the brain. Therefore, the relationship between the subject and the phenomenal content can derive only from the actual mental attributes of those two things, not from a physical substrate. Although Kastrup (2017) suggests that experiences are states of the subject, our observation is that experiences are contained in the set of phenomenal contents, while the subject itself plays only the role of that which witnesses the phenomenal content. Therefore the subject has no specific attributes by which it can be individuated from other subjects. This leads us to the counter-intuitive conclusion that subjects are not individuated, and that the subject is numerically identical in all minds.

The subject is sometimes said to 'have' experiences, as if the experiences were transient attributes of the subject, rather as size, shape, luminosity, and hue are attributes that a patch of color 'has' in the visual field. That is, however, disanalogous. A color patch' attributes are co-located in the visual field: its size, shape, luminosity, and hue visibly sit in the same place. In contrast, the subject does not lie within the sensorium and 'has' no perceptible qualities.

Mental privacy also fuels the intuition that subjects are individuated. As I can access only my own phenomenal content, not yours, my mind appears to have a subject that experiences only my phenomenal content. It is more exact to think in terms of co-mentality than subject privacy: the phenomenal content in my mind is co-mental insofar as I have access to all of it. I can shift my attention within my visual field or other part of my sensorium, and I can be aware of relations between elements of the sensorium: for example, that I am experiencing a red square inside a blue circle at the same time as hearing a beep. I cannot access, or be aware of relations with, phenomenal content in your mind. So, the co-mentality consists in phenomenal contents' being accessible by shifts of attention. We may say that a mind is a set of phenomenal content bound by access consciousness: (a) co-mentality is constituted by a mind's access to the contents of a sensorium; (b) privacy is constituted by one

mind's not having access to another's sensorium. On this view, the existence of co-mentality and its complement mental privacy is consistent with inter-mental identity of the subject.

Hence, in mental monism, there is at most one subject, common to all minds. It coincides with Śankara's doctrine, that the individual Atman is non-different from the universal Brahman. We will address how the personal mind, as opposed to subject, individuates in Corollary 2 below.

This resembles Hume's assertion that the subject is not an apprehensible item within the mind. Is subjectivity subjectless? When saying an experience exists, do we convey anything more by saying that a subject experiences it? Kastrup (2017) proposes that an experience is a state of the subject: "experience is a pattern of excitation of TWE [that which experiences, i.e. the subject]", but that presupposes a mysterious dynamics of the subject, without clarifying or explaining anything.

This paper requires that there be at most one subject. Whether that universal subject is an actual thing, or a notional artefact of the grammar of subjectivity, does not matter here.

2.7 A Naïve Formulation of Mental Monism

Mental monism says that only minds exist. The physical world is considered to be virtual and the persistent structure and dynamics of this virtual physical world are attributed to a background mechanism, which must possess the nature of a conscious mind of some sort because, by hypothesis, there is nothing else in reality besides minds. This background mind has various historical designations (Berkeley's 'God,' Śankara's 'Brahman') that come with religious baggage that is not relevant here. Therefore, I propose to refer to it as the 'metamind.' This model resolves the universe into nothing but a set of minds, namely the personal minds of people and other sentient beings, and the metamind. Intrinsic qualities are ascribed only to the phenomenal content of minds.

(The common, but ludicrous, misapprehension that Berkeley thought that the mundane world was nothing but an assemblage of ideas in our own minds can be dismissed. Berkeley was perfectly clear that there was a world 'out there', but equally clear that 'out there' meant in the mind of God (Berkeley 1713, Dialogue III, p 235).)

The internal structure and dynamics of the metamind are largely unexplored.

Berkeley posited ‘archetypes’ within God’s mind, corresponding to bodies (Berkeley 1713, Dialogue II, p 210). His analogy was that God renders the sensory appearances of bodies by referring to the archetypes, as a musician makes music by referring to a musical score (Berkeley 1710, §71). As the score has no auditory similarity to the music, so the archetypes have no sensory similarity to our perceptions. Berkeley’s decomposition of God’s mind into archetypes is consonant with the object-oriented computational architecture of human-made virtual-reality systems, where a computational ‘object’ encapsulates the data and functions that are required to implement and render the corresponding simulated object. It is a parsimonious hypothesis that Berkeley’s archetypes are likewise encapsulated entities. It may seem arrogant to suppose that God’s mind uses the same designs as human software engineers, but in fact they are both addressing the same computational problem, hence the convergent evolution of software engineer and God.

‘Monolithic mental monism’ says that the metamind is not separable into constituents corresponding to bodies; and an ‘object-oriented mental monism’ maintains the Berkeleyan thesis that the metamind is separable into encapsulated archetypes corresponding to observed physical entities. The latter is a more parsimonious model and in what follows I shall consider only object-oriented mental monism. I shall argue that panpsychism is a special case of object-oriented mental monism, which is a special case of general mental monism.

Does the hypothesis of an object-oriented metamind clash with the interconnectedness of all objects? Shani (2015) argued that the division of the physical world into objects is merely nominal, because objects are not natural kinds. For example, while my laptop is on the table, its atoms touch the table’s, and our separation of this assemblage into two objects is just an artefact of our preferred way of thinking. In addition, quantum physics shows us that everything in this mélange of atoms is entangled. So, Shani argues, there aren’t really distinct physical objects and therefore the mental world cannot naturally divide up in correspondence with objects. To a first approximation, however, the atoms in the laptop are objectively clustered: if I were to tip the table over, the laptop would fall one way, and the table another way. This is why VR physics simulations model distinct objects, and why Occam’s razor suggests that the metamind does likewise, and why the more parsimonious hypothesis is that each observed macroscopic object has a corresponding archetype.

Of course, this hand-waving argument proves nothing on its own, but it does suggest that object-oriented mental monism is more plausible than the monolithic flavor.

To clarify the relationship between panpsychism and mental monism, we will cast mental monism from the naïve formulation into the normalized form. To do this, however, we need to review the concept of a 'subject' within mental monism.

2.8 A Normalized Formulation of Mental Monism

Each of Berkeley's archetypes contains the information that the metamind requires to render the corresponding object in the observational experience of personal minds who witness the object. We do not know, we cannot tell what form that information takes, except that it must take some form of conscious experience, as that is the only contents that exists according to mental monism. So, when a tree has fallen in the forest and nobody hears it, the archetype of that tree contains the metamental thought that the tree has fallen and, when a person walks into the forest and looks at that spot, the archetype of the tree will be rendered in that person's mind as a fallen trunk lying on the ground.

That archetypal content is experienced by the subject of the metamind. It is not a problem for us that the same subject experiences each archetype because, as we have seen above, there is at most one subject in all minds. Therefore each archetype, A, can be regarded as a miniature mind, comprising phenomenal contents C_A, experienced by the common subject S. Each tuple $<C_A,S>$ is an 'object-mind', a mind that is associated with an observable physical object (but not, of course, literally 'inside' that object). Object-minds do not necessarily exhaust the metamind, which may also contain disembodied minds besides the ones embodied in brains or associated with inanimate objects. Now, we can formulate mental monism's position thus:

- There is a one-to-one mapping between observable physical entities and object-minds.
- The physical properties (of physical entities) are constituted by extrinsic relations between physical entities.
- The intrinsic properties (of object-minds) are phenomenal mental qualities.

Using these normalized formulations of the theories, we can more readily compare

them, thus:

- There is a one-to-one mapping of:
  - (in microphysical panpsychism) fundamental physical entities and elemental minds;
  - (in generalized panpsychism) some physical entities and elemental minds;
  - (in object-oriented mental monism) physical entities and object-minds.
- The physical properties (of each physical entity) are constituted by extrinsic relations between those entities.
- The intrinsic properties (of each mind) are phenomenal mental qualities.

We see that: microphysical panpsychism is a special case of generalized panpsychism, which is a special case of object-oriented mental monism, which is a special case of general mental monism. (We also see that micro-panpsychism and micro-idealism are equivalent.)

3. Evaluation 1: Mental Monism is True

In the previous section (*Comparison*) I argued that panpsychism is a special case of mental monism; in this section (*Evaluation 1*), I shall argue that mental monism is true; in the next (*Evaluation 2*), I shall argue that panpsychism is false; and I conclude that reality is a non-panpsychist form of mental monism.

I shall evaluate mental monism in two limbs: First, that consciousness is grounded in physics: my approach will be a version of Berkeley's semantic argument. Second, that a mind-independent world can be referenced.

Anti-physicalism has an extensive literature, a survey of which is outside the scope of this paper. One of the most widely used of contemporary arguments is Chalmers' (2015) use of conceivability to attack the notion that consciousness is grounded in physics (which he defined thus: "truths about consciousness are grounded in physical truths if all truths in the first set obtain wholly in virtue of truths in the second set obtaining"). His conceivability argument has less power than Berkeley's semantic argument, which attains the same conclusion more robustly. (Chalmers mentions in passing the *knowledge argument* and the *structure and dynamics* argument, which are cousins of the *semantic argument*.)

The conceivability argument says that we can conceive of a universe that is physically identical to our own, but contains no consciousness. People in this hypothetical universe behave as we do, but they are so-called 'philosophical zombies'. If we can conceive a world that is physically identical to ours, but devoid of consciousness, then we can infer that consciousness is not a necessary concomitant of the physical workings of the brain, and so consciousness is not grounded in physics. The argument's weakness is its assumption that we know enough about a complete physics to tell whether a zombie universe is genuinely conceivable. In particular, its premise requires that zombies have the same debates about consciousness that we have. They, too, ask what it is like to be a bat, and struggle to understand how conscious minds emerge from insentient matter. If, however, zombie philosophers were to discover, and assert, that there is nothing it is like to be a zombie, then Chalmers' argument collapses. Conversely, if zombies do debate consciousness, then our own such debates are a sham as they too stem from neural activity only, not consciousness. Either way, the conceivability argument fails.

