

## Does physicalism entail panexperientialism?

### Abstract

*In Realistic Monism (2006a), Galen Strawson argues that physicalism combined with a denial of ontological emergence entails a panpsychist ontology. A far more common position in philosophy of mind is that of reductive physicalism. I will defend Strawson's argument and conclude that a form of panpsychism does follow, provided a defence of Strawson's assumptions about fundamental science and the related concept of emergence. However, I will conclude that a substantive panexperientialist thesis appears to end up with many of the same problems it was formulated to avoid.*

In the light of continuing debate about the mind body problem, I will consider the possibility of a solution coming from an often ignored position in philosophy, that of panpsychism. I will do this by defending Galen Strawson's (2006a) argument, that physicalism entails a form of panpsychism. The paper is split into three main sections. Section I unpacks Strawson's argument that physicalism entails (what I will call) panexperientialism. Section II defends the argument focusing on the dismissal mainstream interpretations of physicalism. Section III assesses the plausibility of the resultant panexperientialist thesis. Overall, I will argue that the entailment *does* hold given defence of certain subsidiary metaphysical assumptions. I will conclude that the resultant thesis, if found acceptable has very interesting repercussions but will not be able to put the mind body problem to bed once and for all.

Panpsychism is, at its simplest, is the doctrine that everything that has a physical aspect has a mental or conscious aspect too. This is a way of saying that some kind of consciousness or mentality is fundamental to all reality; this can be understood in various ways. Prima facie it is often thought of as the 'hippy' idea that rocks and tables have human traits like self awareness, representational abilities, goal directedness and such like. It is hard to see what would

motivate this view or why you would want to. However, ‘weaker’ forms of panpsychism have been advanced throughout the history of philosophy<sup>1</sup> as ways of reconciling our ideas about mentality and experience with our seemingly incompatible idea of the physical. The notion Strawson’s argument will be advancing is just that, in some way, everything physical *is* experiential and it is in virtue of this that humans and other animals can have complex and sophisticated conscious experiences. Skrbina points out that this goes back to Spinoza’s view that: “There is one ultimate substance of the world, but it exhibits two faces – ‘physical’ from the ‘outside’ and ‘mental’ from the ‘inside’”(Skrbina, 2006, p153). So, rather than the substance dualistic thesis wherein whole minds are attached to anything and everything; panpsychism is construed as a substance monist position where there is just one kind of substance, the essence of which, is experience.

If a panpsychist position like this can be properly articulated and defended, it seems it should be more widely taken note of, as it could have the potential to dissolve, or significantly simplify, the mind-body problem. However, it is plainly a position which is massively unintuitive to most, for one reason or another. Therefore, I will now advance an argument in defence of panpsychism, in order to root out those metaphysical commitments that ground its *prima facie* rejection by other philosophers of mind. I will do this by deconstructing Galen Strawson’s argument that panpsychism is entailed by physicalism.

First though, I have found discussion of panpsychism to have a slippery relationship with traditional categorisations in philosophy of mind. So, before reconstructing the argument I will discuss carefully of my usage of the terms: physicalism and experience.

## **Physicalism and Experience**

---

<sup>1</sup> See Skrbina, 2003, p4-46. This article highlights panpsychist leanings in philosophy from the Pre-Socratic to the late 20<sup>th</sup> Century. It finishes by identifying five key, ongoing, approaches: “quantum physics”, “information theory”, “process philosophy”, “part-whole hierarchy” and “non-linear dynamics.”(p43) Of these, this essay will broadly advocate the process philosophy approach.

Physicalism grew from the widespread rejection of substance dualism and captures the intuition that all real or concrete phenomena in the universe are physical. But what does the word 'physical' encompass such that this is a meaningful statement? In normal speech it is used to refer to 'material', 'solid' or, 'spatially located' substances, subject to deterministic laws; basically, the stuff posited to exist by physics. However, when we look to modern physics, its posits include things like waves and fields, massless particles and probabilistic interactions as constitutive of those larger, visibly tangible physical substances. So, it seems these things must be considered physical too, perhaps even more paradigmatically so than visible scale objects. Also, given the revisionary nature of science's history<sup>2</sup>, and the still unexplained phenomena within its domain<sup>3</sup> it would be wrong to say modern physics is true or complete. Therefore for physicalism to get off the ground many physicalists have seen it as necessary to limit their claim to one about what a true and complete physical science *would* say. So I will take physicalism to be the broad claim that: All that exists are the posits of a true and complete physical science.

I will use 'experience' to refer to the subject matter of philosophy of mind. Anything conscious; possessing of a phenomenal or qualitative feel; or fundamentally, any object which there is something it is like to be it<sup>4</sup> - has experience, or has an experiential aspect. My usage of the term is not intended to mean anything stronger. Much of the characterisation of a mental state tends to presuppose what are arguably contingent, human attributes. When I say that something has an experiential aspect I do not, necessarily, mean it has a rich representational field, or memories, emotions or even awareness of its own existence. I just mean that there is something it is like to be that thing. I think this is the only kind of definition which will capture distinction between what we think of as mental and non-mental in a way that will

---

<sup>2</sup> Just consider the move from the classical elements (earth, wind, fire etc), via vitalism, Newtonian physics, special relativity and now quantum theory.

<sup>3</sup> For example, gravity's ability to cause action at a distance is inconsistent with scientific laws of motion.

<sup>4</sup> Nagel, 1979, p175

continue to be coherent when discussing the plausibility of the existence of mental states that are radically simple or alien to our own.

