

How Exactly does Panpsychism Help Explain Consciousness?

(References to author's work removed for peer review)

There has recently been a revival of interest in panpsychism as a theory of consciousness. The hope of the contemporary proponents of panpsychism is that the view enables us to integrate consciousness into our overall theory of reality in a way that avoids the deep difficulties that plague the more conventional options of physicalism on the one hand and dualism on the other. However, panpsychism comes in two forms – strong and weak emergentist – and there are arguments that seem to show that weak emergentist panpsychism faces problems analogous to those of physicalism whilst strong emergentist panpsychism face problems analogous to those of dualism. If this proves to be the case, it is hard to see how panpsychism advances the debate.

In this paper, I will develop a new hybrid of the strong and weak emergentist forms of panpsychism, a view according to which subjects of experience are strongly emergent but their phenomenal properties are weakly emergent. I will argue that this hybrid view manages to avoid both the challenges facing physicalism and dualism, and the analogues of those challenges that seem to undermine standard forms of panpsychism. The result is a theory of consciousness we should take very seriously indeed.

Section I gives background. Section II introduces 'hybrid cosmopsychism,' and argues that it avoids the problems that plague other forms of panpsychism. Section III takes a deep dive into the details of hybrid cosmopsychism. Section IV considers further versions of the notorious 'combination problem,' widely seen as the biggest challenge to a panpsychist theory of consciousness. Section V is a brief conclusion.

|

The word 'consciousness' is a little ambiguous. Throughout this paper, I will use 'consciousness' exclusively to mean *phenomenal consciousness*, states which are essentially characterized by what it's like to have them. Pleasure, pain, visual and auditory experiences are fairly uncontentious examples of phenomenally conscious states.

The problem of consciousness is the challenge of accounting for how (phenomenal) consciousness fits into our overall theory of reality. The reality of consciousness is hard to deny: nothing is more evident than the reality of one's own feelings and experiences. And so it seems that consciousness must fit into reality *somehow*; the challenge is to explain exactly how. The two traditional options are physicalism and dualism. Physicalists believe that the facts of consciousness can be accounted for in terms of the facts of physical science.¹

¹ I do not mean by this that physicalists are committed to there being an a priori entailment from the facts of physical science to the consciousness facts; the popular phenomenal concept strategy (Loar 1990/1997, Balog 1999, Papineau 2002, Diaz-Leon 2008) would deny this. Rather I mean that, for the physicalist, the

Dualists believe that conscious states are non-physical properties, residing either in the brain (property dualism) or in a non-physical individual (substance dualism). Both of these two traditional options face deep difficulties, which is what makes the problem of consciousness so hard.

The most discussed worry for dualism is the *causal exclusion problem* (Malcolm 1968, Kim 1989). Many philosophers believe that we have empirical reason to accept that the physical world is *causally closed*, that is to say that every physical event has a sufficient physical cause. If this is true, if everything I do has a sufficient physical cause (e.g. in terms of neurophysiological process in my brain), then it seems that putative non-physical consciousness has nothing left to do, no role to play in generating my behavior. The dualist seems to be driven either to *epiphenomenalism* (the view that consciousness has no causal impact on the physical world) or to *systematic over-determination* (all the effects of consciousness are systematically overdetermined, as every event caused by consciousness also has a sufficient physical cause). Many take these options to be intolerable.

The problem with physicalism is that there seems to be an explanatory gap between the facts described by physical science and the facts of consciousness. The charge is not merely that our current theories are not up to the task, but that there is an in principle bar to the latter being explained in terms of the former, rooted in the very different kinds of concepts we use to characterize physical processes on the one hand and conscious experiences on the other hand. The former concepts are quantitative and third-personal, whereas the latter are qualitative and first-personal. Proponents of the knowledge argument and the conceivability argument argue from this explanatory gap to the conclusion that the postulations of physical science alone are not enough to ground the facts of consciousness.²

Of course, in both of these cases, there is much controversy, with dualists and physicalists claiming that there are satisfactory responses to these arguments.³ However, for philosophers who find these arguments compelling, there is strong motivation to look for an alternative theory of consciousness.

Against this background, the Russellian panpsychist steps in with a promise to avoid both the causal exclusion problem faced by dualism and the explanatory gap problem faced by

postulations made to account for the data of physical science are also sufficient to account for consciousness. I give a more detailed definition of physicalism in XXX. See next footnote for further clarification.

² The 'postulations' of physical science may refer to the dispositional properties expressed by the predicates of physics, or it may refer to categorical properties underlying those dispositions. As I define physicalism, a physicalist may postulate categorical grounds that in some sense take us beyond what physics reveals to us (to this extent, physicalism may resemble Russellian panpsychism, discussed below). But, unlike Russellian panpsychism, physicalists do not do this specifically to account for consciousness but only because they think that in general dispositions need categorical grounds. Hence, the categorical nature of physical properties will not have some special character tailored to account for consciousness. As I have defined physicalism previously (XXX), this comes out as the thesis that physicalists do not commit either to phenomenal or to proto-phenomenal properties at the fundamental level (where proto-phenomenal properties are defined as being involved in facts that a priori entail facts about consciousness, where that entailment isn't wholly dependent on structural features of the grounding fact (where structural features are properties whose essential nature can be captured in a purely mathematico-causal vocabulary)).

³ On the physicalist side, the most popular response is the phenomenal concept strategy (1990/1997, Balog 1999, Papineau 2002, Diaz-Leon 2008). Two defences of dualism are Chalmers 1996 and Gibb 2015.

physicalism. The view is so-called because it is inspired by certain claims Russell made in *The Analysis of Matter* of 1927, although the view Russell defended here was not quite a form of panpsychism. Perhaps the best way to introduce Russellian panpsychism is to say that it has a negative component and a positive component. Let us take each of these in turn.

The negative component is the claim that physics tells us less than we might have thought about the nature of physical reality. Physical science identifies the causal roles associated with fundamental physical properties, such as mass, spin and charge, but doesn't tell us the essential nature of the properties that realise those causal roles. Mass, for example, is characterized in terms of gravitational attraction and resistance to acceleration, and charge in terms of attraction and repulsion. Physics tells us what mass and charge *do* – the causal roles they realise – but not what they *are*.

This negative aspect of Russellian panpsychism identifies a huge hole in our standard scientific story of reality. The positive proposal of Russellian panpsychism is to put consciousness in this hole: physical properties are, in their essential nature, forms of consciousness. Thus, Russellian panpsychism is a radically *non-dualistic* form of panpsychism. It is not the view that matter has physical properties (mass, spin, charge) on the one hand and experiential properties on the other. Rather the claim is that physical properties like mass, spin and charge *are* forms of consciousness.⁴ Physical science tells us what mass does, but, in terms of its essential nature, mass is a form of experience.⁵

Russellian panpsychism assumes a distinction between the causal role a property plays and its essential nature: what a property does versus what it is. Some reject that distinction. Couldn't it be that the essential nature of a property is *given* by its causal role? An affirmative answer to this question is made by proponents of *pan-dispositionalism*.⁶ On this view, once you know everything this is to know about the causal role of mass, you know everything there is to know about what mass is.