Berkeley's semantic argument (Berkeley 1713, Dialogue III, p 222), articulated in a modern form (Lloyd 1999a, 2006), captures this intuition more rigorously, and is roughly equivalent to Foster's (1982) argument against physicalism. The argument presented here follows Berkeley's central thinking on the matter, which is different from his so-called 'Master Argument' (a phrase invented by Gallois (1974)). What is sometimes referred to as his 'semantic argument' is the only argument in his books that works, and he attaches a central importance to it. I shall outline my Berkeleyan semantic argument here, and expand it below.

All physical terms are defined analytically in terms of undefined fundamentals, whereas all mental terms are rooted in private ostensive definition. Therefore physical terms and mental terms form disjoint sets, and so physical propositions and mental propositions are disjoint, hence no mental fact can be grounded in physical facts. This delivers the first limb of mental monism. The second limb stems from the fact that we have direct experience of mental content but not of mind-independent physical systems. This enables us to establish reference to mental entities but prevents us from making reference to those physical things. Any assertion that a mind-independent world exists is therefore an incoherent attempt to refer to what is unreferenceable, or what Berkeley termed an "unknown somewhat" (1710, §80; 1713, p 482). Therefore,

the whole of physical discourse has the character of a text that refers to nothing outside itself, that is to say, it is a fiction, albeit a convenient one. I will unpack that terse summary below.

As Pearce (2014) observes, this linguistic interpretation of Berkeley differs from the standard scholastic readings—the subjunctive and ideational interpretations—both of which are trivially defeated. Pearce shows that the linguistic interpretation is both a better fit to Berkeley's writings, and philosophically more robust.

3.1 First limb: Consciousness is not Grounded in Physics

As we have noted earlier, the terms of physical discourse are defined analytically in terms of fundamental unknowns. For example, an electron is an entity with a designated mass, electric charge, spin, and so on. But the question of what mass is lies outside the discourse of physics. The equations of physics include mass as a term but the actual nature of it—what mass really is—cannot be expressed in the language of physics.

Consequently, all propositions in the discourse of physics are expressed wholly in these physical terms. From any set of propositions in that discourse, further propositions can be inferred by propositional logic and predicate logic, but all such derived propositions will also be expressed in terms from the lexicon of physics. That is, the discourse that expresses itself in physical terms is closed under logical inference.

To be sure, new terms are continually being devised within physics, but these are either defined analytically in existing terms, or denote new fundamental quantities, which are undefined except for their possession by fundamental entities and their dynamic relations to the rest of physics For example, the quark flavors of upness, downness, and strangeness (discovered in 1964), charm (1970), and topness and bottomness (1973)—these were novel physical properties that were defined in terms of behavior observed in other, already defined properties. No physicist is ever going to introduce a property such as hylasity that has no observable effect on other physical things.

In contrast, the discourse of consciousness gives meaning to its basic terms by private ostensive definition; and it builds up further terms analytically from them. That is to say, one mentally attends to some element of experience, and assigns a

designation to it. For example, 'phenomenal redness' (or, simply, 'redness') can be defined only by having an experience of that phenomenal content, and accepting that label for it. An individual with congenital color blindness who cannot see red can never know the meaning of the term 'red'. This is so, even if that individual knows about the physics of red light, and the physiology and psychology of its perception. By the way, an example of a built-up term would be 'yellow-colored, lemon-tasting circle', which is formed analytically from yellowness, lemonness, and circularity.

The discourse of the mental expresses all of its propositions wholly in these terms. From any set of propositions in that discourse, we can infer further propositions by logic, but all such derived propositions will also be articulated in expressions from the repertoire of mental terms.

Thus we find ourselves possessed of two corpora of propositions and hence of facts: the disjoint sets of physical propositions and mental propositions, and hence the disjoint sets of physical facts and mental facts. Since the corpus of physical propositions is closed under rules of inference, no mental proposition can be entailed by any assemblage of physical propositions. Ergo mental facts are additional to the set of all physical facts. To put it bluntly, consciousness is non-physical.

Certainly, one could posit a psychophysical bridging proposition. For example, that a person experiences pain if and only if a certain C-fibre in her brain fires. But any such psychophysical proposition lies outside the corpus of physical facts, and therefore cannot follow from physics. (Conversely, it is also outside the corpus of mental facts, and cannot follow from mental facts alone.) Hence it cannot defeat the conclusion that consciousness is non-physical.

Furthermore, a bridging rule is an add-on to our picture of the world, lacking a principled theoretical base from which to expand into a deeper account of relations between mind and brain.

It tells us nothing about how the psychophysical relationship changes with modification of its correlates. If the neural activity alters—if there is a change in the firing rate or the duration, or the firing of other neurons, or the existence of particular synapses—then no principles predict the effect on the mental experience. Conversely, if we consider a different mental experience, no principle predicts what the correlating neural activity would be. Furthermore, it relies on the verbal reporting of

a human, which has several challenges.

Throughout the history of science, our understanding progressed when we acquired deeper theories that provided a principled explanation of observations. A catalogue of psychophysical correlations, although useful, would fall short of a satisfying explanatory account. This warrants us in seeking an analytical theory, grounded in new principles rather than *ad hoc* correlations.

3.2 Constructs: A Prerequisite for the Second limb

In order to formulate the second limb of the argument for mental monism, we will need to use the idea of a sensory construct. To motivate this, however, I shall first touch briefly on the social construct as an illustration, as it is more familiar (e.g. Searle 1995). Consider, for example, chess. The coordinates of the board map onto the squares of, say, a flat wooden block; the pieces—the King, Queen, Knights, and so on—map onto moveable pieces of wood; the 'move' functions map onto particular movements of the pieces across the board, and predicates such as 'is a valid move' and 'is in checkmate' map onto particular movements and configurations.

Any usable construct needs both a formal system of entities and relations between them (for example, the board and pieces of chess), and a binding (for example, an agreement that this piece of wood 'is' a White Knight, and that one a Black Queen, and so on). To be sure, we may study a formal system without a binding. We might want to study chess strategies *in abstracto*. Likewise, in pure mathematics, we study axiomatic systems without semantically binding them to a model. To use a formal system as a construct, however, we have to bind its terms.

As we lack commonly agreed terminology for some of these issues, I will propose some notation. I shall define a construct as a system of predicate logic C that includes a set of constant terms T_C, functions F_C, and predicates P_C. A semantic binding is a mapping of constants T_C into a set of substrate elements P, functions F_C into relations between elements of P, and predicates P_C to facts about the substrate.

It is obvious that the entities of a construct have no independent existence: they have a purely notional existence that is exhausted by their defining propositions.

The constructs of interest to us in this paper are 'sensory constructs', and in particular 'physical constructs', rather than social ones. Sensory constructs include the world depicted in a dream, or in a virtual-reality computer simulation, or by our

ordinary waking experiences—which we conventionally refer to as the physical world. Given the discourse of physics, a 'physical construct' is defined by a subset of the constant terms in that discourse, together with functions and predicates of the discourse. Those constant terms denote the entities in the construct. Thus, a construct might correspond to a part of, or the whole of, the physical universe. A phenomenal semantic binding maps the entities of a physical construct into the mental world.

We may distinguish two kinds of phenomenal binding, which I will call 'representational' and 'embodying'. To motivate this distinction, consider two examples.

- When I look at my laptop, I have visual images to which the physical laptop maps, varying in shape as I look from different angles. This I am calling the representational binding: a set of sensory images is semantically bound to an object in the physical construct and serves to represent the object. When I dream during sleep, I behold images that do not have that binding. If I dream of my laptop, I might have the same images as when I look at the real thing, but they are not representationally bound to the physical laptop, but rather to a 'dream laptop' in the dream construct, an ephemeral and unstable simulacrum of the physical construct.
- Second, one special object has a dual relationships to the contents of my mind. Namely, my brain. If I open up my skull, and peer at the grey matter, I have a series of visual images that have a representational binding to the brain. My sensory impressions represent the brain in the same way that other impressions represent my laptop. In addition, the contents of my mind have a more intimate connection with the internal structure and electrochemical dynamics of the brain: I shall call this the 'embodying binding'. My physical body is my mind's 'avatar' in the physical construct. If I had no embodying binding, then I could have no avatar, and I would be just a disembodied observer.

Although the physical construct is virtual, we can talk as if it were real, because it is grounded in our phenomenal experience. There is a fact of the matter whether I am sitting on this physical chair, just because we construe this as expressing experiential facts. In contrast, a fiction such as Sherlock Holmes is not grounded and

it makes no sense to talk as if it were real.

Kastrup (2018), following Shani's (2015) terminology, partitions mental contents into 'revealed' experiences, which are perceptions of the outside domain, and 'concealed' experiences, which are private, interior experiences. There are, however, counter-examples to this strict disjunction: open-eye hallucinations (from psychotropic chemicals or during sleep paralysis) may integrate hallucinated objects into a real scene, or may distort a real scene in an hallucinatory manner. Another example would be a phantom limb, generated by the mind but projected into physical space. Therefore, I treat the mental contents as a whole, and allow experiences to have distinct, possibly multiple, bindings.

Formally, then: an embodying binding maps constant terms of the physical construct into the set of phenomenal contents of the mind; and maps the functions into mental operations; and the predicates into facts of those phenomenal contents. The intention is that this binding defines a correlative relationship between the mind and its avatar within the physical construct. In contrast, a representational binding maps constant terms in the physical construct into the power set of the set of phenomenal contents (that is, it maps a constant term to multiple instances of sets of phenomenal content). The intention is that these terms correspond to observable bodies in the physical domain, and each of the target sets of phenomenal content corresponds to one perception of that body.

The only difference between the physical construct and the physical universe is that the former is explicitly defined as a fiction that can function as a model of the patterns observed in the mental world, while the latter has emerged out of history tethered to the metaphysical belief that it is somehow a mind-independent reality. As far as the practice of physical science is concerned, it makes no difference whether we use the physical construct or the physical universe. The additional baggage that physicalism attaches to physics—namely the claim that physics concerns a mind-independent reality—is thus otiose: it reduces to a purely metaphysical appendix, the question of whose existence does not admit of a falsifiable answer and therefore lies outside the precinct of science.