This definition also allows me to advance another piece of terminology. If the ‘psyche’ is removed from panpsychism and replaced by ‘experience’ as I just defined it, we remove the unwanted connotation of there being whole ‘psyches’ present in objects like rocks and tables. ‘Psyche’ implies a mind and our conception of a mind tends to presuppose complexity. So, by using panexperientialism from now on, I distance the discussion from this unwanted extra implication. Panexperientialism is still a *version* of panpsychism because it still says that everything has something it is like to be it. This move has been made before. The term was coined by David Ray Griffin<sup>5</sup>, and is used by Sam Coleman in discussion of Galen Strawson’s argument<sup>6</sup>. Although Strawson claims the term is coextensive with his use of ‘panpsychism’ it is certainly not coextensive with everyone else’s usage so I will stick to panexperientialism.

I must also note at this point that my definition of physicalism superficially allows experience to be defined as physical. We know that experience exists by considering our own case<sup>7</sup>; and we know that ‘the physical’ is everything posited by a true and complete physical science – surely then, experience *must be* one of the posits of a true and complete physical science? However, this seems to remove the sense in which ‘physical’ refers to physical things. If ‘physical’ can mean anything that exists, and it is being invoked as a definition of *what* exists then it is not saying anything at all<sup>8</sup>: “One may debate the exact boundary of physical science: but unless some human sciences, of which psychology will be our exemplar, lie beyond its pale, physicalism, as a doctrine about mind will be vacuous”(Crane and Mellor, p186, 1990). This highlights an ambiguity in my definition of physicalism. The way physicalism is generally used as a substantive theory presupposes that the ‘physical’ does *not*

---

<sup>5</sup> Griffin, 1997

<sup>6</sup> Coleman, 2006

<sup>7</sup> I will be assuming that eliminativism is false and that experience definitely exists in this essay. This is because there is already ample discussion of eliminativism elsewhere and I do not wish to get bogged down with it here.

<sup>8</sup> MacPherson, 2006, p74

include experiential phenomena. So mainstream physicalists generally want to say that: all that exists are the posits of true and complete physical science *and* that these posits are, in themselves, wholly non-experiential. They must then, of course, provide some explanation as to how we come to have, or think we have, experience. The panexperientialist will want to say something different, but which is also substantive.

I will use ‘physicalism’ to denote my initial broad definition of physicalism and ‘non-experiential physicalism’ to refer to physicalism plus the substantive claim made by most mainstream physicalists. So, panexperientialism and non-experiential physicalism are both positions consistent with physicalism though not with one another.

## **I – The Argument**

Strawson argues that: given the reality of experience and a commitment to physicalism, we *cannot* hold non-experiential physicalism, so we *must* admit that the posits of true and complete physical science will include experience. His discussion of why this is the case is necessary to make sense of his following move - from a physicalism including experiential posits to one which yields a panexperientialist ontology. Therefore, I will begin with first part of the argument. In order to do this I must introduce Strawson’s view of completed science and of emergence.

### **Complete Science as Fundamental Science**

Strawson uses the word ‘ultimates’ to refer to: “a fundamental physical entity, an ultimate constituent of reality, a particle, field, string, brane, simple, whatever”(2006a, p9) and says that his substantive assumptions concerning the physical are:

“(1) there is a plurality of ultimates (whether or not there is a plurality of types of ultimates)

(2) everything physical (everything physical that there is or could be) is constituted out of ultimates of the sort we actually have in our universe

(3) the universe is spatio-temporal in its fundamental nature.”(2006a p9)

These are common philosophical assumptions. Many of us hold an atomistic conception of the nature of reality. The methodology of scientific explanation in general appears to be to split up, or unpack, the constituents of complex phenomena. We do this in order to explain their properties in terms of their parts, and their relationships with one another. This seems a natural and intuitively correct process. Why? Well surely because we are confident that the behaviour of a large, complex phenomenon *is* determined by the interactions of its parts. To ask for a scientific explanation is to ask for this kind of explanation, and this methodology implies that we believe that the deeper we go, the closer we get to the truth. So, if science can be true and complete, and contemporary science has its methodology right, then its posits will be the most microscopic, finest grained, ones possible. Therefore, whatever fundamental science talks about will be impossible to further unpack, or divide. Strawson’s comments demonstrate that he holds that ‘true and complete physical science’ would be fundamental in the way I have described. As I have argued, this is a natural and common assumption to make. Indeed, Sam Coleman points out that this assumption:

“...is pervasive amongst philosophers of mind, especially physicalists, who tend to express it by saying that facts about the *microphysical* determine facts about the chemical, the biological and so on.”(2006, p40)

I will refer to ‘fundamental science’ to mean the systemisation of scientific knowledge whose posits are indivisible and which fully determine any complex phenomena. This is consistent with there being conceptual divisions between the various roles of an ultimate in forming different complex phenomena while it is still impossible to divide that ultimate physically.

Physicalism becomes the doctrine that ‘ultimates’ fully determine all that exists. Non-experiential physicalism then says that these ultimates will be not be experiential while panexperientialism says that they will. A position seems to have been missed out here: the one known in the literature as ‘micropsychism’, (although in this essay it will make more sense to call it microexperientialism). This is just the view that only *some* ultimates are experiential while others are not. I will discuss the rejection of microexperientialism presently, but first I will talk about Strawson’s ground for rejection of non-experiential physicalism. He does this by arguing that experience is not the kind of property which can *emerge* from non-experiential ultimates. Emergence is the key concept here, so I will try to explain what emergence means; and what understanding of it is consistent with a fundamental view of science.