Some Russellian panpsychists doubt the coherence of pan-dispositionalism. There is a line of argument, going back to Russell himself, which presses that on a pan-dispositionalist view everything is defined in terms of everything else, which, it is alleged, leads to a kind of vicious circularity.⁷ But even if this argument fails and pan-dispositionalism is a coherent metaphysical option, this is consistent with Russellian panpsychism also being a coherent

⁴ Alternately, the Russellian panpsychist may hold that physical property terms refer to dispositional properties, and hence that physical properties are realized by, rather than identical with, forms of consciousness. The disagreement between this view and the view described in the main text is not one of substance but rather regards how terms in physics are defined. I suspect it is indeterminate whether the linguistic use of physical scientists is such that 'mass' refers to a dispositional property or to a categorical property in terms of the dispositions it realizes.

⁵ Strawson 2006, Chalmers 2015, Brüntrup & Jaskolla 2016, XXX, Mørch 2019, Roelofs 2019, Seager 2019. A closely related view is Russellian *panprotopsychism*, which hopes to explain consciousness by postulating proto-phenomenal properties at the fundamental level (see footnote 2 for a definition of proto-phenomenal): Pereboom 2011, Coleman 2016, McClelland 2013. Russellian panpsychism and panprotopsychism are known collectively as 'Russellian monism'; see Torin & Nagasawa 2015 for a collection of essays on Russellian monism.

⁶ Ellis 2001, 2002, Molnar 2003, Bird 2007, Mumford 2004.

⁷ Russell 1927, Campbell 1976, Robinson 1982, Lowe 2006, XXX.

metaphysical option. The crucial question is whether the theoretical attractions of the latter give us reasons to embrace it.

What are the theoretical attractions of Russellian panpsychism? The ultimate goal, of course, is to account for human and animal consciousness in terms of more basic forms of consciousness. The hope is that by doing this, we can avoid the problems that beset physicalism and dualism. We avoid the problems of physicalism because the arguments that press the explanatory gap target physical-science based account of consciousness, i.e. accounts which aim to explain consciousness in terms of the postulations made to account for the data of physical science, whereas Russellian panpsychists are instead trying to account for consciousness in terms of the 'hidden' essential nature of the physical world.⁸ And we avoid the problems of dualism because consciousness is incorporated into the causally closed physical system; it is only once we distinguish physical processes from consciousness processes that causal exclusion problems arise. At least, this is the hope of Russellian panpsychists.⁹ Unfortunately, it's not clear these problems can be dispensed with so easily.

As we have seen, Russellian panpsychists hope to account for human and animal consciousness in terms of more basic forms of consciousness. But how exactly is this done? There are broadly speaking two options. One option is to postulate basic laws of nature that bridge the gap between consciousness at the fundamental level and the consciousness of an animal. Thus, it might be simply a basic law of nature that when you have conscious particles arranged in such and such a way, consciousness associated with the whole system emerges. This is the *strong emergentist* option. The *weak emergentist* panpsychist, in contrast, tries to account for systems-level consciousness without appeal to such extra laws of nature. On this version of Russellian panpsychism, facts about human or animal consciousness are wholly constituted by facts about consciousness (and perhaps physical structure) at the fundamental level; the latter are nothing over and above the former, in something like that on physicalism the facts about consciousness are nothing over and above the facts of physical science.¹⁰

The problem is that there is reason to think that the strong emergentist panpsychist faces the causal exclusion problem endured by dualism, whilst the weak emergentist panpsychist faces the explanatory gap problem suffered by physicalism. If this proves to be the case, it seems that we've got nowhere.

⁸ Broadly speaking, physicalist accounts of consciousness fit into two categories, which David Chalmers (2002) dubbed *type-A* and *type-B*. Type-B physicalists hold that we do not need to close the explanatory gap to explain consciousness. Type-A physicalists do try to close the explanatory gap, but by attempting to explain consciousness in terms of causal roles. The distinctive approach of the Russellian is to try to close the explanatory gap by postulating a special essential nature to the properties underlying the causal roles identified by physical science. See footnotes 1 and 2 for further clarification of how I understand physicalism in contrast to Russellian panpsychism.

⁹ See, for example, Chalmers 2015 and XXX.

¹⁰ Chalmers (2015) has a similar distinction between constitutive and non-constitutive forms of Russellian panpsychism. However, I want here to focus on whether or not extra laws are needed to account for emergent consciousness facts, and, whilst non-constitutive panpsychists will tend to postulate extra laws, this is not part of the definition of the view. Chalmers (2006) gives an account of strong and weak emergence in epistemological terms.

These concerns are standardly pressed against the background assumption that the fundamental facts for the panpsychist concern fundamental particles bearing very simple forms of consciousness. We can call panpsychism so understood ‘micropsychism.’¹¹ The challenge for the micropsychist is to bridge the gaps between particle-level consciousness and systems-level consciousness. We will later reject micropsychism but can work with it for the moment.

Why think strong emergentist panpsychism faces causal exclusion worries? If we suppose that the micro-level is causally closed, then the new systems-level consciousness that strongly emerges would seem to have nothing left to do in generating behavior, and as a result would seem to be rendered epiphenomenal (Chalmers 2015, XXX).¹² One might also worry that once one commits to explaining human consciousness in terms of special laws of nature, one loses the motivation for adopting panpsychism. Why not just be a property dualist, bridging the gap between physical-science properties and biological consciousness via special laws of nature, rather than postulating consciousness everywhere?

Why think the weak emergentist panpsychist faces explanatory gap worries? One way of pressing this (XXX, Chalmers 2016) is in terms of a variant of the zombie conceivability argument against physicalism. Whereas standard zombies are physical duplicates of humans or animals which lack consciousness altogether, *micro-experiential zombies* are physical duplicates of humans such that (A) all of their most basic parts have conscious experience, but (B) there is no systems-level consciousness, i.e. no consciousness associated with any macro-level part of the organism. Panpsychist zombies seem *prima facie* just as conceivable as regular zombies. If the possibility of regular zombies follows from their conceivability, then the same would seem to be true of micro-experiential zombies; and if the possibility of regular zombies is inconsistent with the truth of physicalism, then the possibility of micro-experiential zombies would seem to be inconsistent with the truth of weak emergentist panpsychism.

This is, of course, a particularly worrying problem given that panpsychism is often motivated via an employment of the zombie argument (to reject physicalism). If that very argument, in a slightly modified form, rules out panpsychism, we seem to have made no progress. Furthermore, there are arguments which purport to show that *any* reductive account of a conscious subject, even one in terms of more basic forms of consciousness, must necessarily fail (XXX, Nida-Rümelin 2014). This is one form of ‘the combination problem’, the banner for

¹¹ Strawson 2006 gives a slightly different definition of ‘micropsychism.’

¹² By saying that the ‘micro-level is causally closed,’ I mean that every event either has a sufficient micro-level cause or has a sufficient cause that is wholly grounded in a micro-level event. One might also worry that if all the causal structure of the physical is at the micro-level, then there is no macro-level physical causal structure for the strongly emergent conscious states to realise, and hence the strongly emergent physical states will count as non-physical rather than physical (given that ‘physical’ states, for the Russellian panpsychist, are the states that realise the causal structure discerned by physical science). However, even if micro-level causal closure is true, there will still be macro-level causal structures realised by the micro-level, and it will be coherent to suppose that those macro-level causal structures are also realised by (and hence over-determined by) strongly emergent consciousness. Even if strongly emergent states are epiphenomenal, and hence do not realise any physical causal structure, we could say that they are physical in virtue of being of the same essential nature as the physical states from which they strongly emerge.

a broad range of challenges to the panpsychist's attempt to bridge the gap from micro-level consciousness facts to the familiar facts of human and animal consciousness.¹³ We will explore some other forms of the combination problem in section IV.

