

The Real Combination Problem: Panpsychism, Micro-Subjects, and Emergence¹

0. Introduction

Panpsychists hold that all matter possesses mental properties of some form. The main motivation for panpsychism is the perceived failure of reductive physicalism—the project of explaining, *inter alia*, the instantiation of conscious experience solely in terms of conventionally understood physical matter and properties. Such physical matter and properties are presumed not to include mental properties at the basic level.² According to panpsychists, the moral of anti-physicalist arguments like the knowledge argument and the zombie argument is that from ingredients entirely lacking in consciousness, consciousness cannot derive.³ To adequately account for the instantiation of consciousness in beings like us, panpsychists aver, we must take mental properties as part of the fundamental furniture of the material world.⁴

I consider overall that the panpsychist explanatory project is well-motivated, and that panpsychists are in the right ballpark for a here-and-now solution to the mind/body problem.⁵ I won't attempt to motivate panpsychism with respect to physicalism, but take it that *if* a viable panpsychist theory could be described, it would represent at least an attractive option on the field of metaphysics of mind. My concern is with difficulties internal to panpsychism, and with where the intuitions driving panpsychism should propel one if those difficulties turn out to be insuperable. The discussion should be of interest to any theorists who still hope for an explanation of consciousness in terms of its material underpinnings:

¹ Thanks to Gregg Rosenberg in particular for helpful comments on this paper, as well as the attendees of the 'Panpsychism on the Reef' conference of July 2012 organised by David Chalmers, and an anonymous referee for this journal.

² Cf. Stoljar 2006.

³ For a recent example of this sort of argument see Strawson 2006. See also Nagel 1979, James 1890.

⁴ As Nagel and Strawson in particular make clear, if the phenomenal can only derive from the phenomenal, then ultimate entities (e.g. basic particles) must be phenomenal: since were phenomenality to first arise at any non-basic level, it would derive from the non-phenomenal.

⁵ As compared with a wait-and-see solution like Stoljar's (2006). Stoljar believes we should await the discovery of a non-mental intrinsic nature to matter which would make the instantiation of macro-consciousness intelligible.

the question at the center of that enterprise, and of this paper, is quite how rich the properties of micro-matter must be in order to account for macro-consciousness.

Panpsychism harbors an unresolved tension, the seriousness of which has yet to be fully appreciated. I capture this tension as a dilemma, and offer panpsychists advice on how to resolve it. The dilemma, briefly, is as follows. Panpsychists are committed to the perspicuous explanation of macro-mentality in terms of micro-mentality (in a sense of ‘perspicuous explanation’ to be elucidated). But panpsychists take the micro-material realm to feature not just mental properties, but also micro-subjects to whom these properties belong. Yet it is impossible to explain the constitution of a macro-subject (like one of us) in terms of the assembly of micro-subjects, for, I will argue, subjects cannot combine. Therefore the panpsychist explanatory project is derailed by the insistence that the world’s ultimate material constituents (ultimates) are subjects of experience. The panpsychist faces a choice of abandoning her explanatory project, or recanting the claim that the ultimates are subjects. This is the dilemma. I argue that the latter option is to be preferred. This needn’t constitute a *wholesale* abandonment of panpsychism, however, since panpsychists can maintain that the ultimates possess phenomenal qualities, despite not being subjects of those qualities. This proposal requires us to make sense of phenomenal qualities existing independently of experiencing subjects, a challenge I tackle in the penultimate section. The position eventually reached is a form of neutral monism, so another way to express the overall argument is to say that, keeping true to their philosophical motivations, panpsychists should really be neutral monists. This shift carries a further boon: as well as being a coherent position—something I argue undiluted panpsychism is not—neutral monism is somewhat closer to physicalistic respectability, since it does not assert that basic material ontology is rife with subjects of experience.⁶

⁶ The argument assumes a robust realism about human subjects—that there really are such subjects, and that they are not merely (in some sense) splinters of a greater, universe-sized subject. I take this premise to have empirical support—we experience ourselves as such subjects (cf. James 1912/2005:13-14). But I don’t expect others to find this evidence uncontroversial, and am therefore prepared to accord this realism the status of *mere* assumption. The assumption

The dialectic revolves around a comparison between panpsychism and emergentism, so I begin by outlining this comparison, before moving to a discussion of the combination problem, and then to an examination of the difficulty positing micro-subjects causes for panpsychism—a difficulty it cannot overcome, and which ultimately motivates theory-change.

1. Panpsychism and Emergentism

Panpsychism's metaphysical and explanatory aspirations can be clarified through comparison with emergentism. The general idea of 'emergence' concerns the appearance of a property at a higher level of being or organisation which is absent at lower levels. Emergentism takes two forms, strong and weak. Weak emergentism amounts to the idea that the novel property is unpredictable with respect to the lower-level base from which it arises. But this is taken to be strictly an epistemological limitation for those studying the properties of the lower-level base. Perhaps they don't possess every scrap of lower-level information relevant to the generation of the novel property. Or perhaps there is some deeper reason why occurrences of the property remain unpredictable.⁷ Weak emergentism stops short of claiming that emergents are *ontological* novelties: in fact they are taken to be of a piece with their bases, and in principle their instantiations are wholly deducible (i.e. for an ideal reasoner). Strong emergentism, by contrast, is the claim that some properties are genuinely ontologically novel with respect to their lower-level bases. Strong emergents are unpredictable in principle. Such properties also possess autonomous causal powers.⁸ For these reasons it is plausible that strong emergentism concerning mental properties implies property dualism.⁹

Panpsychists shun emergentism, especially of the strong variety. They might allow that there exists a contingent gap in our understanding of the generation of macro-mentality from its material underpinnings, but it has been panpsychism's

prevents the panpsychist's dilemma branching into a tetralemma: we can rule out subject monism or nihilism as options. Thanks to Luke Roelofs for discussion of this point.

⁷ Humans might not be intelligent enough to carry through the derivation, for example.

⁸ See e.g. Humphreys 1997.

⁹ See e.g. Brüntrup 1998.

raison d'être to preserve ontological continuity between the world's lower ontological levels and the conscious macro-mentality arising from these levels. With this commitment comes the doctrine that the generation of conscious mentality from material underpinnings is in principle *perspicuously explicable*. This idea can be captured by saying that given a suitably complete catalogue of the world's micro-material contents, it should be inconceivable¹⁰ that ours is a zombie world—one lacking higher-level phenomenal consciousness of the sort we actually enjoy. Given the conceivability of *physical* zombies, panpsychists deem physicalist ontology short on resources for capturing consciousness. The required 'extra' is to be supplied by a distribution of micro-phenomenal properties across the ultimates. The panpsychist seeks thereby to add just enough to the world's micro-ingredients, as conceived by physicalism, to logically guarantee the instantiation of macro-consciousness. The presence of macro-consciousness should *flow from* the properties of the ultimates plus details of their arrangement.

This commitment to 'smooth transition'—i.e. non-emergence—is on display in the main philosophical vertebrae comprising modern-day panpsychism: the writings of James, Nagel and Strawson. James, the father of modern panpsychism, found that:

The demand for continuity has, over large tracts of science, proved itself to possess true prophetic power. We ought therefore ourselves sincerely to try every possible mode of conceiving the dawn of consciousness so that it may *not* appear equivalent to the irruption into the universe of a new nature, non-existent until then.

And the explicit reasoning leading James to panpsychism went: '*If evolution is to work smoothly, consciousness in some shape must have been present at the very origin of things.*'¹¹

¹⁰ The usual restrictions apply: ideally conceivable for an idealised reasoner. See Chalmers 2002 for explanation of these ideas.

¹¹ 1890:148-149.

Nagel is largely responsible for re-igniting interest in panpsychism in recent decades, having been brave enough to examine a theory of mind that, as the twentieth century closed, was largely ignored—or else considered laughable. The key premise of Nagel’s argument for pansychism is his rejection of emergence:

there are no truly emergent properties of complex systems. All properties of complex systems that are not relations between it and something else derive from the properties of its constituents and their effects on each other when so combined.¹²

From this principle, plus the claims that a conscious being is constructible from mere matter, and that a physicalist construal of such matter (as fundamentally lacking mentality) cannot account for consciousness, Nagel infers that consciousness must be present in basic matter.

Strawson’s work brought panpsychism to the attention, and critical scrutiny, of mainstream philosophers of mind, also making a new generation aware of its appeal. His argument for panpsychism departs not far from Nagel’s formulation, but Strawson contributes a clear and forceful presentation, allied to a sustained defence of the non-emergence principle. This he expresses with characteristic pithiness:

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 [i.e. matter as construed by physicalism]¹³

In understanding the production of liquidity from non-liquid components we move, says Strawson, ‘in a small set of conceptually homogenous shape-size-mass-charge-number-position-motion-involving physics notions’¹⁴ These notions provide no conceptual bridge to consciousness—hence the

¹² 1979:182.