## **Emergence**

An emergent property is one which appears in an entity of a certain level of complexity but which is absent from the constituent parts of that entity. A paradigm example is liquidity: A single molecule cannot be a liquid but a group of molecules can, provided they are the right type, are close together and at the right temperature and pressure. When combined in the correct way the excited molecules can slide past one another but cannot fly away; thus liquidity ‘emerges’. Thomas Nagel describes emergence as an: “epistemological condition” (1979b, p82), whereby a property *seems* to appear unannounced at a level of complexity but which, if you look closer, is actually entirely explained by the interactions of its parts. Strawson contrasts this epistemological emergence with what he calls “brute”(2006a, p18) emergence,

whereby a property appears in a complex which is *not* determined by its parts, and says that the latter is plainly impossible. This follows directly from commitment to the fundamental view of science I discussed in the previous section. If such a property did appear, it would not exist according to fundamental science, because it is not derived from fundamental sciences' posits. It could not be, in any way, predicted or explained by science; it would be "by definition a miracle every time it occurs"(2006a, p18). Also, since the actions of the ultimates already causally determine the actions of a whole, any brute emergent feature would appear to lead to overdetermination, or itself be epiphenomenal.

So, *if* a complex phenomenon has an emergent property, *then* that property must, in principle, have an explanation in fundamental science. Even if that explanation is very hard to determine, it must be there. This claim is powerful because it captures the inconsistency in holding a belief in fundamental nature of science at the same time as holding that experience 'supervenes' on the physical, *without* being entirely determined by it. This leaves the possibility that experience appears to emerge in complex phenomena but *is* entirely determined by them, in which case the emergence is only *epistemological*, so, in principle, soluble. So, to argue against non-experiential physicalism one needs to deny that experience can be determined by non-experiential parts. In other words one needs to claim that the property 'experience' *cannot* appear only in complex systems, because, it cannot, even in principle, be derived from non-experiential ultimates.

The orthodox way to express this claim in philosophy is via the concept of reduction. Saying experience cannot emerge from non-experiential parts, is equivalent to saying that experience does not 'reduce' to non-experiential parts. For this reason I will, in places, refer to reduction rather than emergence.

Strawson makes this claim by way of forceful arguments from the unintelligibility of such a reduction:



“You can get liquidity from non-liquid molecules as easily as you can get a cricket team from eleven things that are not cricket teams. In God’s physics, it would have to be just as plain how you get experiential phenomena from wholly non-experiential phenomena. But this is what boggles the human mind.”(2006a, p15)

However, this strategy is bound to be inconclusive because unintelligibility does *not* entail impossibility. This is the premise which will require the most work to defend and I will move on to discuss it at length in section **II**. But, we are now, finally, in a position to construct the argument formally. Here it is:

**P1.** Physicalism states that everything that exists can be derived fully from the posits of a true and complete physical science

**P2** Our idea of a true and complete physical science is of a fundamental one.

**P3.** Experience exists.

**C1.** If Physicalism is true, then all experience must be derivable fully from the ultimates posited by fundamental science.

**P4.** There are no ‘brute’ emergent properties.

**P5.** Experience cannot reduce to non-experiential parts.

**P6.** Experience is not brute emergent.

**P7.** Therefore, at least some ultimates must, themselves, have an experiential aspect.

**P8.** If some do, they all probably do.

**P9.** Everything posited by fundamental science has an experiential aspect.

**C2:** If physicalism is true then everything has an experiential aspect.

**C3:** If physicalism is true then panexperientialism is true.

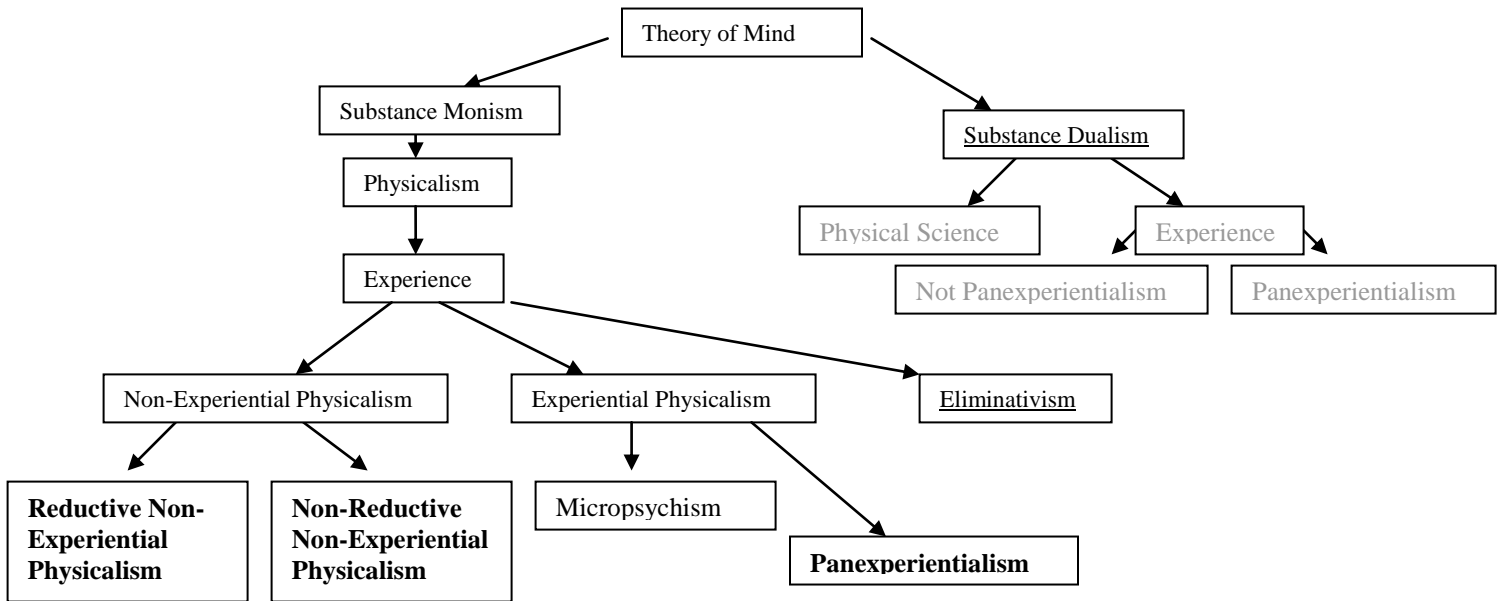
In order to reject the argument you must either: Not be a physicalist, (at least, not hold that my weak description of physicalism is true), so reject **P1**; reject the idea that a true and complete science would be fundamental in its nature i.e. **P2**; be an Eliminativist and reject **P3**; or argue that experience does reduce to non-experiential parts, and reject **P5**.