Let us say that a construct is 'grounded' if it has a semantic binding to a substrate that is known. The physical construct is grounded because it has bindings to the mental world, which is known by direct acquaintance.

Is the converse also possible? To be sure, we could, in principle, define a 'mental construct', which would be a predicate logic whose constants represent elements of the phenomenal content of the mental world. And we could propose a binding of the mental construct in the physical world. But any such construct could not be grounded because we have no direct acquaintance with the physical world. There is thus a profound asymmetry: within the realm of the conscious mind, we can form a grounded physical construct that has all the extrinsic structure and dynamics of physics; but within the physical realm we cannot form a grounded mental construct.

Thus we are led to an internally consistent explanatory framework for the mind-body problem: (a) minds are the only reality, and (b) what we take to be a physical universe is actually a physical construct, which is an ontologically and epistemically grounded virtual world that enables us to model our experiences.

The fact that we can formulate this framework does not mean it is true. For, one could still argue that (as a metaphysical hypothesis) there might actually be a physical universe that may or may not have any effect on our phenomenal content. A common motivation for this is the observation that much of our phenomenal content is not under our personal control and appears to be governed by an unobservable mechanism. Why couldn't that unobservable mechanism just be the physical universe? This leads us to the final step of the argument for mental monism.

3.3 Second Limb: A Mind-Independent World is not Referenceable

With the concept of a 'construct', we can now articulate the second limb of Berkeley's semantic argument. Pearce (2014) presents similar reasoning, but using the less specific terminology of 'quasi-entities' for the contents of what I call the physical construct.

The argument centers on the act of referring. We can grasp this concept from straightforward cases of reference in computer science. A computer program has an address of a memory location, and uses the address as a reference by storing data into, and retrieving data from, that location. Thus, the act of reference is an act of reaching out to something by a determinate means. In like manner I can refer to parts of my sensory field. For example, I can refer to the blue strip above the text I am writing in my visual image of my computer screen. I can also refer to elements of a construct, provided that it has a grounded binding. For example, I can refer to the chess pieces in a game I am playing. Although the King is an abstraction with no independent

existence, I can refer to it, insofar as my reference is construed via the binding onto, say, the particular wooden pieces on a wooden board in front of me. If I assert that my King has taken my opponent's Queen, then my notional reference is to the abstract entities in the game construct of chess: but the substantive reference—the 'cash value' of this statement—is to be found in the physical binding: I mean that my King piece of wood has moved into the little square where my opponent's Queen piece of wood previously stood, and the latter piece has been taken off the board. The effective meaning of the statement is the state of affairs on the physical board. And the reference to the physical pieces and board in turn resolve into references to my perceptions of them.

Likewise I can refer to elements in a physical construct. If I say I am sitting on a chair, then the notional reference is to my physical body and the physical chair (in the construct), but the substantive reference is construed, via the phenomenal semantic binding, as being to my phenomenal content. It means that I can see and feel my body on the chair. An objection is that this conflates epistemology with ontology. In fact, the physical construct is "grounded ontologically and not just epistemically" in the phenomenal world (Robinson 2009, §10.2.3). This echoes Berkeley's strategy to "collapse the truth conditions for a claim into the evidence for the claim" Pearce (2014, 9.3.1).

For illustration, Chalmers (2003) and Lloyd (2003a) both argued that, in the Wachowskis' (1999) film *The Matrix* (as in traditional brain-in-a-vat thought-experiments), statements that notionally refer to an external physical world are effectively referring to a simulated world implemented in a computational rather than physical substrate. My present argument extends that line of thinking to show that notional physical references are effectively referring to the construct, and hence substantively referring to the phenomenal world that supports the construct.

It is possible to make, and understand, references to things within the physical construct, because the phenomenal semantic binding gives us a determinate means of making the references operational, of giving them meaning. In contrast, an attempt to refer to a supposed mind-independent physical universe is a dangling reference. The mind-independent physical world is, by definition, not part of our direct experience. And, by the hypothesis that it is not a construct bound to the phenomenal world, but is rather something independent of the phenomenal world, we are left with

no determinate means of operationalising the reference.

Thus we simply cannot refer to a mind-independent physical universe. The only physical domain that we can refer to is the physical construct, via its phenomenal semantic binding. In all practical matters, both in everyday life and in the physical sciences, when people refer to physical things, they are actually referring to the physical construct. The only people who pretend to be referring to a mind-independent physical universe are philosophers, who attempt—and fail—to refer to something for which there is no mechanism of reference.

This is deeply counter-intuitive. We have an almost unconquerable intuition that the physical domain to which we refer—which is actually the construct—somehow possesses a mind-independent reality. Maybe this is because we conflate the physical construct with whatever external reality is driving our experiences of the natural world. That is, however, a conflation that we must dissolve.

Pearce (2014) uses the terminology of 'quasi-referring' to physical 'quasi-entities', whereas here I prefer to say that we refer to entities in the physical construct but with the proviso that this is an indirect reference via a semantic grounding; but we agree that there can be no reference to a mind-independent domain. This terminology fits better with the linguistic practices around artificial virtual reality and enables us to import the conceptual framework of virtuality into the Berkeleyan philosophy.

Chalmers (2018, personal communication) offered the following objection. If physical terms are topic-neutral, why can't they just 'pick out' whatever is 'playing' those roles, even if they are parts of a mind-independent world? Let us see how this works in chess. As 'White Knight' is a term in the topic-neutral formal system of chess, it can pick out a piece of wood that is playing the role of the White Knight. But a piece of wood plays the role of a White Knight only by being assigned that role in a semantic binding. Being the player of a role of a topic-neutral formal system is not a natural kind. A Martian anthropologist who scoops up samples of human artefacts cannot determine that this piece of wood is a White Knight and another is the Bishop, and another is just a paperweight. The Martian would have to observe the language-game played by humans and discover that they have defined a semantic binding from the game of chess into those objects. In like manner, being the player of the role of an electron is not a natural kind. Given our experiences, with their

regularities, we have collectively devised the formal system of physics, and can bind the term denoting an electron to a particular experiential nexus. This is intelligible because we can extrapolate from our experience to the electron's. But a semantic binding, or in Chalmers' terms, a topic-neutral term's picking out a thing that plays that role, does not give us a mechanism for referencing an unknowable reality. We cannot assign a semantic binding to something that we cannot, in principle, refer to. As nobody can refer to any putative mind-independent substrate, it follows that a topic-neutral formal system cannot pick out entities in a mind-independent substrate, either.

3.4 Conclusion for Evaluation 1: Mental Monism is True

Bringing these two limbs of argument together, we find: first, the domain of consciousness is not reducible to physics; second, the physical domain to which we ordinarily refer is a construct grounded in our phenomenal content; third, any attempt to refer to a mind-independent physical world is incoherent. We must conclude that reality comprises only conscious minds, and that what we take to be the physical world is a construct within this mental world.

3.5 Corollary 1: Mental Monism Entails the Metamind

Most experiences are not under our volitional control, and it is implausible to suppose that their regularities are due to chance. What, then, could be behind it all? We can rule out the physical construct, as it is a derivative of the mental world and therefore cannot be the source of the observed natural order. Also, we can rule out the mind-independent physical world as it is an incoherent notion. So the only option left on the table is that the natural order is driven by something of the nature of a mind, which I have referred to earlier as the 'metamind'.

There are many misapprehensions about mental monism in common circulation. One is that we have two rival hypotheses for the origin of the natural order that we observe around us: either that it is caused by the physical world, or that it is caused by the metamind. And that, whilst the latter cannot be ruled out, by far the most economical and most explanatorily powerful hypothesis is that we are embedded in a physical universe, which causes the observed natural order. Not so. Mental monism is selected, not because it is simpler or because it offers greater explanatory power (as yet), but because the rival theory is philosophically incoherent. That is to say, the claim that the regularities of the manifest world are produced by a mind-independent

substrate does not make sense.

Another misapprehension is that Berkeleyan idealism has "the world as constructed from the sensory contents of individual perceivers" (Robinson 2009). In fact, Berkeley wrote that the (virtual) physical world is constructed primarily in God's mind, rendered in the sensory contents of individual perceivers, and re-constructed in the mental models of those individuals.

3.6 Corollary 2: Personal Minds are Subsets of the Mental Universe

The picture presented above is of some kind of background consciousness, the 'metamind', that runs the natural world. In common with 'cosmopsychist' theories, this poses the decombination problem: how do our private personal minds stand in relation to the metamind? Chalmers (2018) uses the term 'cognitive fragmentation' for this partitioning of the mental universe into personal minds. The question has been addressed for centuries. The Advaita Vedanta anatomizes the individual as a series of sheathes surrounding the Atman, which is non-different from the Brahman. This is picturesque but has no obvious explanatory power. Likewise is Shani's (2015) metaphor of the personal mind as "a 'vortex' surging from the oceanic background" of cosmic consciousness—a metaphor that brings a lot of baggage from fluid dynamics without offering any account of the structure and dynamics of the connection between personal consciousness and the metamind.

Kastrup (2018) draws an analogy from psychiatry: just as an individual may dissociate into multiple personalities (in Dissociated Identity Disorder, DID), so the metamind is supposed to dissociate into the personal minds of people (and, I assume, animals and disembodied beings). This is descriptively vivid but explanatorily unrewarding. Insofar as dissociation involves the partitioning of a larger system of experiences into private sub-systems, each accompanied by a subjective awareness, personal minds are indeed dissociations from the metamind. In human psychiatry, however, this phenomenon is driven by a malfunction of the mind, often produced by psychological trauma. How is a dissociation of the metamind produced and maintained? The absence of explanatory power in this analogy is evident in Kastrup & Kelly (2018), "[I]f something analogous to DID happens at a universal level, the one universal consciousness could, as a result, give rise to many alters with private inner lives like yours and ours;" and, as clinical dissociation is correlated with changes in brain function, "… We posit that this appearance is life itself:

metabolizing organisms are simply what universal-level dissociative processes look like." This says nothing over and above the bare fact that organisms have private streams of consciousness.