¹³ 2006:15.

¹⁴ 2006:13.

conceivability of physical zombies—so, for Strawson, nothing short of positing phenomenality as a property of the ultimates will suffice as an explanatory basis for macro-consciousness.

Panpsychism thus stands opposed to emergentism.¹⁵ It has been said that austere reductive physicalism is suited to those with a taste for desert landscapes. Still, from the panpsychist perspective such physicalism, along with dualism and emergentism, all seem committed to discomfitingly *abrupt* ontological inclines—spots where the material landscape of a sudden manifests mentality, having just previously shown not the least sign of doing so. On these views consciousness irrupts onto the scenery unexpectedly, in sharply jutting outcrops. Panpsychists might then be described as those theorists with a taste for only gently graded landscapes. They decline to scale the sheer slopes of sudden mentality.

2. Constitutive Panpsychism

The panpsychist posits mentality as a fundamental feature of matter from a deep commitment to the non-emergence principle. But how should one understand the relationship between this basic mentality and high-level consciousness? Given the desire for a smooth and intelligible transition from the micro-material to the macro-conscious—what Van Cleve dubs panpsychism’s ‘mereological rationalism’¹⁶—the natural way to construe the relationship between the mentality of the ultimates and the conscious mentality of high-level individuals is to say the former is *constitutive of* the latter.

The broad idea behind this *constitutive panpsychism*¹⁷ is as follows. Consider a macro-subject, one of us, having an experience of being cold, tired and smelling roast beef.¹⁸ This experience is *unified* in the following sense: though the subject can attend now to the sensation of cold, now to her tiredness, and now to the

¹⁵ Cf. Van Cleve (1990), who responds to Nagel by arguing that emergentism is preferable to panpsychism as an outcome of his argument.

¹⁶ 1990: 218.

¹⁷ The name is due to Chalmers (talk at Hochschule of Philosophy, Munich, June 2011).

¹⁸ The example is Goff’s (2009).

smell of roast beef, still, phenomenologically-speaking, these three sensations are given to her all in one go. Her phenomenological point of view appears suffused by all three together, as opposed to experiencing them only discretely, in series. It is the presence of such unified phenomenal fields as we each enjoy, featuring at a time many different sorts of phenomenal element, that any theory of consciousness must adequately explain. On constitutive panpsychism this unified experiential field is literally constituted by parts, ultimates in the subject's brain, instantiating just those qualities that feature in consciousness. So the macro-experience of feeling cold, tired and smelling roast beef is constituted by one set of ultimates instantiating phenomenal coldness, another set instantiating phenomenal tiredness, and a further set instantiating roast beef smell. These groups of ultimates together compose the consciousness-realising portion of the subject's brain.

This is an over-simplified model of what the constitutive panpsychist intends,¹⁹ but the guiding idea is clear enough. Any phenomenal qualities figuring in the macro-experience are put there through being carried by the phenomenally-qualified ultimates jointly composing the relevant part of the subject's brain. Thus we can trace the macro-phenomenology to the phenomenal states of the ultimates in a smooth way, and panpsychists can account for the macro-experiential state in terms of the matter composing it. This explanatory operation is essentially no more difficult than accounting for the overall composition of a painted canvas by reference to the various patches of paint filling it, along with their qualities. In the case of phenomenal paint, this is achieved by taking phenomenal qualities to belong to basic matter, and by taking this matter not only to materially constitute the subject, but to *phenomenally compose* her experiential field as well; that is to say, the phenomenal quality of each experiential field-composing ultimate finds its way into the conscious state enjoyed by the subject.

¹⁹ For one thing, there will have to be some sort of *qualitative blending* or *pooling* among the qualities carried by each ultimate: if each ultimate's quality showed up *as such* in the macro-experience, it would lack the notable homogeneity of (e.g.) color experience, and plausibly some mixing of basic qualities is required to obtain the qualities of macro-experience. I revisit these issues from section 4 onwards, arguing that they are not terribly problematic for panpsychist models of combination.

The experiential field, composed in this way by the relevant micro-phenomenal items, counts as what O'Connor and Wong term a 'structural property', where:

there is nothing more to having the structural property than being composed by parts having certain other properties and bearing certain relations to one another—it is ontologically reducible.²⁰

In our case, instantiating the macro-experience of coldness, tiredness and roast beef smell is nothing more than having a certain brain-state composed of ultimates some of which instantiate phenomenal coldness, some phenomenal tiredness, and some phenomenal roast beef smell.²¹ Constitutive panpsychism thus successfully evades emergentism, providing a satisfactory realisation story for macro-experientiality by its material underpinnings. Put in the terms employed earlier to capture the panpsychist's explanatory ambitions: it ought to be that given full specification of the ultimates composing the subject's experience, including details of the phenomenal quality each carries, it is inconceivable both that this micro-material setup obtains and that the subject does not instantiate an experiential state featuring coldness, tiredness and roast beef smell. Since the phenomenally-qualified ultimates in question (jointly) instantiate these very qualities, and constitute the subject's experience, this test appears to be passed. The absence of the subject's experience seems no more conceivable than the absence of a certain painting once we have stipulated that a set of paint patches is arranged thus-and-so, in these-and-those colors.

Constitutive panpsychism is thus faithful to the panpsychist's core concern: to posit enough in the material fundamentals for a smooth realisation story for macro-consciousness. Indeed, it seems the most natural way of meeting the panpsychist's explanatory requirements. So I will henceforth equate (anti-emergentist) panpsychism with constitutive panpsychism.

²⁰ 2005:10. See also Armstrong 1978. Note that the anti-emergence premise in Nagel's argument for panpsychism expresses the same idea. The panpsychist anti-emergence principle is effectively the claim that all non-basic properties are structural.

²¹ Wong and O'Connor's characterisation requires that the components not possess the higher-level property. This requirement is met in the present case: the higher-level property is of experiencing coldness *and* tiredness *and* roast beef smell, and no ultimate possesses this conjunctive property.

3. Panpsychism and Micro-Subjects

The panpsychist's ultimates possess phenomenal qualities. For example, the ultimates constituting the experience just considered possessed phenomenal qualities like phenomenal coldness, tiredness and roast beef smell. But phenomenal qualities are properties that characterise conscious experience of the sort *we* (macro-subjects) enjoy; they are those qualities such that *there is something it is like*²² to have them, something it is like, crucially, *for* the haver. But if they are like something for the haver, that haver would seem *a fortiori* to be a subject of experience—since a subject is just that sort of entity for whom anything can be like anything at all. This 'for', in other words, seems to drag ineluctably in its wake the notion of subjectivity, and its corollary, the notion of a phenomenological *point of view*. Something can be like something *for you* but not like anything *for me* precisely because we are distinct subjects occupying distinct points of view in (and upon) the world. So it would appear that phenomenal qualities necessarily exist as belonging to subjects. This, in turn, would convert the ultimates into subjects, since the panpsychist's ultimates possess phenomenal qualities. It's unsurprising, then, that most historical panpsychists agree that the ultimates are subjects of experience. As Goff says: 'If panpsychism is true then physical ultimates are subjects of experience...there is something that it's like to be a physical ultimate.'²³ Strawson is equally clear on this point.²⁴

For constitutive panpsychists, then, human experiencers are at bottom constituted by materially fundamental 'micro-subjects'—corresponding to basic physical particles—who are the first holders of the phenomenal qualities we enjoy atop our macroscopic perches.

We must now ask: What is the relationship between the *subjectivity* of these fundamental subjects and the unified, single-perspective subjectivity of the human subjects they compose? For simplicity we'll consider only the relationship between ultimate-subjects and human subjects, without worrying

²² This, and related phrases (e.g. 'what-it-is-likeness') commonly used to evoke phenomenal consciousness derive from Nagel 1974.

²³ Goff 2009:301.

²⁴ See his 2006.

whether the ultimate-subjects also compose subjects composing us, e.g. subjects corresponding to brain-hemispheres.

We grasped well enough, in outline, the constitution of a macro-qualitative state of *coldness, tiredness and roast beef smell* by adverting to composing ultimates which jointly instantiated just those qualities. It was something like the way we understand the composition of a painting by thinking about the various painted patches filling the canvas: we consider their qualities in isolation, and see how, by assembling the qualitative patches, we obtain the complete image. But, we will see, there is no way to understand, in parallel fashion, the assembly of subjects into a larger subject. Subjects do not combine. This leaves constitutive panpsychism in a pickle. To lay the ground for this difficulty, it's time we introduced the notorious *combination problem* for panpsychism.