**P8** is, on the face of it, a very weak premise. Strawson just says that he would: “bet a lot against their being such radical heterogeneity at the bottom of things” (Strawson, 2006, p25). However, this is perhaps not as weak as it sounds. Rejection of just **P8** still leaves a you committed to some form microexperientialism, so panexperientialism just has to be more attractive than this option for the argument to work. Intuitively, it seems likely it would be possible, in principle, build a brain out of just about any material. If broken down far enough and rearranged at a subatomic level one could use anything from hydrogen to plutonium. Therefore, as Sam Coleman points out: “To many it is plausible that brain making ultimates are ‘refundable’”(2006, p48). Consider, as a possible *reductio*, that there could be a planet devoid of those particles but otherwise similar to ours, there is no reason why natural selection should not go on there, rather as it does here, but that it could result only in the creation of ever more sophisticated automatons; creatures with minds could never evolve there. This does not gel very well with our intuitions. In short then, there really does appear to be something less acceptable about the idea of microexperientialism than the idea of panexperientialism. Therefore, even detractors of panexperientialism are unlikely to accept the argument up to **P8** and then not accept the move to panexperientialism. And so, if the rest of Strawson’s argument can be made to stand up, it seems that panexperientialism follows more readily than microexperientialism. For this reason I will now lay this **P8** to rest and go on to concentrate on the two main contentious parts of the argument.

In the literature, the most common move against this argument comes from rejection of **P5**; Peter Carruthers & Elizabeth Schechter(2006), Frank Jackson(2006), David Papineau (2006) and C. Wilson (2006) all argue this view explicitly, while others make mention of it. Some take an eliminativist route, rejecting **P3**; this includes Philip Goff (2006) to an extent and Georges Rey (2006). H.P Stapp (2006) and Michel Bitbol (2003) reject **P2** by offering an alternative account of completed science where facts about the smallest things *do not* have to upwardly determine facts about the larger ones.

More common than rejection of particular premises, however, are claims that no remotely plausible panexperientialist thesis can be formulated; Thomas Nagel (1979b), Colin McGinn (2006); or that attempts to formulate such a thesis fell foul of the assumptions used in the argument from physicalism; Philip Goff (2006), Fiona MacPherson (2006). Nagel expresses this kind of response well, saying that: “[panpsychism] appears to follow from a few simple premises, each of which is more plausible than its denial, though not perhaps more plausible than the denial of panpsychism”(1979b, p182), and after further discussion of the position concludes that: “[panpsychism] should be added to the list of hopelessly unacceptable solutions to the mind-body problem.”(1979b, p193)

I will not discuss rejection of **P1** or **P3** further. The argument is only intended to hold from an acceptance of physicalism and I have already indicated I do not want to get bogged down with eliminativism. Instead, I will spend section **II** addressing **P2** and **P5**, and section **III** attempting to formalise a plausible panexperientialist thesis which does not, itself, fall foul of the argument from physicalism. First though, here is a diagram make it clear which positions I have already rejected (underlined), and which depend on discussion of **P2** and **P5** (in bold):



## II – More on Fundamental Science and Emergence

In this section I will discuss the two central lines of objection, which I identified in the previous section, to the argument from physicalism to panexperientialism. First I will discuss the rejection of **P2**. I will argue that it entails rejection of some fairly dearly held intuitions about the nature of science are wrong; but that it is a plausible position which would remove the motivation for panexperientialism. I will then move on to discuss **P5**. I will argue that there seems to be no way to make sense of experience emerging from non-experiential parts and no way to rule out microexperientiality in the parts of any attempted reduction. I will conclude that the intuition that **P5** can be rejected is grounded in a category mistake<sup>9</sup>.

### Fundamental Science II

In section **I** I pointed out that a fundamental, upwardly determined view of science is a natural view to have. I argued that the way we make sense of things is to divide them up. And this

<sup>9</sup> Bitbol, 2007, p1 and credit to Gilbert Ryle's notion of "category mistake"

leads us inexorably towards a real or utopian goal of finding the ‘ultimates’. However, a glance at the history of metaphysical philosophy suggests that using an atomistic framework to make substantive claims may be a mistake. Hume’s atomism was arguably to blame for the weaknesses in his system of the mind<sup>10</sup>. And Wittgenstein’s led his *Tractatus* into a positivistic mire which he later rejected. In this case, it seems that denial of brute (ontological) emergence was entirely predicated on Strawson’s commitment to fundamental science. If completed science turned out not to be atomistic/fundamental in this way then we would not say phenomena necessarily had to be reducible to anything. They would not have to reduce to description in the language of ultimates because there would be no ultimates. This would make space for the position which Sam Coleman calls ‘macro nonreductive physicalism’<sup>11</sup> in which properties are not determined ‘upwardly’ from lower levels.

