

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

0. Introduction

Panpsychists hold that all matter possesses mental properties of some form. The main motivation for panpsychism is the perceived failure of the reductive physicalist project—the project that aimed to explain, *inter alia*, the existence of consciousness in beings like ourselves in terms only of conventionally understood physical matter and properties. Such physical matter and properties are presumed not to include mental properties at the basic level.¹ 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.² To adequately account for the generation of consciousness in beings like us, the panpsychist avers, we must take mental properties as part of the fundamental furniture of the material world.

I consider on the whole 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.³ I won't attempt to motivate panpsychism with respect to conventional physicalism, but will 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 the panpsychist position, and the question of where the intuitions driving panpsychism ought to propel one should those difficulties turn out to be insuperable.

¹ C.f. Stoljar 2006.

² For a recent example of this sort of argument see Strawson 2006. See also Nagel 1979, James 1890.

³ As compared with a wait-and-see solution like (e.g.) Stoljar's (2006). Stoljar believes we should await the future discovery of a non-mental intrinsic nature to matter which would make the production of macro-consciousness intelligible. The position advocated at the end of this paper is midway between Stoljar and the traditional panpsychist: I propose a here-and-now solution, with basic properties easily recognizable to us from mental contexts, nonetheless it is reasonable to view these properties as falling short of full-blown mentality in their native form.

There exists an unresolved tension within the panpsychist position, the seriousness of which has yet to be fully appreciated. I capture this tension as a dilemma, and proceed to offer some advice to panpsychists on how to resolve it. The dilemma, briefly, takes the following form. 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 typically 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, as I will show, 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 giving up her explanatory project, or giving up the claim that the ultimates are subjects. This is the dilemma. I argue that the latter option is to be preferred. This need not constitute a *wholesale* abandonment of panpsychism, however, since it is open to the panpsychist to hold 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 final section. The position eventually arrived at is a form of neutral monism, so that another way to express the argument of this paper 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 populated by micro-subjects of experience.⁴

⁴ 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 (c.f. 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 prevents the panpsychist’s dilemma branching into a tetralemma: we can rule out monism or nihilism about subjects as options. Thanks to Luke Roelofs for discussion of this point.

The dialectic revolves around a comparison between panpsychism and emergentism, so I spend a few sections outlining this comparison, before moving to a discussion of the combination problem, and from there to an examination of the difficulty that 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 may be usefully clarified by comparison with emergentism. The idea of 'emergence' concerns the appearance of a property at a higher level of being or organization that was not present at lower levels. Emergentism takes two forms, strong and weak. Weak emergence amounts to the idea that the appearance of the novel property was not predictable from knowledge of the makeup of the lower-level basis from which it arose.⁵ However this is taken to be strictly an epistemological limitation on the part of those trying to predict the occurrence of the higher-level property. Perhaps they didn't possess every scrap of lower-level information relevant to its generation. Or perhaps there is some deeper reason why occurrences of the property will remain unpredictable.⁶ But weak emergentism stops short of claiming that emergent properties are genuine ontological novelties; in fact they are taken to be wholly of a piece with the basis from which they emerge—their novelty is at root an epistemic one. Strong emergentism, by contrast, is the claim that some novel properties are not only unpredictable, but genuinely ontologically novel with respect to their lower-level bases. Such properties also possess autonomous causal powers.⁷ For these reasons it is plausible that strong emergentism concerning mental properties implies property dualism.⁸

Panpsychists shun emergentism, especially of the strong variety. They might allow that there exists a contingent gap in our understanding of the generation of

⁵ A sense of 'predictable' that allows prediction in hindsight—so close to 'explicable'.

⁶ As physicalists might say who employ the 'phenomenal concept strategy' to explain why we cannot intelligibly derive consciousness from matter. See for example Loar 1990, Papineau 2002. For present purposes I set aside such 'epistemic' defences of conventional physicalism—we are primarily interested in theories that aspire to explain consciousness as such (not merely to explain our alleged inability to explain it).

⁷ See e.g. Humphreys 1997.

⁸ See e.g. Brüntrup 1998.

macro-mentality from its material underpinnings, but it has been the very *raison d'être* of panpsychism to preserve ontological continuity between the world's lower ontological levels and the macro-conscious 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 inconceivable that ours is a zombie world,⁹ one lacking in higher-level phenomenal consciousness of the sort we each actually enjoy. Given the conceivability of physical zombies, panpsychists take the view that physicalist ontology is 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 thereby seeks to add just enough to the world's micro-ingredients, as conceived by conventional 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 clear display in the main philosophical vertebrae comprising the spine of modern-day panpsychism: the writings of James, Nagel and Strawson. James, surely 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 famous piece of explicit reasoning leading James to panpsychism was as follows: '*If evolution is to work smoothly, consciousness in some shape must have been present at the very origin of things.*'¹⁰

⁹ The usual restrictions apply: ideally conceivable for an idealized reasoner. See Chalmers 2002 for explanation of these ideas.

¹⁰ 1890: 148-149.