4. The Combination Problem

Seager coined the name,²⁵ describing the combination problem as 'the problem of explaining how complex conscious states emerge from the primitive mental states ascribed [by panpsychists] to...fundamental entities'.²⁶ This specification, however, is ambiguous between two distinct difficulties in particular. I will argue that only the second, the *subject combination problem*, need seriously worry panpsychists. Unfortunately, this difficulty is also insuperable, and it is on this rock that traditional panpsychism runs aground.

The combination problem is originally due to James, who famously formulates it thus:

Take a hundred [feelings], shuffle them and pack them as close together as you can (whatever that might mean); still each remains the same feeling it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean. There would be a hundred-and-first feeling there, if, when a group or series of such feelings were set up, a consciousness belonging to the group as such should emerge. And this 101st feeling would be a totally new fact;

²⁵ Seager 1995.

²⁶ MS:6.

the 100 original feelings might, by a curious physical law, be a signal for its creation, when they came together; but they would have no substantial identity with it, nor it with them, and one could never deduce the one from the others, or (in any intelligible sense) say that they evolved it.²⁷

To isolate the two versions of the combination problem, we should examine what James means by 'feeling'. There are two possibilities, which are not usually distinguished. First, James may simply mean by 'feeling' something like 'qualitative element', a notion explained as follows. Consider your current overall phenomenal field—your complete experience, such that this comprises other things which could be called experiences (like the feeling of your bottom on your chair, or your visual experience of this page) and such that this experience is not itself a component of any further experience. It is a maximal property of you. Now isolate a single component of your complete experience: for convenience let it be the visual experience as of this page, a largely white field with black shapes across it. This visual experience of the white and black page is a 'qualitative element' of the overall qualitative phenomenal field you enjoy, across your various senses and faculties.

What is notable about this particular qualitative element, which goes for any other we care to describe, is that we specified its intrinsic character without any essential reference to the subject in fact enjoying it: you. We have no need, in the normal course of things, to mention the subject when enumerating the qualitative elements she enjoys. You are, for common sense anyhow, the only subject enjoying this token visual experience of this page. So we can specify the experience, this component of your complete experience, purely in qualitative terms—in terms of the qualities you experience in having it. These are, as we said, something like *a field of white and smaller black shapes running across it*. Thus considered in abstraction from—without mentioning—the fact of its being

²⁷ 1890:160.

enjoyed by a subject, the experience is a mere qualitative element of your overall experience. By ‘feeling’ James could have meant this sort of thing.²⁸

But when James speaks of the assembly of feelings producing ‘a consciousness belonging to the group as such...a 101st feeling’, it suggests a different sense of ‘feeling’, one more distant from everyday talk, though, one might think, a sense congenial to panpsychists. This is a use of ‘feeling’ which James apparently considers synonymous with ‘subject’, since he employs the terms ‘a consciousness’ and ‘a feeling’ interchangeably. As noted earlier, consciousness entails its being like something for someone, and that someone is a subject.

So, on one reading James’s ‘feelings’ are qualitative elements experienced by a subject, and on the other they are experiencing subjects themselves. The first is the more common understanding, since we typically talk about the feelings a subject enjoys—those qualitative elements ‘before’ her in consciousness—without bothering explicitly to invoke the presence of the subject herself. We simply specify these elements *via* the different qualities they exemplify. *Nota bene*: There is, as yet, no insinuation that qualitative elements could exist outside of being experienced by subjects. The point is only that they can be described without including subjectivity into the description.

Given these two readings of ‘feeling’, we now have two ways of understanding the problem James wants to highlight. Providing his use of ‘feeling’ is univocal throughout the quoted passage,²⁹ it seems James either claims that:

1. *Qualitative elements cannot be assembled so as to have a substantial identity with their product.*

or:

²⁸ I note that some (e.g. Kriegel 2009) believe there to be a distinctive ‘me-ness’ quality present in experience. If this is right there is at least one qualitative element that can’t be specified without mentioning the subject. I’m not acquainted with the ‘me-ness’ quality, however.

²⁹ Perhaps it isn’t. But then I am at a loss to reconstruct the difficulty James has in mind.

2. *Subjects of experience cannot be assembled so as to have a substantial identity with their product.*

I understand James's term 'substantial identity' as follows. James half-unpacks it, as part of his denial that the hundred feelings could form a substantial identity with the one-hundred-and-first, by saying the product could never be deduced from its antecedents. So James means by 'substantial identity' just the idea we have been operating with on the anti-emergentist panpsychist's behalf: it is a product that does not emerge from its ingredients ('emerge' in the bad sense of *strongly* emerge),³⁰ but which is perspicuously and smoothly derived from their arrangement: a structural property. This reading is reinforced by James's final phrase, where he parses a substantial identity as an item that *intelligibly evolves* from its antecedents. Claim 1, then, asserts that a complex qualitative field cannot intelligibly (non-emergently, structurally) be assembled from qualitative ingredients. Claim 2 asserts that a macro-subject cannot intelligibly (non-emergently, structurally) be assembled from micro-subjects.

Before assessing these claims, we can now see how Seager's talk of the 'problem of explaining how complex conscious states emerge from the primitive mental states', did not tell us whether the challenge concerned the 'emergence' (presumably Seager's is the innocuous sense of 'emerge from', equivalent to 'derive from') of complex qualitative fields from simpler qualitative ingredients, or the emergence of a higher-level subject from subject-components. The Jamesian ambiguity is also present, to a lesser degree, in Goff's work on the combination problem. Though Goff tends to avoid ambiguous formulations, he does fluctuate between describing the difficulty in quality-based and subject-based terms, even within the same article. Perhaps he is best understood as holding that both aspects of the combination problem are equally serious for panpsychists.³¹

³⁰ For, recall, weak emergents are in principle fully explicable.

³¹ Goff's 2006, for instance, employs the example of the combination of ten slight pains into a single severe pain. At times Goff talks as if the problem is understanding the qualitative combination (since ten slight pains are not the same thing as one severe pain) and elsewhere he

I do not agree that this is so, however—that both problems are equally serious. The reason is that there is simply no ‘combination problem’ as concerns the assembly of qualitative instances into a qualitative whole. There is little challenging in the notion of taking ingredients, each with a certain qualitative character, and putting them together into a whole whose macroscopic qualitative character is the intelligible product of the qualities of the components plus their arrangement. There are a few different relevant conceptual models we have available from everyday life, and it might be that phenomenal qualitative combination conforms to any one (or more) of these. Again, phenomenal qualitative combination may have its own rules; but there seems little reason to think there is anything much more conceptually challenging in such combination than there is in the everyday models of qualitative combination.³²

We considered already a merely additive model, on which the qualities of the composite are simply the organized set of those instantiated by its components. Our illustrative analogy for this kind of combination was a painting’s composition by qualitatively distinct, determinate paint patches. Perhaps phenomenal qualitative combination is like this. Of course, the additive model is simplistic, even as regards building a painting. In painting an important element, something the artist spends considerable time on, is the *blending* of qualitative instances. A red patch mixed with a dab of blue will give purple, wherein the antecedent qualitative instances survive in their contribution to the blend. Various colored rays of light blend differently (as common sense views things): here the blending of the various qualities gives white—in painting it gives black. Some such blending mechanism will surely be an important component of micro-quality combination into macro-qualia. In any of these cases just considered—additive, or blending, or a combination of both—a synchronous snapshot of the composite can be intelligibly related to its component qualitative elements,

seems to suggest that what blocks combination is the fact that each pain is associated with a distinct subject.

³² This is not to say the rules will be graspable *just* from consideration of uncombined qualitative ingredients. We had to study chemical reactions to understand the generation of the properties of NaCl from its antecedents, for example (at the time he writes, Broad 1925 considers chemical properties emergent). But after the fact we are able to see how the high-level properties are structurally produced. It may well be the same for macro-qualitative combination.

which can be readily conceived as persisting *via* the ongoing contribution they make to the new quality instantiated (take the blueness away from the purple paint-dab and you are left with red, altering the quality of the whole). There is, in short, no serious difficulty associated with the idea of a complex macroscopic qualitative state being the combination of a set of simpler quality instances. The first Jamesian combination problem, then, is no real problem at all. For homely confirmation of this, we can think about the way we easily understand the combination of the various flavours we taste in a meal into a gustatory *gestalt*.

What of the second putative problem, concerning the combination of subjects into a 'larger' subject? This is the real combination problem.