But surely, if we want to hold on to determinism, we cannot deny that complex phenomena are determined by their parts! Well, Schrödinger points out that it is a law of quantum mechanics that: “best possible knowledge of a whole does not necessarily include (best knowledge of) its parts. [...] The whole is in a definite state, the parts taken individually are not”(1935, Section I, 11) It is very clearly not worth getting lost in quantum theory in this essay but it is worth taking on board the gist. This rule appears to mean that when we get down to talking about subatomic particles we are no longer in a position to define the wholes in terms of their parts. Instead, the parts appear to be probabilistic, all we can say definitely about them is that they exactly constitute the whole. So, this would suggest that our intuition has turned out to be (somewhat) mistaken. We can deny that complex phenomena are determined by their parts at least in cases of subatomic particles. Instead, the parts seem to be determined by the complex phenomena. Or rather, as Michel Bitbol puts it: “there is a mutual relation of co-arising between the whole and its parts. The many parts are still taken as constitutive of the

---

<sup>10</sup> See Pears, 1990, p6 for a discussion of this.

<sup>11</sup> Coleman, 2006, p41

whole, but at the same time the whole is irretrievably involved in the definition of the parts.”(2007, p9)

So, perhaps we can reject that a true and complete science would have to upwardly determine experience. If we reject this, we do not have to say that brute emergence is impossible. And if we accept brute emergence is possible then we do not have to worry about reducing experience to non-experiential parts. Therefore, if we reject fundamental science the argument from section **I** does *not* entail panexperientialism (but instead leaves room for the non-experiential non-reductive physicalist as the diagram above shows).

## **Emergence II**

I will now move on to discuss **P5**. I left this in the last section by saying that Strawson claims experience cannot be reduced to non-experiential parts, because such a reduction appears to be utterly unintelligible. But I also noted that unintelligibility is not enough. This is because one can still hold, what Sam Coleman calls ‘a posteriori physicalism’<sup>12</sup>. This is the view that experience does reduce, but such a reduction need not be explicable to us.

So how can we be sure that such a reduction cannot ever happen, not even in principle? Well, I think it is because what we know of science, we know indirectly, via our experience. This knowledge is fundamentally mathematical and law based. Russell puts it like this:

“It is only [the physical world’s] mathematical properties that we can discover. For the rest, our knowledge is negative... The physical world is only known as regards certain abstract features of its space-time structure – features which, because of their abstractness, do not suffice to show whether the physical world is, or is not, different in intrinsic character from the world of mind”(Russell, 1948, p240)

---

<sup>12</sup> Coleman, 2006, p45

This point is not trivial. What science tells us is always ‘extrinsic’ in the sense that all it can ever say about the properties of an object are due to the ways it behaves in relations with other objects. If an object can be broken down then sure, it can be described in terms of its parts and their relational behaviour, but that is just the same thing on a smaller scale. If this process is repeated until we reach ultimates, it seems that all science could ever be able to say is that a particular ultimate moves, or spins, or oscillates or whatever according to a particular law. This has led to some philosophers, like W. G. Lycan, to suggest that maybe the ultimate constituents of reality have *no* intrinsic properties: “perhaps the nature of a subatomic particle is exhausted by the totality of its relations to other things”(2006, p67) However, I would urge that intuitively this is not the case. Were there a universe composed of only one ultimate particle, hence no relata, surely that particle would still have some intrinsic nature<sup>13</sup>! Surely, there is something more than just relational qualities, some underlying character of an ultimate that causes it to have the relational characteristics that it has. There must be something that fixes its characteristics through time. To pick on a classic example: “Socrates is taller than Plato” is a relational property. But this is *because* Socrates is intrinsically Xcm tall, and Plato is intrinsically Ycm tall, and X is greater than Y. Each would still possess the same height without the other.

So, I think that we are right to assume that ultimates, if they exist, must be possessing of intrinsic properties. And, I think Russell is right to say that such properties *cannot* be reduced to their scientific description. It follows that a complete reduction of a brain to its smallest parts and their relationships will be able to say nothing of the experiential nature, or lack of it, in any one part. It will not even be able to say anything positive or negative about experience in any complex of those parts for that matter. It is not until we have something

---

<sup>13</sup> Seagar, 2006, p141

extremely complex and unified like a working human brain that it is able to indicate, through displaying intentionality and through language, that it is, itself, conscious. Even linguistic assertion is vulnerable to sceptical arguments about other minds<sup>14</sup>, so perhaps a safer example is the idea of a complete reduction of your own brain. Whatever convincingly sufficient condition you may find for experience it will not explain *why* experience appears when this condition is met, or prove indubitably that experience is not present lower down. This suggests that the unintelligibility of the reduction really is insoluble.