These are certainly very serious challenges. However, I am not entirely persuaded that the problems they raise are *as* serious as the corresponding challenges facing dualism and physicalism. It is not so clear to me that there is a strong empirical case for *microphysical* causal closure, as opposed to a more general thesis of physical causal closure. This is difficult to assess, as the case for causal closure is often stated but rarely defended. But one defence of causal closure (McLaughlin 1994: 278) is the 'no-gap' argument: an inductive argument which starts from the putative observation that we never find gaps in the physical causation observed in the brain. It is open to the emergentist panpsychist to hold that human conscious states *are* physical states of the brain, just ones whose existence is not wholly grounded in facts about their parts (the whole is more than the sum of its parts). It's not immediately obvious why the no-gap argument would rule out strongly emergent physical states.¹⁴

Regarding the motivation for adopting strong emergentist panpsychism over dualism, I would suggest that whilst panpsychism seems at first to be a rather extravagant thesis, upon further reflection it turns out to be much simpler, and more elegant and unified a picture of reality than that offered by the dualist.¹⁵ On a dualist view, there is a radical division between two fundamentally different kinds of property; on the panpsychist view all things of are the same nature (all determinates are of the same determinable). There may be an analogy here with the many worlds interpretation of quantum mechanics: what initially seems profligate turn out, in the judgment of many, to be the simplest account of the data. Even in its strong emergentist form, therefore, we may have reason to prefer panpsychism to dualism.

Turning to weak emergentist panpsychism, the putative gap between particle-level and systems-level consciousness is arguably *less* severe than the gap between the facts of physical science and the facts of consciousness. In the case of the physical/consciousness gap, the kinds of concepts employed on either side of the gap are very different: third-personal, quantitative concepts on the one side, first-personal, qualitative concepts on the other. In the case of the particle-consciousness/systems-level-consciousness gap, the same kinds of concepts are employed on each side: first-personal, qualitative concepts. It is also noteworthy that, whilst a zombie argument seems to apply to both physicalism and weak emergentist panpsychism, it is much less obvious that a version of the knowledge argument applies to weak emergentist panpsychism.¹⁶

¹³ The term 'combination problem' is from Seager 1995. For more detail on the combination problem, see Chalmers 2016 and XXX.

¹⁴ It is plausible that neurophysiological properties are essentially defined as complex properties composed of micro-physical properties. But this is consistent with the fusion form of strong emergentism discussed below (at least if 'X is composed of Y' indicates merely a part/whole relationship and doesn't entail that Y is more fundamental than X).

¹⁵ XXX

¹⁶ XXX

What I want to consider for the rest of the paper, however, is what options are available if one is persuaded that both strong and weak emergentist panpsychism, at least in their standard forms, fail due to the problems discussed above. I believe that there is a form of panpsychism that avoids these concerns, and it is to this that we now turn.

II

In the last section, we worked with the assumption that fundamental entities exist at the micro-level. However, the panpsychist view I want to defend here is a form of *cosmopsychism*.¹⁷ Contemporary cosmopsychist views build on the *priority monism* developed by Jonathan Schaffer (2010). Philosophers tend to assume that fundamental entities exist at the micro-level, and that facts about wholes (tables, chairs, rocks, planets) are grounded in facts about parts (ultimately fundamental particles). This is the ‘Lego’ theory of reality: take a load of little things, stick ‘em together, and you get big things. However, Schaffer has defended a theory which turns this standard assumption on its head. According to priority monism, facts about parts are grounded in facts about wholes, with the chain of grounding ultimately coming to rest in the maximal whole of the universe. One attraction of priority monism is that it fits well with a field-ontology, which in turns fits well with quantum field theory. On a field-ontology, the fundamental entities are universe-wide fields, and particles are identified with local excitations on these fields. On a priority monist view, we can hold that these fundamental fields are basic attributes of the one fundamental individual: the universe.

Priority monism does not entail panpsychism, but there is a closely related form of panpsychism: cosmopsychism, the view that the universe is a conscious subject and that all facts depend on facts about the universe-subject.¹⁸ On the form of cosmopsychism I want to explore here, the structure uncovered by physics is the structure of the experience of the universe-subject, and fundamental fields (including spacetime) are aspects of the experiential field of the universe-subject. Imagine the complete description of reality in the terms of fundamental physics: an incredibly complicated story of patterns of excitation in fundamental fields. That very complicated structure, on the view under consideration, is realized by the experiential field of the universe-subject.

Strictly speaking, cosmopsychism avoids the combination problem discussed in the last section. If we’re not trying to get from conscious particles to systems-level consciousness, then we don’t have to worry about the putative explanatory gap that holds between these two levels. But a moment’s further reflection makes it clear that we’ve merely pushed the lump to another part of the carpet: we now face an explanatory gap between the

¹⁷ Examples of cosmopsychism include Mathews 2011, Shani 2015, Nagasawa & Wager 2016, XXX, Shani & Kepler 2018.

¹⁸ Forms of cosmopsychism according to which all facts are grounded in facts about the universe-subject are forms of priority monism. However, cosmopsychists who think that animal subjects *strongly* emerge from the universe may deny that animal subjects are *grounded* in the universe, in which case they wouldn’t count as priority monists (as Schaffer defines priority monism). Schaffer himself thinks grounding relations are underwritten by basic laws, and hence holds both that animal subjects are grounded in the universe and that animal subjects strongly emerge (as I am defining strong emergence). However, many others, myself included, would take the strong emergence of animal subjects to entail their fundamentality.

consciousness of the universe and the consciousness of humans and animals. It seems perfectly conceivable that we might have a conscious universe, with experience corresponding to the structure of basic physics, without any of the *parts* of the universe being conscious. Perhaps in some sense the universe would instantiate human experience, or at least experience corresponding to the physical structure of human bodies and brains (more on this soon). But what we surely want to make sense of are *multiple subjects* corresponding to different people and animals. According to our pre-theoretical understanding of things, there are at least seven and a half billion conscious subjects in the world, corresponding to the 7.5 billion people in the world. A mere commitment to a conscious universe seems to give us only one. Whereas micropsychism faces a ‘combination’ problem, cosmopsychism faces this ‘de-combination’ problem (Chalmers 2016, XXX).

Of course, this new explanatory gap is only a problem for a weak emergentist form of cosmopsychism. We could instead adopt a strong emergentist form, according to which there are basic laws of nature which ensure that when the conscious universe is in certain specific states, *new* forms of consciousness emerge corresponding to certain of the universe’s parts. But – here we go again! – it seems we would then be back to worries about causal closure: if the level of basic physics is causally closed, there is no causal work left for these new forms of consciousness to do.