5. Subjects do not Combine

Goff argues that the existence of a set of subjects, however arranged, never *a priori* entails the existence of another subject, thus panpsychism's explanatory project falters.³³ Goff is correct, but the situation for panpsychists is far worse than he estimates, both epistemically and metaphysically. By Goff's argument we cannot see how the combination of micro-subjects could provide for the existence of a macro-subject. I argue that we can see how the combination of micro-subjects could not provide for the existence of a macro-subject. The combination of subjects is a demonstrably incoherent notion, not just one lacking in *a priori* intelligibility or evident necessity. Goff's epistemic argument leaves it open that subjects might combine as the panpsychist envisages; my metaphysical argument rules this out, thus ruling out constitutive panpsychism. Before proceeding with the argument, we need to say a bit more about the concept of combination, and how we understand subjects of experience.

Earlier we connected the notion of a *subject* with that of a point of view.³⁴ Neither notion is easy to get a grip on, but noting one or two basic features of each will help to make their interrelation clear enough for present purposes. The idea of being a subject goes with being an *experiential entity*, something

³³ Goff 2009.

³⁴ A conceptual connection salient in Nagel's work, for example the second half of his 1979.

conscious of phenomenal qualities. That a given subject has a particular phenomenological point of view can be taken as saying that there exists a discrete 'sphere' of conscious-experiential goings-on corresponding to this subject, with regard to which other subjects are distinct in respect of the phenomenal qualities they experience, and they have no direct (i.e. experiential) access to the qualitative field enjoyed by the first subject. A subject, then, can be thought of as a point of view annexed to a private qualitative field (this is a conceptual bifurcation for now—these aspects may be metaphysically inextricable). A thought experiment will exhibit the interrelation between a subject's point of view and those qualities within her experiential ambit. Imagine a hundred qualitatively identical subjects at the 'starting line' of existence—their only difference is that they occupy distinct positions in space-time. They are about to set out on their lives. As time winds on, each takes a unique path through the environment, and is impinged upon differently. These different impingings result in different modifications of each sensory field. Thus each subjectual perspective has access to a qualitatively different array of qualia, as compared with other subjects, over its lifetime. It is the fact of these different points of view, these differently located 'lookouts' on the world, that then grounds the character of the peculiar set of qualities each subject experiences at a given time.³⁵

Our template for *combination* is the assembly of two hydrogen and one oxygen atom into a molecule of water. First we'll consider the combination in *particular*-terms, then in *property*-terms. These two modes function in much the same way, the reason being that they operate in parallel: combination of particulars occurs thanks to the integration of their properties, and combination of properties occurs as the bearers of the properties are combined. When hydrogen and oxygen atoms combine into water they bond covalently, sharing electrons. The oxygen atom completes its outer shell by borrowing an electron from each hydrogen. Thus the three atoms are *deformed*, intrinsically modified, by participating in the combination of water. Yet, importantly, all three atoms

³⁵ Cf. Russell's 'cameras' (1927b:122). What must also help fix the range of qualities a subject is capable of is the subject's own physical constitution.

continue to exist once combination is achieved. Their union modifies them, but they survive in the whole, making an ongoing contribution to its nature. This is clear enough, for were we to remove one of the hydrogens (say) the result would be destruction of the water molecule. Combination, thus, is the formation of a whole from components where the components continue to exist in the whole, but are intrinsically altered by combining with one another. This is as against a mere *aggregation* of items, which need not condition or unite those items in any way. Other intuitive examples are the combination of qualitatively and spatially distinct paint patches into a painting, and the combination of cooking ingredients into a meal. The survival of the components is entailed by saying the product is their combination or union, as opposed to being merely their effect or descendant.

Turning now to properties: a combination or unity possesses novel *systemic powers*. A water molecule forms a dipole: its charge is 'two-faced', being negative towards the oxygen side and positive towards the hydrogen side. This feature accounts for the way water molecules bond, and water's boiling point. An aggregate's powers, by contrast, are a mere jumble of the powers of its parts. What is the relationship of a unity's new systemic power to the powers of its isolated, pre-combination parts? The distinctive power of the combination is novel, in that none of its parts possesses it. It is not however an *autonomous* higher-level power, in the sense of being inexplicable with reference to the powers of the parts. Such autonomous higher-level powers are characteristic of emergents. The systemic powers of unities are, rather, structural properties. The dipolar property of a water molecule is intelligible as the *interactive organization* of the charges of its constituent atoms: these charges are arranged so they interact (thus altering their bearers intrinsically—electrons become prised away and shared), and the upshot of their interaction, their combination as charges, is the dipolarity of the molecule. The electrons loaned to the molecule by the hydrogens spend most time around the oxygen atom, giving this side a relatively negative charge, and the hydrogen side a relatively positive charge. As with particular-combination, it is crucial that the pre-combination ingredients— here properties—survive in the unity: the charges of the pre-combination atoms,

though modified by combining, survive in the ongoing contribution they make to the polarity of the molecule. If any of those charges ceased to exist, the molecule could not have the dipolarity it actually manifests.

If micro-subjects combined into macro-subjects, then, what we should anticipate is their coming together closely enough to interact, thereby modifying their properties, whilst all the while they survived in the whole they formed. Their persistence in the whole would amount to an ongoing contribution, in respect of their natures, to the product, as with atoms in a water molecule. The constitutive panpsychist posits subjectivity—which we analysed as the having of a point of view to which qualities are present—as a fundamental feature of matter, the model being other fundamental properties like mass and charge. Now, it is important to note that basic mass and charge are directly relevant to their higher-level instantiations. The mass of a helium atomic nucleus is the mass of its two protons and two neutrons, minus the energy used to bond them together (a good example of combination). We have already seen that a H₂O's dipolarity is the product of the charges of its atomic parts; their charges in turn are the product of those pertaining to their subatomic constituents, and so on down to the fundamental level. The same pattern goes for instantiations of mass and charge at higher levels of being: both are ultimately the result of mass/charge interactions at lower levels, down to the ultimates. In a way the very reason for the positing of fundamental mass and charge has been to account for higher-level instantiations of these properties, those directly observable by us. Similarly, with subjectivity a high-level feature manifest to us, the constitutive panpsychist posits fundamental subjectivity to account for these high-level instances. For this to work, fundamental instances of subjectivity—ultimate-subjects—would have to 'add up to' bigger subjects, as fundamental charges and masses produce higher-level instances of these properties.

But this cannot work, simply because points of view cannot combine. Consider two micro-subjects, eager for combination into a higher-level individual. How might they combine? Since this is combination, we want the ingredients to survive in the whole. Just as atoms are not obliterated, only deformed, in

building a molecule, subjects would have to persist *as such* within their higher-level product. This means we can rule out immediately a scenario where after integration just a single subject exists. If one subject is left where formerly we had two, this means at least one subject has gone out of existence, which is not combination but a fight to the death. It doesn't improve matters if we try to imagine the two original subjects as somehow each contributing a 'quantity of point-of-view-ness' to the surviving subject.³⁶ For the two to survive in the whole *as subjects*, there would have to persist (at least) two points of view. Yet by hypothesis we have a single subjective point of view left over. Thus one point of view, and corresponding subject, has wholly ceased to exist. There's really no such thing as a 'quantity' of point-of-view-ness, or perspective. Points of view are binary entities: they exist either wholly or not at all. Talk of subjects contributing *quantities of point-of-view-ness* to other subjects could really only mean the contribution of experiential *contents*: Perhaps by telepathy you could add to my stock of experienced qualitative elements, by giving me access to your chocolate-qualia. But this isn't to add any of your *perspective* to mine, only to add a bit of what yours is a perspective *on*. Let's stipulate that the singleton successor to our micro-subject pair requires the donation of one of its predecessors' points of view, so there's a point of view around after integration. But with this contribution made, there's nothing for the other original micro-subject to add to the successor in respect of its having a point of view, on this scenario where we end up with a single unified perspective. All that seems required of her, *qua* point of view, is her quiet disappearance. What could this point of view now add to her successor except her subjective *diversity*, which would precisely be disruptive of its one-ness? Thus if one point of view remains at least one point of view has been eliminated,³⁷ which is not combination.³⁸

³⁶ Thanks to a reviewer for this suggestion.

³⁷ Whether both of the original points of view are eliminated depends on whether we construe the survivor as identical to one of the originals, or as a descendant inheriting a point of view. But since points of view correspond to subjects, it seems that the survivor *will* be identical with one of the originals: whosoever donates the surviving point of view.

³⁸ The proposal that some subjects are eliminated in combination might not seem so terrible to panpsychists, even if it doesn't qualify as 'combination' under our rubric. In the next section I explain why this option (unacceptably) implies emergentism.