Attempts to pass over this problem have been made by way of trying to separate our concepts of experiential phenomena from the experiential properties themselves. This is known as ‘the phenomenal concepts strategy’<sup>15</sup>. Roughly, the claim is that the unintelligibility of reduction stems from conceptual inconsistency between the way we talk about experiences and the way we talk about physics, not from the impossibility of the reduction. Loar distinguishes our *concepts* of phenomenal properties and those properties themselves and claims that: “...We can have it both ways. We can take the phenomenological intuition at face value, accepting introspective concepts and their conceptual irreducibility, and at the same time take phenomenal qualities to be identical with physical-functional properties...”(Loar,1997, p1). I cannot see how any such attempt can avoid entailing the brute emergence of the actual phenomenal ‘what it is like’-ness *from* said phenomenal properties. Michel Bitbol points out that this appears to be built into the starting point:

“Emergence concerns *properties*, to wit features that are intersubjectively accessible, and that can then be described in a third-person mode. [...] consciousness is no ordinary ‘property’ in this sense, but rather a situated, perspectival, first-person mode of access.”(Bitbol, 2002, p1)

---

<sup>14</sup> For instance: Chalmers Zombies

<sup>15</sup> Loar, papineau



He says that this shows we have made a ‘category mistake’ in thinking that the two can be reconciled in this way. But since, by this stage in the panexperientialism argument, brute emergence has been rejected, we cannot leave this as it stands. Either our concept of the non-experiential or of the experiential must yield. This is where panexperientialism offers to plug the gap. There clearly is a conceptual inconsistency here. But perhaps it is due to our inadequate conception of what Loar calls the ‘physical-functional’, rather than to an uncooperative set of experiential concepts. As Russell asserted, our mathematical knowledge of objects is only of their extrinsic properties. So, perhaps these extrinsic properties *all* occur in virtue of the intrinsic character of ultimates. And perhaps there seems to be an inconsistency only because of our habitual, human-chauvinist, *insistence* in not ascribing experience to objects unlike ourselves. If this is the case it is the ‘physical-functional’/scientific concepts which are at fault not our experiential ones.

I think that this is the best defence of **P5** that I can give. It is a hard sell, but I think it makes a fair amount of sense. Rejection of **P5** depends on a defence of the phenomenal concepts strategy or something like it, and nothing appears to be forthcoming. However, in parallel, defending panexperientialism requires making sense of the emergence of human and animal experience from experiential ultimates. If the micro to macro experiential emergence proves as inexplicable as the emergence of experience per se then the argument will not motivate panexperientialism at all. It would just, if anything, tell against the coherence of physicalism in general. For that reason, I will now go on to discuss panexperientialism itself.

### **Section III – Panexperientialism Explained**

So far I have unpacked and defended Galen Strawson's argument to wit physicalism entails panexperientialism. I have identified its major weaknesses and claimed they pertain to assumptions about completed science and the related concept of emergence. I have shown that, provided you have the similar beliefs to Strawson concerning the nature of completed science, the argument does motivate the idea that we should be microexperientialists. And, that panexperientialism is the most plausible formulation of microexperientialism. However, it remains to be shown that we can form a substantial idea of panexperientialist ontology. We need some idea of what we are arguing for, such that the claim is meaningful and does not contradict the argument that motivates it.

Therefore, I will now try to explain somewhat how I think we should understand panexperientialism. In the process of doing this I will identify and discuss what I see as the biggest problems for the view. I will call them **The Combination Problem**, **The Nothing Left Problem** and the **Intrinsic Properties Problem**.

The immediate question is: what on earth is a panexperientialist ontology like? How does a panexperientialist view the world? For my part I think that it means that experience is fundamental and omnipresent and that a mind is just a particular special complex, or unified bundle, of these experiences. This is to say that experience *is* the quality of matter, in virtue of which, brains are able to represent and form memories and complex intentions. Meanwhile in the case of many other complex phenomena, their parts are not arranged in a fashion that leads to these human traits. However, whatever qualities they do have *are* still in virtue of their experientiality. Larger, more complex systems of matter than ourselves could, by the same token have correspondingly larger, richer phenomenal experiences even if we cannot make sense of this from our perspective. In this way panexperientialism has the advantage of avoiding human-chauvinism at the cost of being a theory that is easy to make sense of. It

removes any arbitrarily human-contingent condition for consciousness and allows that any complex bundle of ultimates can, and indeed does, form some kind of mind. But without those arbitrary conditions, any resultant ‘mind’ is likely to be radically, incomprehensibly, different from our own.

### **The Combination Problem**

Well, an immediate problem looms. For something to have experience, for there to be something it is like to be a thing, seems to imply that there is a subject of that experience. And to say that experiential ‘ultimate’ subjects can be combined to form macroexperiences just does not seem to make sense. This has proved the most central and problematic argument against panexperientialism. Here is William James putting the point very strongly 120 years ago:

“take a hundred of them [feelings], shuffle them and pack them as close together as you can (whatever that may mean); still each remains the same feeling as it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean. There would be a hundred-and-first feeling there, if when a group or series of such feelings were set up, a consciousness *belonging to the group as such* should emerge. And this 101<sup>st</sup> feeling would be a totally new fact; the 100 feelings might, by a curious physical law, be a signal for its *creation*, when they came together; but they would have no substantial identity with it, nor it with them, and one could never deduce the one from the others, or (in any intelligible sense) say that they *evolved* it.” (James, 1890, p162)

It does not seem necessary to take this *reductio* much further but, without some limiting law of combination, would not every possible complex of ultimates, not just the ultimates themselves and an arbitrary much larger complex, also be subjects of experience? Even if we were to

swallow the absurdity of there being ‘minisubjects’ in a ‘macrosubject’ would we not have to say there are many, many more overlapping experiences? Without a limiting principle this would amount to every possible combination of any number of the available ultimates. And even worse, the creation of an extra subject, as James, appears to be another case of brute emergence! And of course if panexperientialism needs to evoke the concept of brute emergence then it is no longer motivated by Strawson’s argument. However, the combination problem may not be fatal.