Instead of these metaphors, I suggest an extension of Lloyd's (1999a) reductionist picture of mental monism:

- "In set-theoretic terms, the mental universe or 'metaverse' is defined as the union of all existing minds. So, it contains the conscious minds of all human beings and all animals in the world we see around us. It will also contain any disembodied minds, if they exist. It also contains streams of mental activity that govern the complete panoply of what we think of as natural phenomena … It is a key point that the mental universe is a union of minds, and not a set of minds. For instance, if A = {a1, a2, a3, …} and B = {b1, b2, b3, …} are minds comprising experientia a1, a2, a3, … and b1, b2, b3, … then the union is U = { a1, a2, a3, …, b1, b2, b3, …} whereas the set of minds is {A, B}." (§6.4),
- "An 'ordinary mind' is a subset of the metamind, closed under ordinary operations of mental access. By 'operation' I mean any of the actions that a mind can carry out on its contents' such as perceiving, thinking, recalling memories, imagining things, focusing, paying attention, and so on. By being 'closed' I mean that any operation carried out on any of the contents yields some new contents that are still within that mind. So, whenever you recollect a memory, or focus on some bodily sensation, you remain within your own mind and do not slip into someone else's mind." (§6.5)

Within mental monism, all mechanisms must be reducible to conscious volition and qualia. There is no non-mental substrate that can serve to explain vortices or dissociations. *A fortiori*, the individuation of personal minds must be accounted for in those terms, namely volition and qualia. (By the way, I write here of individuation of minds, not of subjects. As we saw earlier, the agent of the mind—that is, the subject, S—cannot be individuated; but the contents of the mind, C, obviously are. Hence the mind as a whole, <C,S>, is individuated by virtue of the contents, C, being individuated.) Instead of metaphors from fluid dynamics and psychiatry, analogies from computer science seem more apt as it addresses and solves the same logical problems that face the metamind. In a multi-user computer, the operating system

partitions memory into private areas, and each user can access only her own section of memory. Processes in different user areas communicate only through defined channels. Operating systems often achieve this with autonomous software helpers, to which software engineers refer with the evocative name of 'daemons'.

I therefore posit a mental daemon that interfaces the personal mind and the metamind. The daemon must have access to all of the metamind (for otherwise we must posit another daemon to explain the first daemon's restricted access, leading to an infinite regress), and will deliver to the private mind only the experiences that are derivable from the sense organs of its avatar. An empirical implication of this hypothesis is that there could be failure modes of the daemon in which it could deliver experiences from parts of the metamind outside its sensory scope, that is, extra-sensory perception. I shall discuss this in a follow-up paper.

Mathews (2011) argues that individuating personal minds need not mirror the division of the world into objects, as "the boundaries between subjects are not nominal. The individuation of subjects, or centres of subjectivity, is objectively determined." Although strictly speaking the division of the physical world into objects is nominal, nevertheless to a first approximation there are distinct objects and so Mathews' argument does not entail a complete decoupling of mental individuation from object individuation. The relativistic argument below (*Evaluation 2*) shows that (a) the parts of a unified mind must be decoupled from physical objects, while (b) only whole minds can be so coupled.

3.7 Corollary 3: Physical Events Cannot Cause Mental Events

According to mental monism, the brain cannot cause anything to happen in the conscious mind, because the brain has no real existence. This is often cited as a knock-down objection, as brain events of several kinds—trauma, drugs, electrodes—appear to affect the mind. This objection, however, rests on a misunderstanding of the theory.

Mental monism has a universal set of minds $U = \{E_0, E_1, E_2, \ldots\}$ where E_0 is the metamind and $E_1, \ldots$ are whatever other minds exist, including people, animals, inanimate objects, and disembodied minds. Causation operates within and between these minds. A subset comprises minds that are rendered as observable objects, $U_{obs} \subset U$, with an object-mind E_i rendered as a physical entity $P_j = ren(E_i)$. What appears within the manifest world as causation between physical entities is actually a

rendering of a causation within the mental world. In rough terms, the scheme is as follows. Suppose E_m is my mind, ren(E_m) is my brain, E_{LSD} is the object-mind of a tab of LSD and ren(E_{LSD}) is the tab of LSD itself. Ingesting the LSD allows E_{LSD} to act causally on E_m, producing hallucinations. A third-person observer will see ren(E_{LSD}) entering ren(E_m) and consequent changes occur in ren(E_m), namely different brain activity and reports of hallucinations.

This mode of causal connection is just what we already use in virtual-reality systems, such as video games: we use a simulated gun to fire a simulated bullet at the avatar of a simulated enemy, and he falls to the simulated ground. Contrary to what we may feel in the moment, we know that the causation does not operate between the rendered images on the computer screen. We know that the causation operates between the object-modules of the VR software.

3.8 Corollary 4: The Brain is a Virtual Interface

So, what is the brain for? The brain is the mind's avatar in the physical construct, and it needs two interfaces. On the one hand, it has the sensory and motor organs plus the brain tissue that carries out pre-conscious input processing and post-conscious output processing. On the other hand, it has the physical correlate of consciousness, which acts like a portal to the actual conscious mind. That is what the brain is for: an interface mechanism between virtual transducers and actuators on the one hand and the correlate of consciousness on the other; and this interface must operate within the laws that rule the physical construct, hence it has to be a physical object with specific characteristics, within the construct.

I mentioned in the previous subsection the mental daemons that restrict a personal mind to the scope of an avatar: those daemons must exist on the mind-side only, and provide the link between the personal mind and the physical correlate of the mind. For, otherwise, if the daemon were on the brain side then the personal mind would have the freedom to choose whose brain to connect to, and there would not be mental privacy. There is therefore no corpus of brain tissue that implements the mental daemon.

4. Evaluation 2: Panpsychism is False

I argued above (*Comparison*) that panpsychism is a special case of mental monism, and that (*Evaluation 1*) mental monism is true, but in this section (*Evaluation 2*), I

argue that panpsychism is false, and that reality must therefore be a non-panpsychist form of mental monism.

I also argued above that mental activity cannot literally be in space, but that still leaves open the possibility that states of elemental constituents of personal minds might be necessarily correlated with states of physical entities. Here, however, I shall extend the argument and show that even that cannot be so.

In the (slightly) bigger picture, even if panpsychism were true, it would not 'explain' the existence of personal minds, as bare panpsychism does not entail the combination of micro-minds into macro-minds (Goff 2009). The ontogenesis of personal minds requires some additional theory. I don't see anything philosophically problematic in the merging of minds to form larger minds per se, since it could be achieved simply by taking the set-theoretic union of the contents of the merging minds. My concern here is not with combination as such, but with the combination of spatially separable minds into a unified mind.

I will examine three arguments concerning the question of whether mental states can be associated with the states of separable physical entities. First (*Argument from Relativity I*) is a standard argument in favor of localization, which I shall reject. Second (*Argument from Relativity II*) shows that in fact relativity prohibits mental localization. Third (*Argument from Excision*) is a thought-experiment that illustrates a paradox arising from mental localization.

The starting premise of both of the two latter arguments (against localization) is that we cannot mistake the large-scale sequence of perceptions. While two events happening within, say, half a second might confuse us, there is some temporal distance beyond which the sequence is unimpeachable. That interval might be an hour, or a week, or a decade, but it is self-evident that sufficiently separated mental events have a knowable order.

4.1 Argument from Relativity I: Consciousness is Not Spatiotemporal

Lockwood (1992) used Einstein's (1905) special theory of relativity to argue for localization of consciousness. Here I reject that claim, and in the next subsection make a counter-claim, that relativity makes it impossible for consciousness to be separable in space, and hence that the physical correlate of a mind must be a physical simple.

Russell (1927, p 384), Weingard (1977), and Lockwood (1984a,b; 1985; 1989, pp 71-78) used relativity to argue that, as conscious events are in time (proposition R_T), so they must be in space (R_S) since relativity melds time and space inextricably in spacetime. The reasoning from R_T to Rs is sound, but premise R_T is not. Lockwood's defence of R_T is as follows. We continually witness that the mind perceives and causes physical events, but physical events are in physical time, therefore mental events must be in physical time, at least to the granularity of the physical stimulus and response. For example, if I hear a buzzer at t_0 and press a button at t_1, then my perception and volition must have occurred between t_0 and t_1. Now, for any brain event B that correlates with a mental event, Lockwood (1989, p 75) wrote: "If it were now possible to find pairs of physical events standing to B as cause and effect that were separated by ever smaller temporal intervals then it would be possible to define the spatial location of B with any required degree of precision."

For sure, the neural correlates of the acts of perception and volition must be localized in spacetime. But Lockwood's programme of narrowing down arbitrarily the volume of spacetime within which the correlate of any mental event occurs is unfeasible because the speed of light is so high that to localize B even within the 140 mm width of a human brain, the interval between t_0 and t_1 would have to be less than 10^{-10} s but no discernible mental events occur at that scale (even if we allow Pelczar's (2017) claim that mental events can be as brief as 10^{-4} s). Hence Lockwood's program cannot localize the correlates better than the crude observation that they are somewhere in the head.

Nevertheless, we must address Lockwood's core assertion that since mental events can be circumscribed in an interval of time (even if no finer than half a second), they must be circumscribed in some volume of space (albeit a large one)? Lockwood's argument rests on the premise that a physical event can cause a mental one. Any intelligible notion of causation, however, entails temporality, hence assuming that a physical event causes a mental one already presumes that consciousness sits in physical time, which is what the argument is supposed to prove. Hence it is circular.