The two original points of view must survive in combination. But we can also rule out a scenario where precisely their two points of view remain. What we wanted was to assemble our two subjects so as to constitute a *unified* higher-level subject, with its own point of view. If we are left with all and only the original pair of points of view then we still have a multitude, and are nowhere nearer to the genuine combination of subjectivities into one subject. This would be so even were both subjects for some reason to experience just the same range of qualities for a time. There would still be two distinct token *experiencings of*—perspectives on—these qualities.

The serious constitutive panpsychist proposal, therefore, is this: integration of our two subjects must produce a third subject. Might this represent some sort of overarching ‘über-subject’, comprising as constituents the two antecedent points of view, which survive in the whole? As regards experiential *contents*, we can imagine the über-subject’s experiential field to be the *qualitative product* of the experiential contents of the originals, according to our best model of qualitative combination. If one antecedent subject experiences a unitary phenomenal blueness, the other redness, the über-subject experiences both colors, perhaps blended into a purple block, or as stripes.

Yet we are still no closer to combining *points of view*, even if we have made some ground in integrating experiential contents. Under this scenario we have three points of view in play—those of the original duo, plus that of the new subject. Is it possible to understand the original two points of view as *components* of the third? It is not. Here are two ways to see this, both of which bear out the essential discreteness of subject-perspectives:

Consider the original duo’s points of view. One—*Blue’s*—is pervaded by a unitary blueness, the other—*Red’s*—by redness, and that is all they experience, respectively. To say these points of view were present as components in the experiential perspective of the über-subject (*‘Ub’*) would therefore be to say that *Ub* experienced a unitary phenomenal blueness and a unitary phenomenal redness, i.e. had synchronous experiences as of each of these qualities alone, *to*

the exclusion of all others. For it is these qualities each *on their own* that characterise, respectively, the perspectives of the original duo. Experience excludes, as well as includes. Yet nowhere does *Ub* have any such experiences: he precisely *combines* his predecessors' qualitative experiential contents. *Ub* doesn't experience red-to-the-exclusion-of-(blue-and)-all-else, nor blue-to-the-exclusion-of-(red-and)-all-else, let alone—impossibly—both *together*. Thus the original points of view are not ingredients in *Ub*'s subjectivity. Only their contents—the redness and blueness—are. There are some deceptive conceptual possibilities in the neighborhood: We might imagine, first, *Ub* with 'stereoscopic' vision—two visual fields, one filled with blue, the other with red. But that is one point of view experiencing blue and red conjunctively, nothing like the original two points of view. If these visual fields are genuinely *experientially* separated, on the other hand, i.e. not given to the same consciousness, then we just have two separate subjects, much as if *Ub* never came to be. We might also imagine *Ub* experiencing all and only blue, then all and only red, in series. But that, while it might (perhaps) count as occupying now *Red*'s point of view, now *Blue*'s, is not to have their two points of view synchronously compose *Ub*'s point of view, which is what combination requires.³⁹

Another way to see that the original points of view don't compose the third is to imagine subtracting one of them. Recall that in a combination removal of a component compromises the whole—in removing a hydrogen from H₂O we destroy the molecule. Now imagine *Red* disappears. Will *Ub* notice? It seems he need not: provided *Red*'s experiential *content*—the relevant portion of phenomenal redness which *Ub* is experiencing—remains, what it is like for *Ub* won't change. Thus *Red*'s point of view did not phenomenally compose *Ub*'s. Someone might object that if *Red* disappears, so does her experiential content, so *Ub* must lose the redness from his experience—his purple experience will

³⁹ It's true that in describing combination we noted that constituents are modified in combining, so it may be unduly artificial here to consider *Red* and *Blue*'s pre-combination perspectives in unmodified form (i.e. as filled with unitary red and unitary blue, respectively). Perhaps the right picture has *Red* and *Blue* interacting to produce *Ub*, with *all three* now experiencing purple. But this does nothing to get around the problem the artificial setup serves to highlight: if the subjects survive, as they must, there will be three actual experiencings of purple. It's not possible to imagine *Ub*'s unified point of view *as* comprising two synchronous experiencings of purple (pertaining to the perspectives of *Red* and *Blue*).

devolve to blue. But that would at best prove *Red's* experiential content was integrated in *Ub's*. It does nothing to show that *Red's very perspective* was a component of *Ub's* perspective. Indeed, if *Red's* point of view did partially compose *Ub's*, then *Red's* disappearance should make more of a phenomenal difference to *Ub* than the mere loss of the red content. For *that* loss is achievable just by *Red's* reaching a 'zen state' of pure 'contentless' experience, without *Red* as a point of view departing the scene. We might additionally stipulate that *Ub's* existence depends on *Red's* (plus *Blue's*), so that if *Red's* point of view disappears *Ub* dies off. But that's as close as we can get—dependence is not constitution. No sense is to be made of points of view literally composing, *as* points of view, another point of view.

Whatever the relationship between *Ub's* experience and those of the original duo as regards content, it seems *Ub* as the bearer of a point of view can never be said to be the combination of the two micro-subjects, even if he is after some manner their product (by James's 'curious physical law'). While the *qualitative contents* of consciousness may perhaps combine, consciousnesses themselves—subjects—cannot: this is precluded by the metaphysical logic of points of view.⁴⁰

6. From Constitutive Panpsychism to Emergentism

Must micro-subjects survive in the higher-level subject? Perhaps it wouldn't be 'combination', as defined, but what would be wrong with the antecedent subjects generating the über-subject and in the process going out of existence? Then their various points of view wouldn't be around to stick out like such sore thumbs for the single-perspective macro-subjectivity we aim to construct.

Panpsychists inclined to reply thus have something like Humphreys' *fusion* in mind.⁴¹ Humphreys offers fusion as a model of the generation of higher-level items by sets of lower-level items. The lower-level items 'fuse' together, producing the higher-level item, and in the process are obliterated. All that

⁴⁰ Nagel apparently spies the problem: 'How could a single self be composed of many selves?' (1979:194). The preceding section, and the next, seek to bring Nagel's difficulty into focus.

⁴¹ Humphreys 1997.

remains in existence is the higher-level product. In this vein, panpsychists might propose that lower-level subjects fuse to form one higher-level subject, a subject like one of us. After fusion, what remains is the single higher-level subject, with the lower-level subjects of the 'fusion-base' having ceased to exist.⁴² This seems what Seager intends with his concept of panpsychic 'combinatorial fusion',⁴³ and Goff briefly entertains the same idea:

Perhaps the parts of my brain, before they came together to form my brain, had their own individual phenomenal lives. But when they...form my brain they lose their individual conscious identities, and somehow morph into o-experience had by the whole brain.⁴⁴

But fusion cannot help the panpsychist. The problem lies in something already emphasised: the essentially isolated nature of individual subjects with their independent experiential points of view. This feature of subjects ensures their fusion could only provide for the *emergence* of a macro-subject, implying a violation of panpsychism's core metaphysical commitments. To see why this is, we must further examine the constitution of structural (i.e. non-emergent) properties.

Consider as a possible case of fusion the melting of several gold ingots into one big gold lump. The masses of the ingots fuse: Each ingot has its own mass before melting, but having been formed into one chunk of gold, these individual masses no longer exist. Only the mass of the whole exists, formed by the fusion of the antecedent ingots and their respective masses.⁴⁵ What is important to note is this. In general, to avoid emergence, lower-level properties must contribute to their novel product *in virtue of their metaphysical nature*, or, otherwise put, *while*

⁴² Humphreys allows fusion of properties and fusion of things.

⁴³ 'Perhaps we can regard the parts as infusing their properties into the whole and by so doing effacing themselves.' (MS:14).

⁴⁴ 2009:308.

⁴⁵ Someone might dispute this as an example of fusion—it might be thought the mass of the chunk is a straightforward combination: perhaps the product of the (persisting) masses of the ingots, or the masses of constituent gold molecules, which certainly do not cease to exist. Not much hangs on the choice of case, however. The example is only illustrative of the deeper point developed below: that for a structural (i.e. non-emergent) property to result, whether by fusion or otherwise, there needs to be an *intelligible contribution* by the antecedent properties to the product.

remaining true to what they are. In the formation of a structural property, its lower-level antecedent properties, belonging to multiple lower-level bearers, interact with one another due to their character, and the upshot of this self-expressive interaction is the higher-level property. There is thus a requirement of metaphysical *sense-making* in the production of a structural property. Antecedents generate the novel structure by interacting according to, and abiding by, what they *already are*—so that there is a corresponding lack of surprise in the product: it is intelligible. If we look back to the anti-emergence principle as expressed by Nagel and Strawson, in particular, we can see both gesturing towards this idea.⁴⁶

Fusing masses meet this constraint: the masses simply add up, they *amass*, to form the fused mass. The lower-level masses express their nature in producing the larger mass; amassing is a characteristic thing masses do. Other intuitive examples of structural properties conform. We considered earlier the combination of atomic charges into a dipole in water molecules. When charges interact to yield some resultant, group charge, they do so, affect one another, *as charges*. Nothing blocks or defies the influence of each component in respect of *being charged*—each lower-level charge is at liberty to contribute its nature to the atomic interaction, and the resultant molecular dipole is simply the structural upshot of this free, self-expressive interaction. Another example is shape: one can see that in building a Lego penguin, say, the overall shape has the relation of the shapes of the blocks to one another to thank for its determinate nature: their relation *in virtue of their shape*. When colors blend to produce a new color, as with mixing red and blue into purple, it is a colorful interaction with a colorful product.