I think that the way to address this problem is to reject subjecthood, or at least a strong conception of subjecthood, as a necessary part of experience. Consciousness is habitually thought to exist without self-awareness; for instance, in young children and some animals. Few people would be inclined to deny that young children are conscious until they become aware of who and what they are in relation to their surroundings. I have already urged that a good definition for consciousness is a minimal one. Consider something like Block’s distinction between ‘phenomenal’ and ‘access’ conscious<sup>16</sup> in which the former is in some sense ‘raw’, without an associated sense of knowing that it exists, therefore without a sense of itself as a subject. If you take the knowledge of its own subjecthood away from experience; and take away perception, memory, any sort of reasoning ability and leave just a timeless ‘raw feel’ then it does not seem like you have enough left for the application of the word ‘subject’ to pose a significant problem. If experience only starts identify itself, e.g. have a sense of itself as a subject, in areas of patterned complexity (like a human brain) then it is not inexplicable to say that there is simultaneously something it is like to be all sorts of smaller parts of that brain. Indeed, this leaves panexperientialism in no worse a state than many other theories of mind. The ‘binding’ problem - of how parts of conscious experience make a unified whole is a large

---

<sup>16</sup> See Ned Block, 1995

and far reaching problem for philosophy of mind<sup>17</sup>. However rejection of subjecthood leads us to the next problem.

### **The Nothing Left Problem**

Without the characteristics of experience constitutive of subjecthood are we really still holding onto anything that should really be described as experiential? Shouldn't we be calling this *panprotoexperientialism* and reverting to admitting that it is a case of epistemological emergence? In other words: have we just been renaming non-experiential ultimates as experiential for no good reason this whole time? The intuition here seems to be that once you have stripped enough away from our experience to make sense of it as applying to ultimates there is nothing left of what we think of as experience.

Now, it is true that we cannot imagine *what* it would be like to be an ultimate, or some very simple complex, but this does not amount to a rejection of the view because again, unintelligibility does not entail impossibility. We cannot imagine what it would be like to be anything, except creatures like ourselves, with any probable degree of success. Nagel has made this point well; that the extent of our imagination is clearly limited to the projection of our own experiences or simple cases of conceptual rearrangement<sup>18</sup>. If we have not had something like an experience ourselves, then we cannot imagine it. Recall what I said about reductive non-experiential physicalism and phenomenal concepts: Given tension between experiential and non-experiential concepts; and the primacy of experience over experimentation - it is the non-phenomenal concepts which should yield. And in the absence of our ability to image what something is like, we should conclude just that we cannot imagine it, not that there is not something there.

---

<sup>17</sup> See, *Unity of Consciousness* in the *Stanford Encyclopedia of Philosophy*

<sup>18</sup> Nagel, 1979a, p169

Coleman addresses both the combination and the nothing left problem with an attempt to describe how macro-experiential abilities could emerge acceptably given the existence of micro-experiences:

“If the major distinction between the micro-phenomenal and the macro-phenomenal is that the latter has rich representational capacities, then perhaps... ..when suitable unities of intrinsically conscious ultimates are placed in [a brain-world causal] relation, they represent. Starting out from conscious bedrock, perhaps the feeling that naturalizing intentionality leaves something out can now also be overcome” (Coleman, 2006, p51)

However, I still think that this construal of experience in terms of ‘raw feels’ is such a weak one that it remains problematic. It seems plausible to construe raw feels as in, some sense, ‘latent’ or ‘unactualised’ until put in a brain. This is to say, there could be nothing ‘it is like’ to be those ultimates until the point at which they are utilised in a brain. If this is the case the theory is in trouble again. The reductive non-experiential physicalist would have the non-experiential properties he needed to ground the emergence of experience. Even granting their raw experientiality, there is certainly no strong sense in which a ‘raw feel’ could be considered eligible for moral consideration; in which it would be happy to be ‘alive’ or that its destruction could be a morally wrong act. And I think that a large part of why we try to identify what is, and is not alive is tied up in this sense that things that are alive have something to lose. For this reason, it seems like the experientiality we suppose ultimates to possess may be so weak that it does not constitute there being ‘something it is like to be’ them.

### **The Intrinsic Properties Problem**

My attempts at explanation notwithstanding, this is still an undeniably mysterious hypothesis.

It is not at all clear *how* experiential ultimates could combine even if they are not subjects of experience. This is because the conception we have, tentatively, formed is of experience as the intrinsic nature of ultimates *in virtue of which* they have the extrinsic nature they do. This leaves the extrinsic, relational effects which these ultimates have firmly within the realm of physical rather than experiential description. If a complex phenomenon has an acceptably emergent property it seems it must be due to the way its parts, the ultimates, react with one another; on the ‘outside’ so to speak. It may well be the experiential nature of ultimates that cause them to have the extrinsic qualities they do, but it does *not* follow that the experience carries over to the larger phenomenon.

Can we say that the experiential properties of ultimates combine in an additive way, separately to the physical relationships between the ultimates? After all, Fiona MacPherson asserts that what we have here is really a property dualist position: it appears to be consistent with the definition of property dualism to say that: “human beings are composed of one type of substance but have conscious mental properties which are not mainstream physical properties (as well as mainstream physical properties)”(2006, p81). If experiential properties are construed as separate from physical ones then we run into the kind of problems that the theory is intended to avoid: Are the experiential properties causally efficacious? If so how do you reconcile them with the causal completeness of the physical? And if not, must they be overdetermined or epiphenomenal?

Strawson replies to MacPherson by asserting that he does not see his view as property dualism; that he sees the distinction between an experiential and physical description of an object as merely conceptual: “like the distinction between triangularity and trilaterality.”(2006b p196) He continues, saying that the property-object distinction is one that he feels *cannot* be applied in all cases of metaphysics. To support this he mentions F.