As we saw in Corollary 3 above, physical events cannot cause mental ones under the theory of mental monism. So, what is happening in the simple experiment where I press a button whenever I hear a buzzer? According to mental monism there is a

mental event M_{Buzz} that is rendered in the physical world as ren(M_{Buzz}), the physical sound waves. But it is M_{Buzz} that acts on my mind, not ren(M_{Buzz}). M_{Buzz} yields the mental sensation of the buzzing noise, but its doing so is a cause-and-effect in the mental domain and it has no coordinate in physical time. My volition to press the button is again a mental event M_{Press}, which is rendered in the physical world as ren(M_{Press}). Thus, in mental monist, there are two decoupled temporal sequences: on the one hand, the mental event M_{Buzz} causes me to hear the sound, and my volition then performs M_{Press}; on the other hand, within the virtual physical world, ren(M_{Buzz}) occurs, and then ren(M_{Press}), which occur at specific physical times. Relativity theory pertains to the rendered events ren(M_{Buzz}) and ren(M_{Press}), which must indeed occur in spacetime; but the mental process from M_{Buzz} to M_{Press} subvenes on that physical world. As ren(M_{Buzz}) cannot cause M_{Buzz}, the causal link that Lockwood wants to use to tie consciousness into spacetime does not exist, and his argument fails.

The theory of relativity explicitly concerns the relationship of physical *observations* in spacetime. If mental experiences were observable events in spacetime then Lockwood would be right, but they are not third-person observable, so there can be no operational meaning in ascribing physical time or space coordinates. Relativity says nothing about physically unobservable things, and certainly does not require them to be embedded in spacetime. Gibbins (1985) tried to soften the blow by writing, "The temporal results of relativity will apply to mental events only indirectly via the space-time events with which they are associated," but being 'associated' with an event in spacetime does not constitute actually being in spacetime. In fact, mental events are not in physical time any more than they are in physical space.

4.2 Argument from Relativity II: Physical v Mental Simultaneity and Sequence

I argued above that mental events cannot be in spacetime under the theory of mental monism. In this subsection I present an argument for this conclusion that does not rely on mental monism and so defeats panpsychism more directly. In fact, relativity offers us two anti-physicalist arguments, as follows.

(a) A mind can have distinct, absolutely simultaneous experiences but the brain cannot, because simultaneity is relative to the observer's frame of reference. Russell (1926, p 130) briefly mentions this as a puzzle but draws no conclusion from it; de Silva (1995,1996) likewise; Lee (2007) concludes that mental simultaneity is an illusion; Pelczar (2017) concludes that consciousness cannot supervene on events in

spacetime. These authors disregard the problem that a human mind of normal acuity cannot tell whether two mental events are precisely simultaneous, hence there can never be any empirical evidence for strict mental simultaneity, and the premise of the argument fails.

(b) A mind has experiences in an absolute sequence, but the brain cannot, because the sequence of spacelike-separated physical events is relative to the observer's frame of reference. Lee (2007) also mentions this. Arguing from sequence, rather than simultaneity, is stronger as the sequence of sufficiently separated mental events is discernible and unambiguous.

This argument, however, is flawed because events in a brain of normal size are not spacelike separated. For example, the separation of two brain events at $(x_0,0,0,t_0)$ and $(x_1,0,0,t_1)$ is spacelike if $(x_1-x_0)^2 > c^2(t_1-t_0)^2$ where c is the speed of light, about $3x10^8$ ms^{-1}. For a brain diameter of x = 0.14 m, t cannot exceed about 10^{-10} s, but this is way below the ability of any human mind to discern. Lee (2007) did not recognize this as a problem, and Phillips (2009, p 222) rightly rejected Lee's relativistic argument because of it; but Phillips wrongly supposed that there could be no such argument from the relativity of sequence. In fact, we can consider a thought-experiment in which the brain is expanded to make its parts spacelike separated.

Panpsychism does not prescribe any maximal volume of space that must contain the physical components whose elemental minds combine to form a personal mind. So let us consider a human brain expanded to the size of the Earth's orbit around the Sun, with the same architecture as a regular brain but the axons stretched over millions of miles. A neural pulse travelling at a hundred meters a second will take ninety-five years to traverse the brain. Local transmissions, between, say the visual and auditory cortex would take a few decades. Obviously this is an outlandish thought-experiment but the logic of panpsychism and relativity must still apply.

In a brain this big, neural events are spacelike separated: it would take light sixteen minutes to cross the diameter of the brain.

Considering the mind that is associated with this brain in accordance with panpsychism, suppose that it experiences a bright flash and, several years later, a loud bang. We cannot doubt the mental sequence. In relativistic physics, however, there is no such thing as absolute sequence of spacelike-separated events.

Given two spacelike-separated brain events B_1 and B_2, in some frames of reference, B_1 will precede B_2; in some B_1 and B_2 will be simultaneous; and in some B_2 will precede B_1, depending on the velocity of the observer. Yet, if the phenomenal contents M_1 and M_2 that are associated with events B_1 and B_2 are combined into a unified mind, then M_1 and M_2 will have a determinate temporal sequence. The fact that the observers would have to be travelling at very high velocities to change the brain events' relative sequence is not to the point. What matters is the brute physical fact that sequence is not absolute for spacelike-separated events, and therefore the combination of spatially distributed elemental minds into a unified mind is untenable.

4.3 Corollary 1: Non-Simultaneity v Non-Unified Metamind

Let us step back briefly from this argument to consider the bigger picture. We have seen that relativity theory's prohibition of simultaneity excludes the personal mind from being mapped to the tissues of the brain. Does it also exclude the possibility of the universal metamind's being mapped to physical objects (that is, having an object-oriented architecture, as suggested earlier)? The answer is no, because the metamind's constituent object-minds (which map to individual physical objects) are not collectively unified in a single stream of consciousness. Relativistic non-simultaneity is problematic only when one tries to map a unified stream of consciousness to separable objects. The personal mind is a unified stream of consciousness, whereas the metamind is not; hence the personal mind cannot be mapped onto separable objects whereas the metamind can. Nevertheless, as we shall see below, further considerations arise if 'telepathy' is permitted between remote minds.

This does not deny the core intuition—shared by panpsychists and object-oriented mental monists alike—that we should expect the structure of the mental world to mirror the physical world, but panpsychists and mental monists require it to be at the micro and macro levels respectively.

4.4 Corollary 2: Hypothesis of Universal Mental Time

Relativity prohibits faster-than-light communication (Tolman 1917), but the decoupling of mental time and physical time seems to open up the possibility of instant communication. We need to check this potential conflict.

(a) *Physically instant communication is impossible.* Relativity states that if two observers, A and B, travel in uniform motion relative to each other, then in relation

to A's frame of reference, time will slow down in B, and vice versa. This is well established experimentally. Consequently, superluminal communication would lead to contradiction. Suppose at t_3, A sends an instantaneous signal to B. Because of time dilation, it arrives at B at $t_2 < t_3$. Upon receiving this signal, B sends an instant message to A, saying "Don't send the signal," which arrives at t_1. Because of time dilation, $t_1 < t_2 < t_3$. On receiving this message, A decides not to send the first signal at t_3, and we have a contradiction. (b) *Mentally instant communication is permitted by mental monism.* Since minds are not in space, communication between minds is not required to travel through any intervening space. Therefore mental monism, in principle, permits communication between two minds in an instant of mental time. Furthermore, mental monism, in principle, also permits disembodied minds. Therefore even if we try to block instant communication between two embodied minds by saying it cannot 'travel' faster than light between the respective avatars, this could be circumvented by using a disembodied mind as an intermediary. (c) *Therefore, if mentally instant communication entails physically instant communication, then mental monism cannot be true.*

What exactly do we mean by 'instant communication' from A to B? In a physical system, it means that P_A originates a signal at time t and it arrives at P_B at the time t' that, by the Lorentz time dilation equation, exists in P_B when P_A is at t. In a mental system, it is not so well defined, as we have no units or instruments for measuring mental time. The most plausible model is one that has been in currency from James (1890) to Hameroff and Penrose (2014), namely that consciousness comprises a series of discrete moments of awareness. Whether we accept this discrete model, or one in which consciousness flows continuously in time, the 'clock' that drives consciousness is the succession of experiences. For an embodied mind that is interacting with its environment, that 'clock' will in turn be correlated with the succession of neural impulses entering the sensory cortex, or arising from internal brain activity. Therefore, the physical correlate of elapsed mental time will be the elapsed physical time within the local frame of reference of the avatar (be it a brain or an artificial device). If, at time t, mind M_A sends a mentally instantaneous signal to M_B, then it will arrive at its destination at the same local time t (not the Lorentz dilated time $t' < t$). Whereas physically instantaneous communication goes backward in time, mentally instantaneous communication does not. Hence mentally

instantaneous communication, as permitted by mental monism, does not entail a relativistic paradox.

To illustrate this, we can consider Einstein's thought-experiment of two lamps at opposite ends of a high-velocity train. One observer, Passenger, sits in the middle of the train; and another, Stationmaster, in a station. As the train passes the station, each lamp makes one flash. If Passenger receives both light signals at the same time, she concludes that—in her reference frame—the lamps flashed simultaneously. Meanwhile Stationmaster sees the rear lamp flash before the front lamp, but both signals reach Passenger together. (Photons from the rear lamp must travel farther. In the time it takes the photons to travel from the rear lamp to Passenger, she will have moved forward slightly with the train.) If a third observer, Aeronaut, is travelling faster in the same direction (in an aeroplane, say) alongside the train, then she will observe the opposite sequence: in her reference frame, the front lamp signals first.