Contrastingly, paradigm cases of emergence involve the metaphysical defiance or *thwarting* of the natures of the lower-level properties. Consider two examples Strawson employs: the emergence of the concrete from the abstract, and of

⁴⁶ Nagel, recall, says 'All [non-emergent] properties of complex systems...derive from the properties of its constituents and *their effects on each other when so combined*' (1979:182, my emphasis). These interactions between constituents of the system are expressions of their metaphysical nature, i.e. of their characteristic powers.

extension from the unextended. These are cases where the underlying properties cannot generate their product structurally, because inputting *what they are*—abstraction, non-extendedness—makes no contribution towards what results. No amount of interaction or accumulation of different *loci* of abstraction can *by itself* yield concreteness, and likewise for non-extendedness. Given what they are, abstraction and non-extendedness just don't have *what it takes* to produce concreteness and extension. Quite the reverse: the product positively defies the nature of the antecedents. Thus concreteness and extension must emerge, in these cases. The 'bruteness' philosophers associate with emergence is precisely this metaphysical defiance of the natures of the properties in the emergence base. If the antecedents' natures don't constrain their product, *anything* is possible, so the actual product is bound to appear arbitrary.

We can support this account by defusing a counterexample: *liquidity* is emergent by these lights, it might be said, since (following Strawson) it derives from non-liquid components. Surely the 'expression of their nature'—non-liquidity—cannot structurally account for the production of liquidity? But it's important to see that liquidity is a mere matter of the *dynamics* between a liquid's parts: molecules are sufficiently loosely bonded that they can slide over one another, in effect. Thus liquidity is really a macro-dynamic structuring of micro-dynamical components: its generation is a matter of dynamics all around. As Strawson says of such cases, when considering the micro-ingredients and their macro-product we move in 'a small set of conceptually homogenous shape-size-mass-charge-number-position-motion-involving physics notions'⁴⁷ Thus liquidity from the non-liquid is not emergence, and our account explains why.

Subjects being *essentially* discrete, pre-existing subjects can make no intelligible contribution to the unified subjectivity of a subject they generate *in respect of their* (the predecessors') *subjectivities*. The end subject's unity is precisely in the most direct defiance of the diversity and mutual ontological independence of the antecedents, which is what their being a set of subjects (at least partly) consists in. To create an über-subject from them, each predecessor's essential

⁴⁷ 2006:13.

subjectivity—hence, discreteness—is precisely the characteristic to be suppressed. To put the thought more revealingly: a set of points of view have nothing to contribute *as such* to a single, unified successor point of view. Their essential property defines them against it: in so far as they are points of view they are experientially distinct and isolated—they have different streams of consciousness. The diversity of the subject-set, of course, derives from the essential oneness of any given member: since each subject is essentially a oneness, a set of subjects are essentially diverse, for they must be a *set of onenesses*. Essential unity from essential diversity—what would be involved in panpsychic fusion—is thus a case of emergence, by our ‘defiance criterion’. ‘Consciousness fusion’ is guaranteed to be emergence in a way that mass (etc.) fusion is not: Masses can bequeath their massiness to their fusion. But subjects cannot bequeath their subjectivity: it is the destruction of their property of each having a point of view that is needed.⁴⁸

If points of view are annihilated so as to produce a macro-subject, the macro-subjectivity is not a structural property.⁴⁹ Instead, the only way to construe what happens is to say the micro-subjects brutally *cause* the successor to come into being, thereby expending their own claim to existence. This sort of causation of the new entity, as opposed to its formation by antecedent elements, is a hallmark of emergence for O’Connor and Wong:

[emergents are] those properties whose instantiation does not even partly consist in the instantiation of distinct properties by the entity

⁴⁸ With *liquidity* a ‘deeper’ (dynamic) property base explains why the liquid can be the structuring of the non-liquid. Similarly, when ocean waves meet to form a bigger wave, it is not their *being waves* that contributes, but their being composed of units that can be recombined into a single wave. The smaller waves are destroyed, *qua* waves, and their components made into the larger wave. Thus it is the properties of these lower-level components that are operative in the formation of the big wave (it is not just from *waves* of these components that you could generate the big wave). This ‘bypassing’ option is not available in the case of subjectivity, since the panpsychist posits subjectivity as a fundamental property, and fundamental properties are directly involved in producing their higher-level instances. After all, the reason the panpsychist posited fundamental subjectivity was to account for its higher-level instantiations. If a deeper, non-subjective, nature accounted for macro-subjectivity, this would make the panpsychist’s micro-subjects explanatorily—hence ontologically—superfluous.

⁴⁹ With *respect to* the micro-subjectivities: it might be structural with respect to some other lower-level property—but that would leave the micro-subjects explanatorily otiose, see previous note.

or its parts [i.e. structural properties]...[emergents] must be explained in terms of a causal, not purely formal, relationship to underlying, immediately preceding structures.⁵⁰

Panpsychist fusion clearly fits this rubric. Therefore, by insisting on micro-subjects the panpsychist slides from constitutive panpsychism into emergentism. Recap of the last two sections: To avoid emergentism the panpsychist must render macro-subjectivity as a structural property of an assembly of ultimates. But if ultimates are subjects they cannot combine in respect of subjectivity. If, alternatively, they fuse, their product is not structurally related to its antecedents. The resulting macro-subject can thus be at best only causally related to its micro-subject base. Hence the macro-subject emerges.⁵¹

Yet if panpsychists find themselves endorsing emergence at this juncture—to account for the production of high-level subjects—it's hard to see why they fought so shy of it earlier. Recall that the struggle to avoid emergence was at the heart of panpsychism, which sought a smooth ontological transition from micro-matter to macro-consciousness. Since panpsychism is driven by an aversion to emergentism, the result we face is that a panpsychism which posits ultimate-subjects is structurally unsound—it breaks its own anti-emergentist pledge.

To the extent one considers panpsychism committed to micro-subjects, one will conclude that the view is doomed: it suffers from an internal self-contradiction.⁵² And to the panpsychist who concedes that macro-subjects indeed emerge somehow out of micro-subjects, but who has no better reason for postulating micro-consciousness than to avoid its emergence in us, we can justifiably say

⁵⁰ 2001:10-11. This isn't to say that causal generation of properties is a necessary condition for emergence, only a sufficient condition. See Wong 2006 for helpful classification of varieties of emergentism. Another point to note is that the causal production envisaged in this formulation of emergentism may well be synchronic causation, which perhaps fits better with traditional construals of emergence.

⁵¹ There was perhaps a cheaper way of getting to this conclusion, *via* the fact that Humphreys explicitly intends fusion to be a variety of emergence. What this cheaper way would not have purchased, however, was detailed explanation of what was wrong with the proposal for panpsychists in particular.

⁵² The contradiction is performative or methodological rather than strictly logical: in attempting to evade emergentism the panpsychist who posits micro-subjects flees into the arms of another emergentism.

that the emergence of macro-phenomenality from non-phenomenal physical components (classic physicalism) cannot now consistently be dismissed as a possibility. As Seager observes ‘if panpsychism...requires a mechanism of emergence then why not take the theoretically more economical route of letting consciousness emerge directly from the physical basis itself rather than from a mental basis.’⁵³ Indeed.^{54,55}

7. The Neutral Monist Alternative

Panpsychism’s internal tension is encapsulable as follows. Panpsychists hold, effectively, that all non-fundamental properties are structural: they are reducible to more basic properties plus the arrangement of their bearers. This is the non-emergence principle. It drives theorists to panpsychism since they consider that, short of bearing phenomenal properties, physical micro-matter is an inadequate basis for consciousness. If physical micro-matter lacked phenomenal properties, macro-phenomenal properties could not be structural—they would, unacceptably, have to emerge. Yet panpsychists also insist that phenomenal properties—qualia—necessarily exist *for* someone, a subject who experiences

⁵³ MS:4.