The way forward I want to explore is a hybrid of the strong and weak emergentist approaches: strong emergentism about subjects combined with weak emergentism about the phenomenal properties of emergent subjects. We can call this view ‘hybrid cosmopsychism.’

Let’s begin with by clarifying the terminology. I take *phenomenal properties* to be ways of experiencing, distinguished by what it’s like to have them. *Conscious subjects* are the bearers of phenomenal properties: a conscious subject is a thing such that there’s something that it’s like to be that thing. I will understand an *experience* to be an event of a conscious subject bearing phenomenal properties.

According to hybrid cosmopsychism, there are basic laws ensuring that, in certain conditions, certain parts of the universe become conscious subjects. But these strongly emergent subjects do not appear with their own phenomenal property instances; rather they ‘inherit’ phenomenal property instances which, before they emerged, belonged to the universe. That is to say, there are phenomenal properties, $P_1, P_2 \dots P_n$ and an emergent subject E , such that at T_1 , $P_1, P_2 \dots P_n$ belong to the universe and at T_2 , $P_1, P_2 \dots P_n$ cease to belong to the universe and instead belong to E . As E persists through time, it continues to possess a small ‘bubble’ of the phenomenal properties of the fundamental fields, constantly gaining some from/losing some back to the universe around the edges.¹⁹ At the moment E ceases to be a conscious entity in its own right (e.g. at the death of the organism), it relinquishes its phenomenal properties back to the universe. In this way, although there are strongly emergent subjects there are no strongly emergent phenomenal properties; rather

¹⁹ I have framed the view in non-relativistic terms for the sake of ease of illustration. We might ultimately want to frame the laws discussed below in terms of spatiotemporal locations of emergent subjects. I’m grateful to Brad Saad for raising this issue.

phenomenal properties of the fundamental fields are transferred from the universe subject to emergent subjects, before being relinquished back to the universe.

The crucial advantage of hybrid cosmopsychism is that it accounts for the fact that systems with emergent subjects behave exactly the same as systems without emergent subjects. The fundamental properties driving the evolution of physical reality are the properties of the fundamental fields; their causal powers are tracked by fundamental physics, and they are unchanged by the presence or absence of emergent subjects. All that changes when a new subject emerges is that some properties once borne by one subject come to be borne by another subject. But given that the new subject is just bearing properties that would have been borne by the old subject if emergence hadn't occurred, and is bearing them in the exact same location that they would have been borne if emergence hadn't occurred, there are no grounds for thinking that the evolution of physical reality will be affected by the emergence of local conscious subjects.

In one stroke, this removes any of the empirical difficulties associated with dualism and strong emergentist cosmopsychism. The worry with these views is that one would expect that such a radical change in nature as the emergence of a macro-level subject of experience and/or macro-level consciousness would show up in physical systems associated with that emergence. We would expect physical systems that involve macro-level consciousness to behave very differently to physical systems that don't; but this is not what we seem to find, and thus dualism and strong emergentist panpsychism would seem to be disconfirmed. In contrast, the expectation engendered by hybrid cosmopsychism matches what we in fact observe: systems with macro-level consciousness behave no differently to systems that lack macro-level consciousness.²⁰

At the same time, the strong emergentist element of hybrid cosmopsychism also removes the worries that plague weak emergentist forms of panpsychism. Although it is conceivable that that a conscious universe might exist without any of its parts being conscious, it is *not* conceivable that:

- There is a conscious universe and basic laws of nature determining that, in actually instantiated conditions, certain parts of the universe become conscious subjects and inherit certain phenomenal properties of the universe.
- None of the parts of the universe are conscious.

The extra basic laws committed to the by hybrid cosmopsychism bridge the explanatory gap between fundamental and non-fundamental consciousness facts.

In summary, the strong emergentist element allows hybrid cosmopsychism to avoid explanatory worries, whilst the weak emergentist element allows it to avoid causal closure worries. This is exactly the result we want.

²⁰ Is this solution not ad hoc? Haven't we just designed the theory such that a radical change in nature goes unobserved? This change is only unobservable *from the third-person perspective*. It is not at all surprising that a partial description of reality should leave some facts unknown.

III

In this section, we will explore in much more detail the underlying metaphysics of hybrid cosmopsychism.

There is a crucial metaphysical component of the view, something we might call ‘qualia transference.’ In a case of qualia transference, a single phenomenal property instance is transferred from one subject of experience to another. To take a toy example, we might imagine a wicked witch who regularly has headaches. They don’t bother her, however, as she immediately transfers the horrible phenomenal properties involved to a hapless passerby, who suffers the pain in her place. As I am imagining the story, it is not simply that the witch’s headache disappears to be replaced by a qualitatively indiscernible headache in the passerby. Rather, the token phenomenal property instance that leaves the witch’s consciousness is *numerically identical* with the instance that appears in the consciousness of the passerby.

Making sense of this requires taking a stance on the perennial dispute between trope theorists and realists about universals. To illustrate this debate, consider two negatively charged electrons, E_1 and E_2 . According to a realist about universals, the negative charge of E_1 is numerically identical with the negative charge of E_2 . According to a trope theorist in contrast, what we have here are two numerically distinct (but perfectly resembling) negative charges. For a realist about universals, property instances are individuated relative to the bearer of the property. This would bar us from making sense of qualia transference; in the above example, we have two distinct bearers – the witch and the passerby – and hence two distinct property instances. We would be forced, therefore, to re-describe the story as one in which one property instance (the witch’s) disappears to be replaced by another property instance (the passerby’s). It is only if we take property instances to be entities in their own right, i.e. only if we take properties to be tropes, that we can make sense of one phenomenal property instance (one phenomenal trope) transferring from one conscious mind to another. Hybrid cosmopsychism, therefore, is committed to properties being tropes rather than universals.

Whilst it is necessary to take this stance on the metaphysics of properties, I don’t think we need to take a stand on the perennial debate in the metaphysics of substance between bundle theorists and substance-attribute theorists. My characterization of the view so far perhaps suggests a substance-attribute theory, according to which the strong emergence laws transfer phenomenal properties from one substratum to another, e.g. from the cosmic subject to a local subject.²¹ But we might also interpret hybrid cosmopsychism as a kind of bundle theory according to which at the fundamental level there are only *phenomenal tropes* and *facts about which tropes are co-experienced*.²² On the bundle-theoretic version of the view, the emergence of the first subject non-identical with the universe results from phenomenal properties that were previously co-experienced with all phenomenal

²¹ We might think of the composite entities which are non-cosmic subjects as coming into being at the moment of emergence, or we might think of them as merely becoming conscious at the moment of emergence. Which option we take will depend on one’s views about the metaphysics of composition.

²² I am grateful to Tom McClelland for suggesting this possibility.

properties in existence coming to be co-experienced only with each other. Thus, we have a situation in which:

- At T1, phenomenal properties $P_1, P_2 \dots P_n$ are co-experienced with every other phenomenal property in existence
- At T2, $P_1, P_2 \dots P_n$ are co-experienced only with each other.

In this way, we can think of subjects as non-fundamental entities, derived from fundamental facts about co-experiencing. Despite this, this is still a form of strong emergentism about animal subjects, given the reliance on fundamental laws to bring them into being. Whether one goes for the substance-attribute or the bundle-theoretic version of the view will depend on one's preferred view as to the metaphysics of substance. For the rest of the paper I will assume the substance-attribute version, for the sake of simplicity.