So far, this is standard relativity theory in the physical domain. Now let us consider the mental domain, where mental events are not in physical spacetime, and therefore not embedded in frames of reference. According to object-oriented mental monism, a lamp's flashing is a physical rendering of a state change in the corresponding object-mind. Given that the sequence of physical flashes depends on each observer's frame of reference, we naturally want to ask: In what sequence do the object-minds actuate these flashes? As the object-mind's state change is not in physical time, there is no fact of the matter as to which mental state changed physically first; but we can still ask which occurred first in mental time. Imagine (for the sake of argument) that observers can 'telepathically' communicate with object-minds. Now the observers can interrogate the object-minds of the lamps and discover which object-mind triggered a signal first. Suppose, for example, that the rear lamp's object-mind triggered its signal first. Now, the Aeronaut will be surprised to discover that, although front lamp flashes first in physical time (in her reference frame), the rear lamp's object-mind triggered a signal first. Thus she has foreknowledge of a future physical event, which is spooky but not paradoxical. Next, imagine (again, for the sake of argument) that an observer can 'telekinetically' control the object-mind of a lamp. Let us vary the experimental protocol: as soon as Stationmaster discovers that the rear lamp's object-mind has actuated its flash, he decides whether or not to veto (telekinetically) the front lamp's flash. This potentially creates a problem for the

other two observers. For, Passenger should see both flashes at the same time, and Aeronaut should already have seen the front flash first, which Stationmaster has just cancelled telekinetically.

This paradox arises only if telepathy is physically instantaneous; but as we have seen, it is instantaneous only mentally, not physically. To be sure, in Aeronaut's reference frame, the time at which the front lamp is scheduled to flash can precede the time when Stationmaster decides whether to cancel the front flash. Aeronaut can telepathically ask Stationmaster if he has done so, and even ask the front lamp if it has flashed, and therefore can know whether the front lamp will flash, at a time physically before the lamp would, but the light signal from the front lamp (travelling, obviously, at the speed of light) will not reach Aeronaut until after that time. So if the front lamp is vetoed, Aeronaut will not see the light, and there is no paradox, and no violation of causality. Admittedly, in Aeroaut's frame of reference, it seems that Stationmaster's mind retrocausally went back in time to cancel the front lamp's flash, but in Stationmaster's own frame of reference, there is no retrocausation.

In an earlier subsection (*Argument from relativity II*), we saw that mental events that are inside a unified stream of consciousness must be in absolute order (not relative to an observer), and therefore cannot be mapped onto separable physical events. We now see that the existence of instant communication ('telepathy') between remote minds is enough to impose an absolute ordering on all mental events in the universe. This has the weird (but not paradoxical) result that events in two remote physical avatars could occur in any order (depending on observer) but mental events in their corresponding minds must occur in a single, definite order. This may seem unpalatable as it hints at a privileged frame of reference, which would go against the spirit of relativity. Nevertheless, we should remember that the notion of no privileged frame of reference is not an axiom, but a consequence of verificationism (since there is no way to tell whether you are at rest or in uniform motion), and nothing proposed here violates verificationism.

4.5 Argument from Excision

We now resume the main thread of this paper and examine the third argument about mental location.

We may say that 'the quickness of the neuron deceives the mind.' Leibniz (1714, §17) was perhaps the first to remedy this by imagining a vast thinking engine whose

tangible components take away that deception. Davis (1974) likewise imagined a brain-like organization in which each neuron was simulated by an office and axons by telephone lines, while Block (1978) proposed a brain simulated by the population of China. Lloyd (1999a, §2.4) argued that a thought-experiment that introduced macroscopic delays in the brain led to a contradiction between the unity of consciousness and an identity theory of mind and brain. Here, I will reframe the argument for panpsychism.

Returning to a brain of normal size, consider two pieces of brain tissue B_1 and B_2 whose quiddity (according to panpsychism) is constituted by two mental experiences M_1 and M_2, and suppose that M_1 and M_2 are co-conscious at the same time. Let's say M_1 is seeing a flash and M_2 is hearing a bang. Let us surgically pull B_2 out of the cranium, while retaining its full connectivity. Assume that all nerve fibers to B_2 are extended by, say, a metre of microscopically thin optic fiber. Keep this fragment of brain, B_2, alive and working as normal in a petri dish. Part of the person's mind is now in the petri dish, and the rest of it is in the person's head. This is undoubtedly weird, but it is not paradoxical: both brain function and subjective experience are unchanged.

According to panpsychism, B_1 and B_2 by themselves constitute the bare existence of M_1 and M_2 but do not constitute the co-consciousness of M_1 and M_2. Panpsychists have not explained how distinct experiences such as M_1 and M_2 combine into a unified mind, but as panpsychism rests on the notion of consciousness as the interiority of the physical, the combination mechanism must be a physical process that also has a mental consequence. So, logically, there must be a third physical event, B_0, that combines M_1 and M_2 by making them co-conscious. Panpsychists seem to have no account of what B_0 is, but to be plausible it must involve some causal connection between B_1 and B_2. For the sake of concreteness we may imagine it involves the transmission of some neural signals between B_1 and B_2. That is, M_1 and M_2 are co-conscious if, and only if, B_0 occurs.

Let us delay the onset of co-consciousness by inserting microdevices into the fibres connecting B_2 to B_0 in the brain. These microdevices can delay signals by anything from zero seconds to twenty-four hours. We perform a series of tests in which we give the person the flash and bang stimuli, and on each trial we increase the delay for B_0. At a delay of zero, the situation is as normal; at a delay of, say, five

minutes, it is paradoxical. Say the stimuli are applied about half a second before noon, so that at 12:00 precisely, B_1 and B_2 occur and hence (by panpsychism) M_1 and M_2 occur. But B_0 occurs five minutes later, at 12:05. Hence M_1 and M_2 become retrospectively co-conscious five minutes after they happened. This contradicts the principle that the macro-ordering of mental events is indubitable.

It is impossible to imagine what this person experiences. At 12:00 she sees the flash, but hears no bang. Meanwhile, the isolated mind of B_2, sitting in the petri dish, hears the bang. At 12:05, it retrospectively becomes the case that the bang M_2 was co-conscious with the flash M_1. But the subject does not hear the bang at 12:05, for the bang already happened at 12:00, in its own mini-mind. In fact, the subject will not notice anything at all. Yet, according to panpsychism, M_1 and M_2 have become co-conscious.

Either mental events are co-conscious when they occur, or they are not co-conscious. Since panpsychism entails retroactive co-consciousness, it entails a contradiction, so we must reject it.

4.6 Spatially Separable Panpsychism is Untenable

I have given two thought-experiments showing that the panpsychist premise of associating mental states with the states of separable physical entities leads to contradictions. Therefore, a mind's correlate within the physical construct must be a spatially non-separable 'simple', a physical entity of some sort that, if it has parts, then those parts cannot be separated out. The nature of that simple is beyond this paper, but quantum-mechanical correlates of consciousness, such as the state of a microtubule, might work (e.g. Hameroff and Penrose 2014).

To avoid this problem with panpsychism, some writers have advocated 'cosmopsychism', associating a universal mind with the state of the whole physical universe. For example, Mathews (2011), Jaskolla and Buck (2012), Shani (2015), Nagasawa and Wager (2016). At first, this seems like an extrapolation of panpsychism, but jettisoning any structural mirroring between the physical construct and the mental world loses the core intuition of panpsychism. Cosmopsychism is another special case of mental monism, for the same reason as panpsychism is, namely because the mind cannot really be in space, which is part of a topic neutral system. Cosmopsychism is not, however, equivalent to the most general form of mental monism as the latter allows for disembodied mental entities.

Berkeleyan idealism seems more plausible, and offers more explanatory power, than cosmopsychism, but it remains to be seen which theory is right.

As we saw earlier, the metamental structures responsible for macroscopic bodies are object-minds. By analogy with human minds, we might consider the plausible supposition that the object-mind also has a unified stream of consciousness, although we cannot tell this for sure. If so, then, despite the protestations of panpsychists that tables and chairs do not have recognizable minds, they would in fact do so. Obviously there is no cognitive circuitry inside items of furniture, so the table cannot think, nor the chair speak, but it seems the table may have a rudimentary but unified conscious mind, as does its companion chair.

5. General Conclusion

I have argued that panpsychism is a particular case of mental monism, that mental monism in general is true, but the particular case of panpsychism is not; that reality consists of minds, and what we take to be the physical world is a construct grounded in the mental world; that there is at most a single common subject for all minds; that each mind correlates with a spatially non-separable simple in the physical world; and that the natural order is driven by a metamind, which resolves into encapsulated object-minds that are responsible for macroscopic bodies.

6. Postrequisite: Hierarchical Language-Games

The earlier sections (*Comparison* to *Evaluation 2*) lead to the *General Conclusion*, and one can accept or reject this reasoning as it stands. Even if one accepts this reasoning as valid, however, the conclusion may seem unbelievable. Faith in the real existence of the physical world is rooted so deeply that many people find it hard to let go, even in the face of logic. In this final section, I will address what appears to be a central obstacle to the credibility of mental monism. This does not add anything to the arguments *per se* that lead to the general conclusion, but is a 'postrequisite' in the sense that is required after the conclusion to make it credible.

Given the conclusion that only conscious minds exist, how can we speak correctly about the everyday world? According to mental monism, the table at which I sit does not exist. How then, can I meaningfully say I am sitting at this table, writing this paper? Generations of critics have protested that mental monism, whilst it may be logically irrefutable, is pragmatically untenable. This objection was anticipated by

Berkeley (1710, §51), who wrote that, in the 'vulgar acceptation', it is correct to assert the existence of the table, whilst in the 'learned acceptation' it is correct to deny its existence. Following Lloyd (2006, §1.4), I shall recast Berkeley's differentiation of 'acceptations' into the modern terminology of Wittgenstein's language-games. Flew (1974) has also recognized in Berkeley (1732, §7) an adumbration of Wittgenstein.

Wittgenstein (1953) coined the term 'language-game' to denote a rule-bound pattern of social activity where utterances are integrated into a wider apparatus of action. His motive was to jettison the naïve Lockean (1690) notion of words as labels and the early Wittgensteinian (1922) notion of propositions as pictures, and to forge a new theory of integrated language and behavior. In this later and more nuanced account, Wittgenstein wrote (1953, §43), "For a large class of cases … in which we employ the word 'meaning' it can be defined thus: the meaning of a word is its use in the language." Thus a word is no longer considered as having a single fixed meaning but as having different roles in different language-games. Wittgenstein compared words to the levers of a steam engine, which look and feel the same, as they are made to be handled, but perform functions of different kinds. Likewise, words might share the same look-and-feel but play roles of different kinds.