⁵⁴ There are panpsychists whose main motivation is not aversion to emergence. Some are moved by Russell’s idea that physics doesn’t describe the intrinsic natures of its theoretical postulates, and that phenomenal properties provide a parsimonious way of filling this gap (e.g. Rosenberg 2004). Such theorists might evade the thrust of my argument; however it is clear that the ‘intrinsic nature argument’ (see Seager 2006) is highly controversial—far more so than the mundane thought that conventional physicalism lacks the conceptual resources to account for consciousness. Moreover the explanatory (anti-emergentist) motive has been by far the main driver for panpsychism historically, so most panpsychists are vulnerable to the present argument.

⁵⁵ Someone might claim that since panpsychist emergence at least occurs *within a class* of properties, i.e. the emergence of (unified) subjectivity from (plural) subjectivity, such emergence is in better shape than ‘physicalist emergence’—and not the unacceptable option I deem it. But the relevant contrast is between *structural* and *non-structural* generation of a property. Macro-subjectivity is not a structural property with respect to the postulated micro-subject base. ‘Structurality’—or non-emergence—is an all or nothing affair; once we say a property is non-structural, it doesn’t matter what it happens to have in common with its base (it will also have ‘being spatio-temporal’ in common, if we are non-dualists; in the physicalist case both properties count as *cognitive*, on the assumption that phenomenal consciousness is cognitive). The point is that the new property is not derived from its base; its instantiation is more than a matter of the organization of the basal items. Panpsychist and physicalist versions of emergence equally violate the structurality requirement, and once the panpsychist relinquishes this she has no solid ground upon which to dismiss the physicalist version, which, lest it be forgotten, is ontologically cheaper in not positing all-pervading micro-subjects. To this soft reply might be added a harder one: talk of ‘subjectivity’ at higher and lower levels disguises the deeper contrast—what really is required to emerge is *essential unity* from *essential diversity*, and this shatters the illusion of ‘intra-property emergence’.

them. Hence the ultimates are held to be subjects. But a macro-subject cannot be a structural entity with respect to a lower-level basis of micro-subjects, as we've seen. It transpires, therefore, that the panpsychist's commitment to micro-subjects is at war with her commitment to all higher-level properties—particularly consciousness—being structural. Something has to give.

From our historical survey, we can take it that the panpsychist's anti-emergentism (her 'mereological rationalism', in Van Cleve's memorable phrase) runs deep. Therefore what must go is the proliferation of subjects in micro-ontology. I now explore the sort of doctrine that results from this concession. What I sketch below is something of a roadmap, indicating the future direction of travel for those motivated by panpsychism's anti-emergentist impulse. Working out the details of the theory is something for future research.

Panpsychists are plausibly correct in holding that qualia—the qualitative properties we find in experience—are irreducible, and so (given the structurality of higher-level properties) fundamental. Where they err is in attaching subjectivity essentially to qualia: the notion that there must exist 'someone' to experience any given quale. Making the ultimates little subjects blocks them from being able to constitute a macro-subject. The moral of this failure is that qualia must be divorced from subjectivity—the awareness of qualia by subjects. Macro-qualia considered merely qualitatively, like red patches in the visual field, can be rendered as structural properties with respect to qualitative micro-underpinnings. But since subjects cannot combine into larger subjects, the only way to preserve the panpsychist anti-emergence principle when it comes to high-level subjecthood is to allow that, while quality is a fundamental affair, subjectivity must be susceptible of a reductive treatment.

A position taking qualities as fundamental features of matter, but which makes subjective awareness of qualities a relational (thus reducible) affair, is *neutral monism*, of the sort James later inclines towards.⁵⁶ On neutral monism, unexperienced qualities permeate basic matter. Certain portions of matter

⁵⁶ James 1912/2003.

exhibit a configuration which provides for awareness of the qualities they bear: matter, when specially arranged, can apprehend its own quality, in effect. This is consciousness. I now take these ideas in turn, explicating first a conception of qualia whereby they can exist unexperienced, then a relational conception of consciousness. Overall, the proposed movement of thought takes us some way back from panpsychism towards physicalism, so it won't be surprising if physicalists find the end position significantly less off-putting than panpsychism. The difficulty is persuading panpsychists that neutral monism is the way to get what they want most, as well as most of what they want.

When discussing Jamesian 'feelings', we encountered the conceptual possibility of qualitative elements of experience existing apart from subjects. This is a conceptual possibility to which we must now extend metaphysical seriousness. Of course, the notion of unexperienced qualia is by no means anathema to mainstream physicalism. When a normal-sighted person sees a blue sky, there is a blueness present in her visual experience: one way or another, and regardless for now of its location, there seems to exist a blue expanse of which the subject is aware. This blueness is a quale instantiation. Now, direct realists who are also color realists explicitly externalise our qualia, holding that the blueness apprehended in seeing the sky belongs to the sky and not to the mind of the seer.⁵⁷ These theorists don't consider the blueness in question to evaporate once nobody is looking at the sky, for they are decidedly *not* Berkeleians. This commits them to the thesis that qualia can endure unexperienced. That is just the proposal we need to make sense of here, although panpsychists are more likely to consider the qualia we are aware of as inhering in the matter of the brain. Panpsychists tend to be indirect realists about perception, holding that one is perceptually aware of external items by first being aware of internally instantiated qualia.⁵⁸

⁵⁷ See e.g. Tye 2009.

⁵⁸ A neutral monist, like a panpsychist, doesn't have to be an indirect realist about perception; nevertheless, these views seem naturally to go with indirect realism, given their emphasis on the reality and irreducibility of qualia.

Neutral monism holds that all matter possesses qualia, which need not be experienced in order to exist. Just as the sky's blueness or a fire-engine's redness persist when no-one is around to see them, on a commonsense or naïve realist perspective, so the qualities of the ultimates persist whether or not any subject is aware of them. A particular ultimate, then, might be phenomenally blue in the way a direct realist (color realist) believes the sky is blue.⁵⁹ On neutral monism when one is aware of a phenomenal blueness, one is directly aware of the qualities carried by ultimates in one's brain.

Strawson represents the panpsychist mainstream in holding that 'There cannot be an experience without a subject [to experience it].'⁶⁰ But the neutral monist thesis that the qualities exhibited in experience can exist outside of experience is not the oxymoronic claim that any *experience* can exist unexperienced. It is rather the claim that some ingredients of an experience need not be experienced. Rejection of this thesis would convert direct realists into idealists, so it cannot be completely objectionable. If Strawson's claim, alternatively, is that such qualities are *ipso facto* experiential, then that's simply the claim neutral monists deny, as well as the claim blocking panpsychist explanatory ambitions, since it entails micro-subjects. To my knowledge no panpsychist has provided a serious argument for the claim that qualia cannot exist unexperienced. That claim remains, therefore, in the realm of dogma or mere intuition. Opposing this intuition is common sense, where colors exist without perceivers, and pains sometimes wake us in the night. If pain qualia can cause one to come to consciousness of them, then clearly the qualia existed before one was conscious of them. So much for my attempt gently to accommodate panpsychists to the idea of subjectless qualia. The plain fact is that retaining the thesis that higher-level properties are structural makes subjectless qualia compulsory, as long as qualia are also held to be irreducible.

⁵⁹ Cf. the particulate imaginings of Unger 2005. There are theorists for whom the subatomic entities postulated by physics stand in need of a categorical nature, to underpin the relations physics catalogues with its equations (see Russell 1927a for an influential account of this idea). On neutral monism, this nature is constituted by (subjectless) qualia.

⁶⁰ 2008:152.

The neutral monist considers that qualities don't require subjects to experience them—the whole material world is en-qualified. What of those rare spots where subjects do exist who are aware of the qualia carried by matter, subjects like us? Panpsychists require macro-subjectivity to be a structural property, and neutral monism duly obliges, by rendering conscious awareness as a relation. James argues that:

Consciousness connotes a kind of external relation...not...a special stuff or way of being...The peculiarity of our experiences, that they not only are, but are known, which their “conscious” quality is invoked to explain, is better explained by their relations.⁶¹

For neutral monists, rather than a magical bubble or intrinsic phenomenal glow, consciousness is simply that relation whereby qualities are brought into subjectival awareness: it is the curling of the material world back upon itself, to apprehend its own character. What is the relation in question? This is an area ripe for innovation, for those sympathetic to the present argument. James is vague about the relation he envisages: he explains that qualia instances ('pure experiences') can exist both in subject and object 'position', as it were. They can be either the objects of 'knowledge', by other qualia, or the knowers. Such knowledge is the consciousness-constituting relationship. James also talks suggestively of *representation* playing a role:

As “subjective” we say that the experience represents; as “objective” it is represented. What represents and what is represented is here numerically the same...Its subjectivity and objectivity are functional attributes solely, realized only when the experience is “taken”, i.e., talked-of, twice...by a new retrospective experience (1912/2003:12)

This is difficult. If we decamp to neutral monism, we'll need to develop detailed models of the relevant representational mechanism. For now I restrict myself to one tentative concrete suggestion along these lines.