Let us spell out in more detail what hybrid cosmopsychism might look like in practice. Contemporary neuroscience suggests that the structure of human consciousness corresponds to high-level information structures in the brain. The global workspace theory (Baars 2002), for example, holds that consciousness corresponds to information that's 'broadcast' throughout the brain, i.e. that's broadly available for many different systems in the brain. While they are in some sense realized by electro-chemical processes, these information structures abstract away from neurophysiological details, i.e. from facts about calcium channels, chemical composition of neurotransmitters, etc.²³

The structures of human consciousness, therefore, are not the structures of basic physics. But plausibly these structures are *in some sense* present in the structure of physics. Consider the complete story of the universe in the language of basic physics. Focus on the bit of that structure that's located in my head, and abstract away from a lot of micro-level detail leaving only coarse-grained causal structure. If you did this in the right way, you'd be able to find information structures isomorphic with my conscious experience. Given this, cosmopsychism implies that the experience of the universe-subject contains structures isomorphic with the structure of animal experience. In other words, if you take the rich and complex experience of the universe – which *ex hypothesi* underlies the structure of physics – focus on the bit of that experience that's located in a particular human head, abstract away from a lot of micro-level detail, you'll be able to find structure isomorphic with the structure of that human's experience.

Of course, just because the structures of animal experience are in some sense present in the structure of cosmic experience, it doesn't follow that there are multiple subjects corresponding to each such structure. What we need to do if we want to account for the facts of animal consciousness is to formulate laws that determine that there are subjects that inherit those aspects of cosmic experience that are structurally isomorphic with animal experience. How could this be done?

²³ It is compatible with cosmopsychism that higher-level information structures are realised by more fine-grained physical states, so long as the more fine-grained physical states are states of, or grounded in states of, the universe.

I suggest two principles specifying the kind of laws we need: the *Localization Principle* and the *Thinning Principle*. The Localization Principle is fairly straightforward: it says that we need laws ensuring that, in addition to the universe subject, certain local physical systems are also conscious (namely, those which we know correspond to subjects of experience). In other words, the Localisation Principle ensures that there are subjects in the brains of humans and other animals. The Thinning Principle is a little more nuanced; it says that we need a law ensuring that emergent subjects inherit a 'thinned-out' version of the experience contained in the spatial region they occupy, such that animal subject inherits only those aspect of experience that realise the right information structures (where the 'right information structures' are the ones contemporary neuroscience tells us correspond to animal experience). This notion of 'thinning out' is a conceptual innovation that we now need to clarify.

In a sense, the simplest case of qualia transference would be one in which all the phenomenal properties of one individual are transferred wholesale to another individual, such that what it's like to be the first individual at T1 is exactly the same as what it's like to be the second individual at T2. But we might also imagine that *some aspects* of the total phenomenal properties of one subject could be transferred to another subject.²⁴ Consider the following case. Subject A has a detailed visual experience as of a country scene; perhaps she sees a lake surrounded by trees, with birds flying overhead. Subject B then inherits some aspects of the A's phenomenal properties, and hence ends up with a similar experience but with less detail filled in; let's say she has an experience as of a lake but with no trees or birds. We can say that B has a 'thinned-out' version of A's experience.

We can now apply this to the theory under consideration. According to hybrid cosmopsychism, there is an incredibly complex experience in my brain, corresponding to all the details of its physical structure right down to each electron and quark. But that's not *my* experience; rather, it is experience borne by the universe. As discussed above, it is plausible that structures isomorphic with my experience are to be found in the experience borne by the universe in my head, if you abstract away from enough detail. The purpose of the thinning law is to ensure that the subject of my experience inherits only those aspects of the phenomenal properties of the universe-subject which realise the right information structures (i.e. the information structures that contemporary neuroscience tells us correspond to my experience). The result: the subject of my experience has a thinned-out version of the universe's experience, with just enough detail taken out to ensure that I have experience with precisely the structure that contemporary neuroscience tells us I have.

To be clear, the Localisation and Thinning Principles are not statements of putative fundamental laws, but principles which state what is required from fundamental laws. What the specific laws will be is in part an empirical question. To make it more concrete, we can explore the theory in more detail by conditionally assuming the truth of a specific scientific proposal concerning the correlation between physical facts and the facts of conscious experience, namely that associated with the integrated information theory (Oizumi, Albantakis, Tononi 2014) of consciousness, also known as 'IIT.'

²⁴ I'm not thinking of aspects as being a different category of thing from phenomenal tropes. Rather some phenomenal tropes are aspects of other phenomenal, e.g. the phenomenal trope of phenomenal hue is an aspect of the phenomenal trope of phenomenal red.

IIT revolves around the notion *integrated information* – the amount of which is represented by the Greek letter ‘ ϕ ’ – a concept which proponents of IIT attempt to give a mathematically precise definition of. We can avoid technical details here, except to say that, according to IIT, consciousness is correlated with *maximal* ϕ . That is to say, consciousness exists at the level at which there is most ϕ . If we want to know if some X is conscious, we need to ask two questions:

1. Does X have certain proper parts such that there is more ϕ in those proper parts than there is in X?
2. Is X part of some greater whole such that there is more ϕ in that greater whole than there is in X?

If the answer to both (1) and (2) is ‘no’, then X is conscious, according to IIT. If there is more ϕ in the molecules making up my cup of coffee than there is in the cup of coffee considered as a whole, then IIT predicts that my cup of coffee is not conscious. If there is more ϕ in the cerebral cortex than there is either (i) in any of the neurons making it up or (ii) in any whole of which the cerebral cortex is a part, then IIT predicts that the cerebral cortex is conscious. IIT not only tells us which physical entities are conscious, but also identifies the physical structures that correspond to the structures of human consciousness: roughly, they are those structures which support high levels of ϕ .

IIT is strictly speaking inconsistent with hybrid cosmopsychism, because on the latter view both the universe and I can be conscious at the same time. According to IIT, this is impossible: either the universe has more ϕ than me or vice versa; if the former, then I am not conscious as I am part of a greater whole which has more ϕ than I have; if the latter, then the universe is not conscious, as it has a part which has more ϕ than it itself has.²⁵

However, we can consider a slightly modified form of IIT according to which *non-fundamental consciousness* is correlated with maximal ϕ . Thus, we take it as given that the universe is conscious and hold that the principles of IIT tell us where there are conscious subjects over and above the universe subject. This modified theory, call it IIT*, will be empirically indiscernible from original IIT, and thus in so far as we are judging matters on empirical grounds, this change ought not to make a difference.²⁶

²⁵ At least, hybrid cosmopsychism is inconsistent with IIT if we are thinking of the local subjects as parts of the universe, which I am. One could envisage a more dualistic version in which the local subjects are non-physical and hence not strictly speaking parts of the universe; the question of compatibility of this view with the ‘exclusion postulate’ of IIT (according to which two conscious subjects cannot exist in a part-whole relationship) might be less clear.