Here, I will use language-games to differentiate the roles played by terms of two particular classes: *physical* and *mental*. For instance, the preposition 'in' has the same look-and-feel in sentences such as these: "I have some food in my stomach" and "I have a pain in my stomach." The meaning of 'in' is, however, fundamentally different: in the first case, it denotes the spatial containment of an object; in the other, it denotes a psychophysical correlation. A shared linguistic appearance hides that difference, inducing the naïve belief that, in this example, the pain is actually inside the stomach.

This differentiation of meaning is illustrated in the Wachowskis' *The Matrix* (1999), whose protagonists move between an immersive virtual-reality and a physical world (Lloyd 2003a). By Tarski's (1933) definition of truth, a statement such as "Neo is sitting in a noodle bar" is true if he is sitting in a noodle bar. Now, consider (a) sitting in a noodle bar in the physical world and (b) sitting in a simulated noodle bar in the computer-generated world. In both Neo can truthfully say, "I am sitting in a noodle bar," but what does the sentence actually *mean* in situations (a) and (b)? Consideration of the divergent 'truth-conditions' and 'truth-tests' of that sentence

will give us two different ideas of the meaning. On the one hand, the truth-*conditions* of that same sentence in the situations (a) and (b) are mutually exclusive: if Neo sits in the physical noodle bar, then he is not wired into the virtual-reality and he cannot be sitting in the simulated noodle bar; and vice versa. On the other hand, the truth-tests of the sentence in (a) and (b) are identical: Neo looks around and sees the tables and bowl of noodles, and smells and tastes the food; if he telephones Trinity and invites her to join him, she can have sensory impressions consistent with his. Therefore, when Trinity asks "Are you in the noodle bar?" and he says "Yes," he conveys information about his sensory impressions, not the unobserved putative substrate. So, empirically, the meaning intended by the utterance is not the truth-conditions but the truth-tests. The common academic doctrine that truth-conditions, rather than truth-tests, constitute meaning misses this point.

In the physical language-game, statements refer to truth-tests within the physical world; in the other, they refer to truth-tests within the simulated world. A proposition may belong to one or the other language-game according to whether it is intended to be about the physical or simulated world. In both cases, the truth-tests may be the same, but the tests are carried out in different situations.

These language-games, however, are not equal. For, Trinity and Neo are acquainted with both worlds and can move between them, so when Neo invites her to join him, she may ask, "Which noodle bar? The real or the simulated?" In Tarski's terms, the physical language-game is a metalanguage to the simulational language-game. This metalinguality consists in the fact that the results of truth-tests of the simulational language-game are logically supported (in Foster's term) by facts within the domain of the physical language-game; and not vice versa. The contents of the computer database (which are facts within the physical world) define the truth or falsity of statements (hence facts) within the simulational language-game. Thus Neo's being in the simulated noodle bar is a fact in the simulated world, but it rests on data that have been populated in a physical computer; and those data are physical facts. Consequently, from the point of view of the metalingual language-game (in this case, the physical one), the world denoted by the subordinate language-game (in this case, the simulational one) is a construct.

This extends Chalmers' (2003) analysis of *The Matrix*. For individuals who are locked into the Matrix, Chalmers wrote that the claim that their world is a computer

simulation is not a *sceptical* hypothesis but a *metaphysical* one. "The Metaphysical Hypothesis here tells us about the processes underlying our ordinary reality, but it does not entail that this reality does not exist. … Both the Physical Hypothesis and the Metaphysical Hypothesis tell us about the processes underlying chairs. They do not entail that there are no chairs. Rather, they tell us what chairs are really like." For individuals, such as Neo and Trinity, who understand what the Matrix is and can move between the simulated and natural worlds, that analysis is not sufficient. For them, the claim that the world inside the Matrix is a computer simulation is no longer just a metaphysical hypothesis. They can leave the Matrix and visit Machine City where they can see the machines that run the simulation. It is to accommodate this broader perspective that I propose that we should view the discourse of Neo and Trinity as a hierarchy of language-games, for the simulated and physical worlds. In fact, Lloyd (2003b) argued that what seemed to be the real world in *The Matrix* is revealed in the sequel, *The Matrix Reloaded* (Wachowski & Wachowski 2003) to be another computer simulation inside another world, which implies a hierarchy of three language-games.

Let us pause to recapitulate. First, a given sentence may entail fixed truth-tests but also entail multiple different truth-conditions that depend on which language-game it is uttered in. Second, if one language-game is metalingual to another (in the above sense) then the truth-conditions of statements in the subordinate language-game describe a construct that is logically supported by facts in the metalingual language-game.

This conclusion is pivotal to removing the common misconception that mental monism is pragmatically refutable since it requires disbelief in the necessities of everyday life. How, for example, can I write this essay on mental monism if I believe that this laptop is unreal? The answer is that (a) in the physical language-game, the laptop can be said to exist; but (b) in the mental language-game, the laptop can be said to be a fiction, and (c) the mental language-game is metalingual to the physical one. The physical language-game is subordinate insofar as the results of truth-tests of physical statements are supported by facts in the mental world. For example, a truth-test for the (physical) statement that there is a potential difference of 1.5 volts between the terminals of the battery in this laptop would involve reading a voltmeter; but my consciously seeing a reading of 1.5 volts is a fact in the mental world, which is

articulated in the mental language-game. It is in precisely this sense that all physical truth-tests are supported by mental facts; and that the physical language-game is therefore subordinate to the mental language-game.

With that distinction of language-games in mind, we can now dispel a common objection to Berkeleyanism. According to mental monism, it is true that (p) the laptop on which I am writing is real, but on the other hand, it is also true that (q) the laptop on which I am writing is a fiction. The proviso is that the proposition (p) is taken in the physical language-game, and the proposition (q) is taken in the mental language-game. This is a distinction of precisely the same kind as the one we make when dealing with virtual-realities such as the Matrix, which allows us to make the true statement (r) I am in the noodle bar (in the simulational language-game), and (s) I am not in the noodle bar but wired into a virtual-reality (in the physical language-game). Differentiating the language-games obviates the self-contradiction that many impute to Berkeley's philosophy, and we see that mental monism is pragmatically serviceable as well as logically sound.

7. Discussion: Potential Applications

The motivation for looking at mental monism, and the arguments defending it, are philosophical, but that does not mean the theory has no value outside philosophy. Although mental monism changes nothing in physics—the laws governing the physical construct are those that have been uncovered, and will continue to be uncovered, by the disciplines of the physical sciences—nevertheless there is considerable evidence for anomalous phenomena (e.g. Utts 1996, Hyman 1996, Radin 1997). Psi phenomena—principally telepathy, telekinesis, and telecognition—resist explanation in physical terms at a deep level. If mental monism provides a way to develop naturalistic explanatory models of psi that do not violate physics, then that is worth following through. On a different front, the problem of consciousness engineering—of building artificial conscious minds—is one that requires a fundamental framework in which consciousness is understood as a real component of reality rather than eliminated. Both of these potential applications are examined in follow-on papers in preparation.

Acknowledgement: This paper was improved by critical comments on an early

draft, from Bill Adams, David Chalmers, and John Gregg.

In memoriam: Michael J. Lockwood (1944-2018) was my tutor at Oxford, and my intellectual debt to him is great. His death is a great loss to the study of consciousness.

8. References Cited

Ayer, Alfred (1963). The concept of a person, in Ayer, A. (ed.), *The Concept of a Person and Other Essays*, London: Macmillan.

Berkeley, George (1710). *The Principles of Human Knowledge*, Volume I, Dublin: James Pepyat.

— (1713). *Three Dialogues between Hylas and Philonous*, London: Henry Clements.

— (1732). *Alciphron, or the Minute Philosopher*, London: T. Jonson.

Blamauer, Michael (2011) (ed.). *The Mental as Fundamental: New Perspectives on Panpsychism*, Frankfurt: Ontos Verlag.

Block, Ned (1978). Troubles with Functionalism, in Savage, C.W. (ed.), *Perception and Cognition: Issues in the Foundations of Psychology*, Minneapolis: University of Minnesota Press.

Bruntrup, Godehard & Ludwig Jaskolla (eds.) (2017) *Panpsychism*, Oxford: OUP

Chalmers, David (1996). *The Conscious Mind*, Oxford: OUP.

—(2003). The Matrix as Metaphysics [originally in philosophy section of official website of The Matrix], in Grau, C. (ed.) (2005), *Philosophers Explore the Matrix*, Oxford: OUP. http://consc.net/papers/matrix.html

— (2013). Panpsychism and Panprotopsychism, Amherst Lecture in Philosophy, in Torin Alter & Yujin Nagasawa (eds.) (2015) *Consciousness in the Physical World: Perspectives on Russellian Monism*, Oxford: OUP. Reprinted in Godehard Bruntrup & Ludwig Jaskolla (eds.) (2017) *Panpsychism*, Oxford: OUP. http://consc.net/papers/panpsychism.pdf

— (2017a). The combination Problem in Panpsychism, in Godehard Bruntrup & Ludwig Jaskolla (eds.) (2017) *Panpsychism*, Oxford: OUP. [Chalmers: "written on the same weekend as the other paper on panpsychism, but … published four years later due to the vagaries of publishing."] http://consc.net/papers/¬combination.pdf

— (organizer) (2017b) *Workshop on Idealism and the Mind-Body Problem*, NYU-Shanghai, June 2017, Shanghai, China. (No proceedings published, but individual papers can be found online. The workshop replaced Plenary Session 7, *Idealism and Panpsychism*, June 8, 2017, of the cancelled Shanghai conference *The Science of Consciousness* , which was not reinstated in the San Diego replacement conference, 5-10 June 2017.)