⁶¹ 1912/2003:13.

Many find the *higher-order thought theory* of consciousness (HOTT) unsatisfactory.⁶² On HOTT, it is when a sensory state in the brain (e.g. a visual percept as of the blue sky) is appropriately represented by a thought that this state is brought to conscious awareness. There isn't space to delve deeply into HOTT's perceived shortcomings, but one of these is certainly its apparent struggle to capture the phenomenology. An internal representation that some sensory state obtains seems rather a dry thing, inadequate to provide for immediate, technicolor phenomenology. Yet neutral monists are placed to make one or two helpful modifications to HOTT, which may well end up making it more appealing as an analysis of consciousness. First, we now take the qualia of sensory states to be ultimately irreducible features, the product of the structuring of qualitative ultimates composing these states in the brain. Thus qualia, which HOTT struggles to generate, are already built into our theory. The other side of the difficulty is the seeming distance between higher-order thought (HOT) and represented sensory state: why would an assertoric *thought*, even about an irreducibly en-qualified substrate, provide for the immediacy of qualitative conscious experience? The HOT might do better were it more intimately associated with the sensory state it represents: specifically, we might venture to *embed* qualia-carrying sensory states within the thoughts asserting them to be present. This would be something akin to Papineau's quotational model for phenomenal concepts,⁶³ but deployed now to capture subjectival awareness of qualities. A HOT would then have the structure "This state obtains "
" with a slot into which the (irreducibly) sensory state literally enters. That would be one way of getting qualia right into the mechanism that provides for awareness of them, and rendering such awareness suitably 'immediate'.

I've sketched one relationist/neutral monist attempt to capture that aspect of subjectivity which is the sheer awareness of qualities—a.k.a. consciousness.⁶⁴

⁶² For HOTT see e.g. Rosenthal 1991.

⁶³ Papineau 2002.

⁶⁴ It is worth noting that unrepentant panpsychists will require their own mechanism to explain macro-subjectival awareness, since they don't hold that *every* ultimate makes it into a subject's experience. Panpsychists don't typically think we experience the ultimates in *our toenails* or in *Saturn's rings*, so some relation amongst the privileged ultimates (in the brain?) will have to account for our awareness only of that limited set. Neutral monism is simply more parsimonious

What about the other aspect of subjectivity I have emphasized, the *separateness* of subjects, with their concomitant points of view? How does this fit into the sort of neutral monism under development? Here is a suggestion. If the conscious awareness relation is one of the representation (in some form) of qualia carried by certain portions of the brain, the brain is to be viewed as a device evolved with two key tasks regarding consciousness: It firstly marshals micro-qualia into macroscopic instances useful to the organism, such as macroscopic pains and visual percepts, and it enables awareness of these qualia for the organism. The set of qualia of which a given brain can be aware is limited to those qualitative instances within reach of its representational system—on the ‘neutral monist HOTT’, to those qualitative material conglomerations of cranial ultimates which can be embedded in the relevant higher-order thoughts. Each of our brains, then, houses a point of view, a representational system, which can take up only the qualia instantiated in that brain. The qualia of another person’s brain are simply out of direct representational reach. Brains are effectively a breed of self-directed microscopes, revealing their own inner qualitative character.⁶⁵

The neutral monist HOTT is an example of a relational theory of subjectivity the neutral monist might produce. Many other possibilities exist, awaiting development. The success of the neutral monist—and, formerly, panpsychist—anti-emergentist project frankly hangs on development of an adequate relational model of subjectivity. To the extent one considers such a model impossible, one must be dubious about the prospects for a structural explanation of consciousness, comprising qualitative and subjective elements. The alternatives we would then be driven towards would be emergentism on the one hand, and dualism on the other.

8. Conclusion

I have argued, overall, that to remain true to their anti-emergentism,

in having that awareness-constituting relation do duty for ‘consciousness’ as well. Another model of awareness that suggests itself is the self-representational theory, see e.g. Kriegel 2009 – though I doubt he would have much time for irreducible qualia (for an argument that irreducible qualia are compatible with physicalism see Coleman forthcoming).

⁶⁵ Thanks to Giovanni Merlo for the microscope metaphor. The self-representational model of neutral monism perhaps even more neatly explains the limited scope of subjectivity.

panpsychists must relinquish micro-subjects of experience, even as they take qualia to be irreducible, and should aim to construct a relational account of high-level subjectivity. This combination amounts to a form of neutral monism, so panpsychists should be neutral monists.⁶⁶ We can be to some degree optimistic about the chances of this shift, since even panpsychist Nagel seems open to the leading idea: 'Presumably' he says 'the components out of which a point of view is constructed would not themselves have to have points of view.'⁶⁷

References

- Armstrong D. M. (1978) *Universals and Scientific Realism. Vols. 1–2* (London: Macmillan).
- Broad, C. D. (1925) *The Mind and Its Place in Nature* (London: Kegan Paul).
- Brüntrup, G. (1998) Is Psycho-Physical Emergentism Committed to Dualism? The Causal Efficacy of Emergent Mental Properties. In *Erkenntnis* 48 (2/3):133-151.
- Chalmers, D. J. (2002) Does conceivability entail possibility? In T. Gendler & J. Hawthorne (eds.) *Conceivability and Possibility*, (Oxford: Oxford University Press).
- Coleman, S. (forthcoming 2013) Neuro-Cosmology. In P. Coates and S. Coleman (eds.) *Perception and Sensory Consciousness: The Nature of Phenomenal Qualities* (Oxford: Oxford University Press).
- Goff, P. (2009) Why Panpsychism doesn't help us explain Consciousness. In *Dialectica* 63 (3):289-311.
- (2006) Experiences don't Sum. In *Journal of Consciousness Studies* 13 (10-11):53-61.
- Humphreys (1997) How Properties Emerge. In *Philosophy of Science* 64: 1-17.
- James, W. (1912/2003) *Essays in Radical Empiricism* (New York: Dover).
- (1890) *The Principles of Psychology*.
- Kriegel, U. (2009) *Subjective Consciousness: A Self-Representational Theory* (New York: Oxford University Press).
- Nagel, T. (1979) Panpsychism. In his *Mortal Questions* (Cambridge: Cambridge University Press).
- (1974) What is it like to be a Bat? In *Philosophical Review* 83:435-50.
- O'Connor, T. and Wong, H. (2005) The Metaphysics of Emergence. In *Nous* 39: 658–678.

⁶⁶ For much more on neutral monism and its varieties see Stubenberg's admirable article (2010).

⁶⁷ 1979:194.

- Papineau, D. (2002) *Thinking about Consciousness* (Oxford: Oxford University Press).
- Rosenberg, G. (2004) *A Place for Consciousness: Probing the Deep Structure of the Natural World* (Oxford: Oxford University Press).
- Rosenthal, D. M. (1991) The Independence of Consciousness and Sensory Quality. In *Philosophical Issues* 1: 15-36.
- Russell, B. (1927a) *The Analysis of Matter* (London: Kegan Paul).
- (1927b) *An Outline of Philosophy* (London: George Allen and Unwin).
- Seager, W. E. (MS) Panpsychism, Aggregation and Combinatorial Infusion.
- (2006) The 'Intrinsic Nature' Argument for Panpsychism. In *Journal of Consciousness Studies* 13 (10-11):129-145.
- (1995) Consciousness, Information and Panpsychism. In *Journal of Consciousness Studies* 2 (3):272-288.
- Strawson, G. (2008) What is the Relation Between an Experience, the Subject of the Experience, and the Content of the Experience? In his *Real Materialism and Other Essays* (Oxford: Clarendon Press).
- (2006) Realistic monism: Why Physicalism entails Panpsychism. In *Journal of Consciousness Studies* 13(10-11): 3-31.
- Stoljar (2006) *Ignorance and Imagination* (New York: Oxford University Press).
- Stubenberg, L. (2010) Neutral Monism. In *The Stanford Encyclopedia of Philosophy* available at <http://plato.stanford.edu/entries/neutral-monism/#5>
- Tye, M. (2009) *Consciousness Revisited: Materialism without Phenomenal Concepts* (Cambridge, MA: The MIT Press).
- Unger, P. (2005) *All the Power in the World* (New York: Oxford University Press).
- Van Cleve, J. (1990) Mind-Dust or Magic? Panpsychism versus Emergence. In *Philosophical Perspectives* 4: 215-226.
- Wong, H. (2006) Emergents from Fusion. In *Philosophy of Science* 73, 3: 345-367.