Ramsey's claim that: "the whole theory of universals is due to a mistaking [...] a characteristic of language[...] for a fundamental characteristic of reality". This conforms to my feelings concerning universals. The distinction between an object and its properties appears to me, to be a conceptual one, and one created, primarily, for reasons of linguistic convenience. And *linguistic convenience* is not a good reason to multiply your posited existents. Therefore, I do not think that the conceptual distinction between an object and its properties should be conflated with an ontological difference. Perhaps this is why I am attracted to panexperientialism. While panexperientialism *is* technically a case of property dualism, to call it property dualism is rather disingenuous. The view has to be that 'experiential' is coextensive with 'physical' and they are distinct only in the way that triangularity is distinct from trilaterality. So, the intrinsic nature of reality is 'experiential' while the extrinsic is 'physical'; or, as I quoted from Spinoza at the beginning: "There is one ultimate substance of the world, but it exhibits two faces – 'physical' from the 'outside' and 'mental' from the 'inside'"(Skrbina, 2006, p153).

So, what I have called the combination problem, the nothing left problem and the intrinsic properties problem really all seem closely related. What they amount to is a strong attack on the coherence of a panexperientialist thesis. No good answer is forthcoming as to how microexperientiality is supposed to constitute macroexperientiality even if you ditch subjecthood. Nor is it easy to decide if there is really any point in calling such minimal things as ultimates to be, 'experiential'. The only halfway satisfactory answer to the above problems seems to be just to assert that implications of panexperientialism's actuality do not easily fit into our standard (dualistic) conceptual framework. But then, this is unlikely to be enough to convince those who initially thought of panexperientialism as 'massively unintuitive'.



## Conclusions

I will now conclude my discussion of panexperientialism. I have argued that Strawson's argument does work but that panexperientialism is not enough of a substantive theory to convince many, even if they accept the argument. I think that the best grounds for dismissal of the argument should come from rejection of Strawson's scientific assumptions. The scientific community is moving toward a different metaphysic and perhaps as philosophers we should anticipate this. But, regardless, a panexperientialist thesis is open to the charge of incoherence and mysteriousness, as much or more than any other extant theory of mind. It is attractive to some because it appears to avoid chauvinism about what things have minds. This is a virtue that, arguably, many other theories of mind do not have. In short, I think that while its defence has proved inconclusive, panexperientialism should *at least* be: "added to the list of hopelessly unacceptable solutions to the mind-body problem." (Nagel, 1979b, p193)

## **Bibliography**

**Block, Ned 1995.** *On a confusion about the function of consciousness.* In *Behavioral and Brain Sciences* Vol. 18 pp227-247.

**Bitbol, Michel 2007.** *Ontology, Matter and Emergence.* In *Phenomenology and the Cognitive Science*, 6, pp293-307.

**Brook, Andrew and Rayment, Paul 2009.** *The Unity of Consciousness.* In *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), Edward N. Zalta (ed.). Sourced at: <http://plato.stanford.edu/archives/spr2009/entries/consciousness-unity/> on 26/3/09.

**Carruthers, Peter and Schechter, Elizabeth 2006.** *Can Panpsychism Bridge the Explanatory Gap?* In *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp32-40.

**Coleman, Sam 2006.** *Being Realistic.* In *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp40-53.

**Griffin, David Rey 1997.** *Panexperiential physicalism and the mind-body problem* in *Journal of Consciousness Studies* Vol. 4 No. 3 pp248-68

**Jackson, Frank 2006.** *Galen Strawson on Panpsychism* in *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp62-65.

**James, William 1890.** *The Principles of Psychology*, 1983, Harvard University Press.

**Loar, Brian 1997.** *Phenomenal States* (Second Version) from <http://www.nyu.edu/gsas/dept/philo/courses/consciousness97/>, 23/03/09.

**Lycan, William G. 2006.** *Resisting ?-ism.* In *Journal of Consciousness Studies* Vol. 13, No. 10-11. pp65-72.

**MacPherson, Fiona 2006.** *Property Dualism and the Merits of Solutions to the Mind-Body Problem.* In *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp72-90.

**McGinn, Colin 2006.** *Hard Questions.* In *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp90-100.

**Nagel, Thomas 1979a.** *What is it like to be a bat?* In *Mortal Questions* (1979), pp165-180, Cambridge University Press

**Nagel, Thomas 1979b.** *Panpsychism* In *Mortal Questions* (1979), pp181-196, Cambridge University Press.

**Papineau, David 2006.** *T* in *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp100-110

**Rey, Georges 2006.** *Better to Understand Human than World Psychology* in *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp110-117.

- Russell, Bertrand 1948.** *Human Knowledge: Its Scope And Limits*, London: Routledge.
- Sbrinka, David 2003.** *Panpsychism as an Underlying Theme in Western Philosophy* in *Journal of Consciousness Studies*, 10, No. 3, pp4-46
- Sbrinka, David 2006.** *Realistic Panpsychism* in *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp151-158.
- Schrödinger, E. 1935.** *The Present Situation in Quantum Mechanics* in *Quantum Theory and Measurement*, 1935, J.A. Wheeler and W.H. Zurek edition, Princeton University Press, 1983.
- Seager, William 2006.** The 'Intrinsic Nature' Argument for Panpsychism in *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp129-146.
- Stapp, H. P. 2006.** *Commentary on Strawson's Target Article*. In *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp163-170.
- Strawson, Galen 2006a.** *Realistic Monism: Why Physicalism Entails Panpsychism*. In *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp3-32.
- Strawson, Galen 2006b.** *Panpsychism? Reply to Commentators with a Celebration of Descartes*. In *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp184-281.
- Wilson, Catherine 2006.** *Commentary on Galen Strawson* in *Journal of Consciousness Studies* Vol. 13, No. 10-11, pp177-184.