— (2018). Idealism and the Mind-Body Problem, in Seager, W. (ed.) *The Routledge Handbook to Panpsychism*, London: Routledge. http://consc.net/papers/idealism.pdf

de Silva, Frank (1995), Consciousness and Special Relativity?, letter, IEEE Engineering in Medicine and Biology, July/August 1995, p 358.

— (1996), Consciousness and Special Relativity?, IEEE Engineering in Medicine and Biology, May/June 1996, pp 21-26.

Davis, L. (1974). unpublished colloquium at MIT cited by Dennett, D., Toward a Cognitive Theory of Consciousness, in Dennett, D. (1978) *Brainstorms: Philosophical Essays on Mind and Psychology*, Cambridge, MA: MIT Press.

Einstein, Albert (1905). Zur Elektrodynamik bewegter Körper [On the Electrodynamics of Moving Bodies], Annalen der Physik. 17 (10), pp. 891–921. Translated by Jeffery, G.B. & Perrett, W. (1923) *The Principle of Relativity*, London: Methuen and Company.

Flew, Anthony (1974). Was Berkeley a precursor of Wittgenstein?, pp 153-163, in: William B. Todd (ed.), Hume and the Enlightenment: Essays presented to Ernest Campbell Mossner, Edinburgh: Edinburgh University Press.

Foster, John (1982). *The Case for Idealism*, London: Routledge & Kegan Paul.

Gallois, Andre (1974). Berkeley's Master Argument, *Philosophical Review* 83, pp. 55-69.

Gibbins, P.F. (1985). Are mental events in spacetime?, *Analysis* 45.3, pp. 145-147.

Goff, Philip (2009). Why Panpsychism Doesn't Help Us Explain Consciousness, *Dialectica* 63, pp. 289–311.

Goldschmidt, Tyron, and Kenneth L. Pearce (eds.) (2017). *Idealism: New Essays in Metaphysics*. Oxford: Oxford University Press.

Hameroff, Stuart & Roger Penrose (2014). Consciousness in the universe: A review of the 'Orch OR' theory, *Physics of Life Reviews* 11, pp. 39–78. https://doi.org/10.1016/j.plrev.2013.08.002

Hyman, Ray (1996). Evaluation of Program on Anomalous Mental Phenomena, *Journal of Scientific Exploration* 10(1), 31-58.

James, William (1890), *Principles of Psychology*, New York: Henry Holt & Co.

Jaskolla, Ludwig .J. & Alexander J. Buck (2012). Does panexperiential holism solve the combination problem?, *Journal of Consciousness Studies*, 19 (9–10), pp. 190–199.

Kastrup, Bernardo (2014). *Why Materialism Is Baloney: How True Skeptics Know There Is No Death and Fathom Answers to life, the Universe, and Everything*, Iff Books.

— (2017). An ontological solution to the mind–body problem, Philosophies 2:2.

— (2018). The universe in consciousness, *Journal of Consciousness Studies*, 25 (5-6) pp. 125-155.

Kastrup, Bernardo. & Edward F. Kelly (2018). Could Multiple Personality Disorder Explain Life, the Universe and Everything?, *Scientific American* blog, 18th June 2018. https://blogs.scientificamerican.com/¬observations/could-multiple-personality-disorder-explain-life-the-universe-and-everything/

Lee, Geoffrey (2007). Consciousness in a space-time world. *Philosophical Perspectives* 21(1), pp 341–374. https://doi.org/10.1111/J.1520-8583.2007.00130.X

Leibniz, Gottfried (1714). La Monadologie [The Monadology], in Nicholas Rescher (ed.) (1991) *G.W. Leibniz's Monadology: An Edition for Students*, Pittsburgh: University of Pittsburgh Press.

Locke, John (1690). *An Essay Concerning Human[e] Understanding, London:* Thomas Basset.

Lockwood, Michael (1984a). Einstein and the identity theory, *Analysis* 44.1, pp 22-25.

—(1984b). Reply to Gordon, *Analysis* 44.2, pp 127-128.

— (1985). Einstein, Gibbins and the unity of time, *Analysis* 45.3, pp 148-150.

— (1989). *Mind, Brain, and the Quantum: The Compound 'I'*, Oxford: Blackwell.

— (1992), lectures, Rewley House, Oxford.

Lloyd, Peter B. (1999a). *Consciousness and Berkeley's Metaphysics*, 281 pp. London: self-published.

— (1999b). *Paranormal Phenomena and Berkeley's Metaphysics*, 353 pp. London: self-published.

— (2003a). *Exegesis of the Matrix*, 301 pp. London: self-published.

— (2003b). Glitches Reloaded, in: Kurzweil, R. (ed.) *Kurzweil Essays* [online], http://www.kurzweilai.net/glitches-reloaded.

— (2006). Mental Monism Considered as a Solution to the Mind-Body Problem, pp 101-144 in Batthyany, Alexander & Elitzur, Avshalom (eds.) *Mind and its Place in the World: Non-Reductionist Approaches to the Ontology of Consciousness*, Frankfurt: Ontos Press. http://www.peterblloyd.com/consciousness/Mental Monism Considered as a Solution to the Mind-Body Problem.pdf

Mathews, Freya (2011). Panpsychism as paradigm, in Blamauer, Michael (ed.) *The Mental as Fundamental*, Frankfurt: Ontos Verlag.

McGinn, Colin (1995). Consciousness and space, *Journal of Consciousness Studies*, 2, pp 220-230.

Nagasawa, Yujin, & Khai Wager (2016). Panpsychism and priority cosmopsychism, in Godehard Brüntrup & Ludwig Jaskolla (eds.) *Panpsychism*, Oxford: Oxford University Press.

Pearce, Kenneth L. (2014), *Language and the Structure of Berkeley's World*, PhD thesis, University of Southern California. http://writings.kennypearce.net/diss.pdf

Pelczar, Michael (2017). *Sensorama: A Phenomenalist Analysis of Spacetime and Its Contents*. Oxford: OUP.

Phillips, Ian B. (2009). *Experience and Time*, PhD thesis, University College, London. http://discovery.ucl.ac.uk/16139/1/16139.pdf

Radin, Dean (1997). *The Conscious Universe: The Scientific Truth of Psychic Phenomena*, New York: HarperEdge.

Robinson, Howard (2009). Idealism, in Beckermann, A., McLaughlin, B.P., and Walter, S. (eds.) *The Oxford Handbook of Philosophy of Mind*, Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199262618.003.0011

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

Russell, Bertrand (1926). The world of physics and the world of sense, pp 106–134 in: *Our Knowledge of the External World as a Field for Scientific Method in Philosophy,* 2nd edn, London: George Allen & Unwin.

—(1927). *The Analysis of Matter,* London: Kegan Paul.

Searle, John (1995). *The construction of Social Reality*, New York: Simon and Schuster.

Shani, Itay (2015). Cosmopsychism: A holistic approach to the metaphysics of experience, *Philosophical Papers*, 44 (3), pp. 389–437.

Śankara, Adi (8th Century). *Atmabodha* [*Self-Knowledge*], translated by Swami Nikhilananda (1974), New York: Ramakrishna-Vivekananda Center.

Strawson, Galen (ed.) (2006). *Consciousness and Its Place in Nature: Does physicalism entail panpsychism?* Imprint Academic.

— (2013). Q&A session in: Galen Strawson on Panpsychism, talk given on 19th September 2013, at St Paul's Church Centre, London. https://www.youtube.com/watch?v=F-I8UABk7KA [7 Apr 2018].

— (2015). interview with Robert Wright on *The Wright Show*, July 2, 2015, https://www.youtube.com/watch?v=ZCo9fJ9EP_k&t=17s

— (2018). interview with Robert Wright on *The Wright Show*, February 23, 2018, https://www.youtube.com/watch?v=1ydqAhPNuyY

Tarski, Alfred (1933). Pojęcie prawdy w językach nauk dedukcyjnych [The concept of truth in the languages of the deductive sciences], Prace Towarzystwa Naukowego Warszawskiego, Wydzial III Nauk Matematyczno-Fizycznych 34,

Warsaw; translation by Woodger, J. H. (ed.) (1956) *Logic, Semantics, Metamathematics: Papers from 1923 to 1938 by Alfred Tarski,* Oxford: Oxford University Press.

Tolman, Richard C. (1917). *The theory of the relativity of motion*, Berkeley: University of California press.

Utts, Jessica (1996) An Assessment of the Evidence for Psychic Functioning, *Journal of Scientific Exploration*, 10 (1), 3-30.

Wachowski, Andy, & Larry Wachowski (1999). *The Matrix*, Warner Brothers motion picture.

— (2003) *The Matrix Reloaded*, Warner Brothers motion Picture.

Weingard, Robert (1977). Relativity and the spatiality of mental events. *Philosophical Studies* 31, pp 279-284.

Wittgenstein, Ludwig (1922). *Logisch-Philosophische Abhandlung* [*Tractatus Logico-Philosophicus*], translated by Ogden, C.K., London: Kegan Paul.

— (1953). *Philosophische Untersuchungen* [*Philosophical Investigations*], translated by G.E.M. Anscombe, London: Macmillan Publishing Company.

— (1958). *The Blue and Brown Books*, Oxford: Blackwell.

---

[1] Chalmers (2018) gives three variants: panpsychism asserts "microphysical entities have mental states" (p 6), or "some fundamental physical entities have mental states" (p 11), or "some fundamental microphysical entities have mental states." (p 12). Chalmers has said that the second one is right. But the differences are not germane to the present discussion. He also uses Strawson's term "micropsychism" (p 12) to mean "the thesis that some fundamental microphysical entities have mental states", in other words microphysical panpsychism.

[2] Chalmers (2018, p 11) also defined it thus: "Micro-idealism is the thesis that all concrete facts are grounded in facts about the mental states of fundamental microscopic entities, such as quarks or photons." That, however, would mean that micro-idealism is not a form of idealism, since it would make the microphysical entities fundamental rather than the conscious minds and their contents. Were micro-idealism to be defined in this way then it would trivially entail panpsychism, but that would not be relevant here as it would say nothing about idealism.