²⁶ IIT is at least partly supported by appeal to five ‘axioms’ of consciousness, justified on the basis of introspection, which are then translated into five corresponding postulates, which proponents of IIT allege that a system must exemplify in order to satisfy the corresponding axiom. It’s the fifth postulate (the ‘exclusion postulate,’ according to which two conscious subjects cannot exist in a part-whole relationship) which is inconsistent with IIT*. The justification for the collapse postulation seems to me decidedly weak (as, for example, argued by David Chalmers <http://consc.net/slides/iit.pdf>), so I’m not too concerned that IIT* violates it.

Assuming IIT* as the correct theory as to how non-fundamental consciousness is correlated with physical processes, our localisation and thinning laws will come out roughly as follows:

- Localisation Law: For any local (i.e. non-identical with the universe) physical system P, P is conscious subject iff (there is more ϕ in P considered as a whole than there is in any proper parts of P) and (there is more ϕ in P considered as a whole than there is in any whole of which P is a proper part).
- Thinning Law: For any non-cosmic subject S, S inherits those aspects of cosmic experience located within it that support high levels of ϕ .²⁷

If IIT* is the correct theory of how non-fundamental consciousness is correlated with physical processes, then the above two principles will predict that humans and non-human animals have the kind of consciousness experience they actually do have. Again, this is exactly the result we want.

IV

David Chalmers (2016) divides combination problems for panpsychism into three categories:

- Subject combination problems How do we derive animal subjects from fundamental subjects?

²⁷ Further complexities may be introduced when it comes to the *persistence* of emergent subjects. I am grateful to Brad Saad for prompting me to think about this. Suppose we are working with an *endurantist* model of persistence, such that a subject is wholly present at each moment at which it exists. We could say that a subject continues to exist so long as its parts compose a system with maximal ϕ . But there are tricky cases. Consider, for example, the following possibility: at T1 certain atoms compose a system of maximal ϕ , at T2 all but one of those atoms cease to compose a system of maximal ϕ , whilst at the same moment one of those atoms joins some other atoms to compose a system which at that moment achieves maximal ϕ . Should we say that the subject that was wholly present at T1 is also wholly present at T2? Or has a new subject been brought into existence? Or suppose at T1 certain atoms compose a system of maximal ϕ , and at T2 50% of those atoms compose one system with maximal ϕ whilst 50% of those atoms compose a distinct system of maximal ϕ . Presumably at T2 we have a situation in which there are two distinct subjects, X and Y. Should we say that one of these subjects, either X or Y, is identical to the subject that existed at T1? If so, which one? X and Y cannot both be identical with the subject that existed at T1, as in that case (by the transitivity of identical) X and Y would be identical with each other, which *ex hypothesi* they are not. Or should we say that the subject that existed at T1 ceases to exist and two brand new subjects come into existence? There may be no principled way of decided which of these options is to be preferred. Matters are simpler if we adopt a four-dimensionalist view, according to which subjects are temporal parts of spacetime worms. We can then say that there are spacetime worms corresponding to all of the options outlined above, leaving us the conceptual choice of which temporal parts we want to group together and label as parts of a single person. The downside of this option, as I have explored in XXX, is that subjects of experience – the entities which have consciousness non-derivatively – do not persist through time. Why not identify subjects with spacetime worms? Perhaps there's something that it's like to be a spacetime worm, but it would be a very strange kind of consciousness not at all like the kind of consciousness we pre-theoretically associate with human beings. At best, we could say that spacetime worms derivatively instantiate normal human consciousness (i.e. there is a spacetime worm that 'has human experience' in the sense of having temporal parts with human experience) and is thereby a human subject in a derivative sense. In any case, developing a detailed account of the persistence of subjects/persons will go hand in hand with developing a specific empirical account of the emergent laws (these issues may, for example, give us grounds for doubting IIT), whereas in this paper I am merely sketching the general view.

- Structure combination problem: How do we derive the structure of animal experience from the structure of fundamental consciousness?
- Quality combination problem: How do we derive the qualities of human experience from the qualities of fundamental consciousness?

In the context of cosmopsychism, these are ‘de-combination’ rather than combination problems, but the essential challenge is the same: getting from fundamental to non-fundamental facts of consciousness.

We have already explored in detail the hybrid cosmopsychist’s answer to the subject de-combination problem: the gap from the fundamental to non-fundamental subjects is bridged by basic laws of nature. We can now consider the structure and quality de-combination problems.

The structure de-combination problem

Some readers might have been wondering why the view under consideration needs to be construed as a form of *cosmopsychism*? Couldn’t we conceive of a *micropsychist* version according to which in the beginning, as it were, there were only conscious particles, and then emergent subjects inherit the phenomenal properties that they would have had if emergence hadn’t occurred? This view would be a form of the *fusion* view defended by Hedda Hassel Mørch (2014) and William Seager (2016). We talked earlier about the dispute concerning whether wholes are grounded in parts (the standard ‘Lego’ view) or parts are grounded in wholes (the priority monist view). On Mørch’s view, we find examples of both within the one universe. For a system which isn’t an emergent macro-level subject, facts about the system are grounded in facts about its parts; for a system which *is* an emergent macro-level subject, facts about parts are grounded in facts about the whole. On Seager’s view, the particles simply fuse during the creation of the emergent subject, ceasing to exist in the process. Thus, whilst standard strong emergentism has two fundamental layers of reality (the micro-phenomenal and the macro-phenomenal) co-existing at the same time – we can call this ‘layered emergentism’ – the fusion view holds that the micro-phenomenal layer of reality is in some sense *replaced* by the macro-phenomenal layer when fusion occurs.

On either form of the fusion view, it is not required that we think of the micro-level properties as persisting post-emergence just as they would have done if emergence has not happened. But we *might* construe the fusion view in these terms, and if we do, we will end up with a micropsychist version of the view we have been exploring in this paper. We can call this the ‘hybrid fusion’ view.

The trouble with the hybrid fusion view is that it suffers from the most pernicious version of the structure combination problem, a problem Daniel Stoljar dubbed ‘the structural mismatch problem.’²⁸ Roughly speaking, this is the difficulty panpsychism has in accounting for the apparent mismatch between the structure of the brain and the structure of human

²⁸ Nagasawa & Wager (2016) also hold that cosmopsychism can help avoid the structural mismatch problem.

consciousness. David Chalmers (2016) has recently construed the structural mismatch problem as being raised by four *prima facie* plausible but inconsistent propositions:

1. Micro-phenomenal structure is isomorphic to microphysical structure
2. Micro-phenomenal structure constitutes macro-phenomenal structure
3. Microphysical structure constitutes macrophysical structure.
4. Macro-phenomenal structure is not isomorphic to macrophysical structure.

I'm not sure this is quite the right way to put the problem. As discussed earlier, it's plausible that structures corresponding to human consciousness are in some sense present in the physical structure of the brain. As I see it, the structural mismatch problem arises because it's very hard to see how we could build up to consciousness with the structure of human (or non-human animal) experience merely by aggregating the intrinsic phenomenal properties of particles. An experience that is formed just by unifying a huge number of 'pinpricks' of consciousness would have very little structure.

Perhaps not every micropsychist view need think about mental combination as mere aggregating, but it seems that the proponent of the hybrid fusion must do, given that they think phenomenal properties are unchanged by the emergence of a macro-level subject. Crucially, if my consciousness is wholly formed from inheriting the intrinsic phenomenal properties of particles, it follows that the *relations* between those particles do not feature in my experience. But it is an obvious empirical fact that the structure of my consciousness corresponds to high-level structures in the brain, and that the latter structures in some sense reflect relationships between particles.²⁹ On the micropsychist view under consideration, relations between particles may play a role in determining *when* emergence happens, and perhaps *which* intrinsic phenomenal properties are inherited by the emergent subject, but the relations cannot themselves feature in the consciousness of the emergent subject. The only kind of consciousness that could be formed in this way is a structureless aggregate of a large number of 'pinpricks' of consciousness all experienced at once.

The proponent of the fusion view could avoid this worry by holding that the micro-level properties *change* as they are fused into the emergent subject. Understood this way, the fusion view would not be a form of the hybrid view (strong emergence of subjects combined with weak emergence of phenomenal properties) but simply a form of strong emergence (of both animal subjects and animal phenomenal properties).

Unfortunately, this version of fusion view faces a serious empirical concern. The problem is that one would expect such a radical change in nature – the basic properties of physics beings removed and replaced – to result in a significant change in behavior. We'd expect it to be the case that whilst systems without strongly emergent subjects/phenomenal properties evolve according to the laws of physics, systems with strongly emergent conscious subjects/phenomenal properties evolve according to a different set of laws. But this is not what we seem to see: systems evolve according to the same basic laws of physics regardless of whether or not they involve emergent conscious subjects.

²⁹ I use the word 'reflect' so as to avoid committing to these structures being dependent on more fundamental facts about particles, which cosmopsychists deny.

This is essentially the analogue of the causal exclusion problem faced by dualists and layered strong emergentist panpsychists. A dualist or a layered strong emergentist panpsychist has all the basic properties of physics instantiated in the brain, but also has an extra 'layer' of properties, corresponding to human consciousness. We then have the worry that the effects of the higher layer are systematically over-determined by the lower layer. Strictly speaking, the fusion view avoids over-determination worries, as there is only ever one layer of a system: pre-fusion the one layer is at the level of particles, post-fusion it is at the systems-level. However, this way of avoiding the letter of the causal exclusion worry leads to the empirical concern discussed in the last paragraph.

Hybrid cosmopsychism is able to avoid both of these problems. It avoids the structural mismatch problem by starting not with a large number of pinpricks of consciousness, but with a unified field with rich structure, structure that corresponds not only to particles but also to the relationships between particles.³⁰ The thinning law then abstracts away from much of the detail of that structure to get to the structure of human/animal consciousness. Hybrid cosmopsychism also avoids the empirical problem. Given that it is the properties of fundamental physics running the show before and after the local subject emerges, one would not expect systems containing emergent subjects to behave any differently to systems not containing emergent subjects.

The quality de-combination problem

The most discussed way of pressing the quality combination problem is via what Chalmers calls 'the palette problem.' On a standard micro-based form of Russellian panpsychism, the fundamental forms of consciousness correspond to the fundamental properties of physics: mass, spin, charge, etc. The palette problem arises from the fact that there are only a very small number of fundamental properties in physics, whereas there seem to be an incredibly wide range of qualities in human experience: at the very least the seemingly very different qualities involved in the different sense modalities (colours, for example, seem very different from sounds). It is at the very least puzzling how such a diverse set of qualities could be derived from such a small number of qualities. How is the richness of human experience painted from such a meagre palette?

The palette problem doesn't seem to apply to cosmopsychism. At the fundamental level, rather than having merely a small number of fundamental properties of particles, we have a smallish number of fundamental fields. And there is no reason to think that a given field will instantiate a single homogenous quality throughout the whole universe. The cosmopsychist can hold, and presumably should hold, that a field bears a wide variety of qualities depending on the structure it exemplifies, and on how it interacts with other fields. For sure, fields exemplify intrinsic properties at points or regions, and a cosmopsychist will need to identify those intrinsic properties with distinctive forms of consciousness. Crucially, however, a cosmopsychist does not suffer from the micropsychist's limitation of *only* being able to postulate fundamental forms of consciousness corresponding to intrinsic properties of points or small regions. She can also postulate fundamental forms of consciousness corresponding to the distinctive structures and interactions of whole fields, multifaceted

³⁰ I don't mean to imply by this that particles are fundamental: see last footnote.

kinds of phenomenology shaped by how the fields blend and dance. This allows for a rich palette of qualities without having to add structure beyond that of physics itself. For the cosmopsychist, the qualities we find in human consciousness are simply a thinned-out version of what it's like for the cosmic fields to exemplify the structure of the human brain.

The fact that cosmopsychism avoids the palette problem is great advantage of the view. However, avoiding the palette problem is not the same thing as offering a full solution to the general quality combination problem, i.e. a detailed account of what the fundamental qualities are and how human consciousness emerges from them. In order to do this, we'd need a theory that told us:

- For any structure of the fundamental fields S , what qualities are borne by the fundamental fields when they exemplify S .
- For any structure S and any way of thinning out that structure T , what qualities are borne by a physical system that inherits T .

I'm not sure whether it is possible for human beings to formulate a full solution to the quality combination problem, as articulated above. The big problem is that we don't have access to the qualitative character of so many experiences. As Nagel famously observed, we cannot fully know what it's like to be a bat, and the same is true of very, very many animal experiences, never mind the non-biological forms of emergent experience a panpsychist theory may commit to. Moreover, even with regards to the consciousness in our own heads, we have access only to thinned out forms of the qualities enjoyed by the universe and no direct access to the 'fully filled-in' version. Whilst lacking so many of the pieces of the jigsaw, it's not surprising that we can't see how they all fit together, and it may forever be beyond our ken. As Sam Coleman (2015) has speculated, drawing on the work of Charles Hartshorne, it could be that all the qualities of human experience – colours, sounds, smells, tastes – lie on a continuous spectrum, in the way the visible colours do, but that we are incapable of apprehending this fact because we have no conception of certain qualities in that spectrum. Perhaps there are certain qualities 'in between' colours and sounds, such that if we only had experiences involving those qualities we could move along the spectrum of colours and sounds as seamlessly as we can move between different shades of blue. And perhaps we would need to apprehend these full spectra of qualities in order to form a true understanding of how the qualities of biological experience are generated.

I don't want to be too pessimistic. Maybe careful work, involving introspection and neuroscience, could find hints of a bigger quality structure within which our experiences lie; Coleman is optimistic in this regard. Perhaps scientists of the future will be able to test predictions as to the qualitative character of certain structures not naturally exemplified by human brains by artificially creating those structures in a human brain. Unless and until we have significant numbers of the science of consciousness community working within a panpsychist framework, we won't have a good sense of what potential progress might be made on the quality combination problem.

In any case, formulating a detailed answer to the quality combination problem is a task for (philosophically informed) science to answer. The crucial question for armchair philosophers is whether there are any a priori reasons to doubt that a given panpsychist theory could in

principle account for the diverse range of qualities we find in human experience. The fact that cosmopsychism avoids the palette problem removes such doubts.



The discussion of this paper has been implicitly assuming that a theory of consciousness should be subject to two constraints:

The External Constraint – The theory should be consistent with the empirical data (this constraint is in the background of causal exclusion worries).

The Internal Constraint – The theory should avoid explanatory gaps, by ensuring that one could in principle deduce the facts of animal consciousness from the facts from which they strongly/weakly emerge (where we include in the emergence base, the fundamental laws, if there are any, which underwrite the emergence).

The hardness of the problem of consciousness can be traced to the difficulty of satisfying both of these constraints at the same time. In hybrid cosmopsychism, we find a theory that is up to the job. It satisfies the internal constraint via its appeal to basic laws determining the emergence of animal subjects and specifying which phenomenal properties they inherit from the universe (the strong emergentist component). It satisfied the external constraint because the emergence of local subjects does not introduce new fundamental properties (the weak emergentist component). No doubt there are further problems to address and further details to be worked out. However, in broad outline, hybrid cosmopsychism is a theory of consciousness we should take very seriously indeed.

References

Albahari, Miri (2020) 'Beyond cosmopsychism and the great I am: How the world might be grounded in universal "advaitic" consciousness,' in Seager 2020.

Alter, Torin & Yujin Nagasawa (2015) (Eds.) *Consciousness and the Physical World*, Oxford University Press.

Baars, Bernard J. (2002) 'The conscious access hypothesis: Origins and recent evidence,' *Trends in Cognitive Sciences*, 6: 1, 47-52.

Balog, Katalin (1999) 'Conceivability, possibility, and the mind-body problem,' *Philosophical Review* 108: 4, 497-528.

Bird, Alexander (2007) *Nature's Metaphysics: Laws and Properties*, Oxford: Oxford University Press.

Brüntrup, Godehard & Ludwig Jaskolla (2016) *Panpsychism*, New York: Oxford University Press.

Campbell, Keith (1976), *Metaphysics: An Introduction* (Dickenson).

- Chalmers, David J. (1996) *The Conscious Mind: In Search of a Fundamental Theory*, New York: Oxford University Press.
- Chalmers, David J. (2002) 'Consciousness and its place in nature,' in his *Philosophy of Mind: Classical and Contemporary Readings*, New York: Oxford University Press.
- Chalmers, David J. (2006) 'Strong and weak emergence,' in P. Davies & P. Clayton (Eds.) *The Re-Emergence of Emergence: The Emergentist Hypothesis From Science to Religion*, Oxford University Press.
- Chalmers, David J. (2015) 'Panpsychism and panprotopsychism,' in Alter & Nagasawa 2015.
- Chalmers, David J. (2016) 'The combination problem for panpsychism,' in Brüntrup & Godehard & Ludwig Jaskolla 2016.
- Coleman, Sam (2015) 'Neuro-cosmology,' in P. Coates & S. Coleman (Eds.) *Phenomenal Qualities: Sense, Perception and Consciousness*, Oxford: Oxford University Press.
- Coleman, Sam (2016) 'Panpsychism and neutral monism: ow to make up one's mind,' in Brüntrup & Jaskolla 2016.
- Diaz-Leon, Esa (2008) 'Defending the phenomenal concept strategy,' *Australasian Journal of Philosophy*, 86: 4, 597-610.
- Ellis, Brian (2001) *Scientific Essentialism*, Cambridge: Cambridge University Press.
- Ellis, Brian (2002) *The Philosophy of Nature: A Guide to the New Essentialism*, Montreal: McGill-Queen's University Press.
- Gibb, Sophie (2015) 'Defending dualism,' *Proceedings of the Aristotelian Society* 115: 131-146.
- Kim, Jaegwan (1989) 'Mechanism, purpose, and explanatory exclusion,' *Philosophical Perspectives*, 3: 77-108.
- Loar, Brian (1990/1997) 'Phenomenal states,' originally published in J. Tomberlin (Ed.) *Philosophical Perspectives 4: Action Theory and Philosophy of Mind*, Atascadero, CA: Ridgeview; reprinted in substantially revised form in N. Block, O. Flanagan, & G. Guüzeldere (Eds.) 1997 *The Nature of Consciousness: Philosophical Debates*, Cambridge, MA: MIT Press.
- Lowe, E. J. (2006), *The Four-Category Ontology: A Metaphysical Foundation for Natural Science* (Oxford University Press).
- Malcolm, N. (1968) 'The conceivability of mechanism,' *Philosophical Review* 77: 44-72.

Mathews, Freya (2011) 'Panpsychism as Paradigm?' in M. Blamauer (Ed.) *The Mental as Fundamental* (Process Thought Series; edited by N. Rescher, J. Seibt and M. Weber), Heusenstamm: Ontos Verlag, 141-56.

McClelland, Tom (2013) 'The neo-Russellian ignorance hypothesis: A hybrid account of phenomenal consciousness,' *Journal of Consciousness Studies* 20: 3-4, 125-51.

McLaughlin, Brian (1994) *Epiphenomenalism, A Companion to the Philosophy of Mind*, in S. Guttenplan (Ed.), Oxford: Blackwell, 277-288.

Molnar, George (2003) *Powers: A Study in Metaphysics*, Oxford: Oxford University Press.

Mørch, Hedda Hassel (2014) *Panpsychism and Causation: A New Argument and a Solution to the Combination Problem*, Ph.D. thesis, University of Oslo.

Mørch, Hedda Hassel (2018) 'Is Consciousness Intrinsic?: A Problem for the Integrated Information Theory,' *Journal of Consciousness Studies* 26: 1-2, 133-162.

Mørch, Hadda Hassel (2019) 'Is consciousness intrinsic? A problem for the integrated information theory,' *Journal of Consciousness Studies* 26 (1-2): 133-62.

Nagasawa, Yujin & Wager, Khai (2016) 'Panpsychism and Priority Cosmopsychism' in Brüntrup & Jaskolla 2016, 113-129.

Nida-Rümelin, Martine (2013) 'The Argument for Subject Body Dualism from Transtemporal Identity Defended,' *Philosophy and Phenomenological Research* 86: 3, 702-714.

Oizumi, Masafumi; Albantakis, Larissa; Tononi, Giulio (2014) 'From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0,' *PLOS Computational Biology* 10: 5.

Papineau, David (2002) *Thinking about Consciousness*, Oxford: Clarendon Press.

Robinson, Howard (1982) *Matter and Sense* (Cambridge University Press).

Roelofs, Luke (2019) *Combining Minds: How to Think about Composite Subjectivity*, New York: Oxford University Press.

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

Schaffer, Jonathan (2010) 'Monism: The priority of the whole,' *Philosophical Review* 119: 1, 31-76.

Seager, William, E. (1995) 'Consciousness, information, and panpsychism,' *Journal of Consciousness Studies* 2: 272-88.

Seager, William, E. (2016) 'Panpsychism Infusion,' in Brüntrup & Jaskolla 2016, 229–248

Seager, William, E. (2019) *The Routledge Handbook of Panpsychism*, Routledge.

Shani, Itay (2015) 'Cosmopsychism: A Holistic Approach to the Metaphysics of Experience,' *Philosophical Papers* 44: 3, 389–417.

Shani, Itay & Keppler, Joachim (2018) 'Beyond Combination: How Cosmic Consciousness Grounds Ordinary Experience,' *Journal of the American Philosophical Association* 4: 3, 390-410.

Strawson, Galen (2006) 'Realistic materialism: Why physicalism entails panpsychism,' 13: 10-11.