Nagel is largely responsible for re-igniting interest in panpsychism in recent decades, having been brave enough to take a serious look at a theory of mind that, towards the closing of the twentieth century, was largely ignored—or else simply considered laughable. The key premise of Nagel’s argument for pansychism is his rejection of emergence. He says

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.¹¹

From this principle, plus the claims that a conscious being is constructible from mere matter, and that a conventionally physical construal of such matter (as not possessing mentality in its fundamentals) cannot account for the production of consciousness, Nagel infers that consciousness must be present in basic matter.

Strawson’s work has brought panpsychism to the attention, and critical scrutiny, of mainstream philosophers of mind, as well as making a new generation of philosophers aware of its appeal. His argument for panpsychism departs little from the Nagelian formulation, but Strawson contributes a clear and forceful presentation, allied to a sustained defence of the key non-emergence principle. He expresses this thesis 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 conventional physicalism]¹²

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’¹³ These notions provide no conceptual bridge to consciousness—witness the

¹¹ 1979:182.

¹² 2006:15.

¹³ 2006:13.

conceivability of physical zombies—and so nothing short of positing phenomenality as a property of the ultimates will suffice as an explanatory basis for consciousness as we know it. So Strawson maintains.

Panpsychism thus stands opposed to emergentism.¹⁴ At panpsychism's root is a commitment to a smooth and intelligible generation of macro-consciousness by material ingredients. It has been said that austere reductive physicalism is suited to those with a taste for desert landscapes. Still, from the panpsychist perspective such reductive 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 said to be those theorists with a taste for only gently undulating landscapes. They decline to scale the sheer slopes of sudden mentality.

2. Constitutional Panpsychism

The panpsychist posits mentality as a fundamental feature of the material world, we've seen, due to a deep commitment to the non-emergence principle. But how should one construe the relationship between this basic mentality and high-level consciousness? In light of the panpsychist desire for a smooth and intelligible transition from the micro-material to the macro-conscious—what Van Cleve dubs panpsychism's 'mereological rationalism'¹⁵—the natural way to construe the relationship between the mentality of the ultimates and the conscious mentality of high-level individuals is to say that the former is *constitutive of* the latter.

The idea behind this *constitutional panpsychism*¹⁶ is roughly as follows. Consider a macro-subject, one of us, having an experience as of being cold, tired and

¹⁴ Thus Van Cleve (1990) is able to argue that emergentism is preferable to panpsychism as an outcome of Nagel's argument.

¹⁵ 1990: 218.

¹⁶ The name is due to Chalmers (talk given at the Hochschule of Philosophy, Munich, June 2011). A similar idea is Seager's (himself a panpsychist) 'constitutional supervenience', wherein the

smelling roast beef.¹⁷ This macro-level 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 smell of roast beef, still, phenomenologically-speaking, these three sensations are given to her all together, in one go. Her point of view is, or seems to be, suffused by all three at once, 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 mind must seek adequately to explain. On constitutional panpsychism this unified macro-experience is literally constituted by parts, groups of ultimates, instantiating just those qualities that feature in the experience. So the macro-experience of feeling cold, tired and smelling roast beef is constituted by one set of ultimates instantiating phenomenal coldness, another set that instantiates phenomenal tiredness, and another set that instantiates the smell of roast beef.

This is a somewhat simplified model of what the constitutional panpsychist has in mind, but the guiding idea is clear enough. Any phenomenal qualities figuring in the overall macro-experience are put in place there through being carried by the phenomenally-qualified ultimates that jointly compose 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 in its essence no more difficult than accounting for the overall composition of a painted canvas by reference to the various patches of paint that fill the canvas, and 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 relevant, experiential-field-composing, ultimate finds its way into the overall phenomenally-qualified experience enjoyed by the subject.

properties of the supervening level of an existent are implied in an intelligible manner by the properties and arrangement of items at the subvening level.

¹⁷ The example is Goff's (2009).

Constitutional panpsychism thus successfully evades emergentism, by providing a satisfactory realisation story for macro-experientiality by its material ingredients. The macro-phenomenal 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.¹⁸

In the present case, the instantiation of the unified macro-experience of coldness, tiredness and roast beef smell is nothing more than having a certain state composed of ultimates some of which instantiate phenomenal coldness, some of which instantiate phenomenal tiredness and some of which instantiate phenomenal roast beef smell.¹⁹ Put in the terms we employed earlier when spelling out the panpsychist's explanatory ambitions, it ought now to be the case that given the relevant 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 macro-subject does not instantiate an experiential state with the phenomenal qualities of coldness, tiredness and smelling roast beef. Since the phenomenally-qualified ultimates in question (jointly) instantiate these very qualities, and together constitute the macro-subject's experience, it does seem that this test is passed. The absence of the macro-subject's experience appears 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 and in these-and-those colors.

Constitutional panpsychism is thus the optimal version of panpsychism, the construal most faithful to the core concerns of the theory: to posit enough in the

¹⁸ 2005:10. See also Armstrong 1978. Note that the same idea is expressed by Nagel's anti-emergence premise in his argument for panpsychism. The panpsychist anti-emergence principle is effectively the claim that all non-basic properties are structural.

¹⁹ Note that Wong and O'Connor's characterization requires that the component parts not possess the property instantiated at the higher level. This requirement is met in the present case: the higher-level property is of experiencing coldness *and* tiredness *and* roast beef smell, and none of the relevant ultimates possesses this conjunctive property.

material fundamentals for a suitably smooth realisation story for macro-consciousness. Panpsychists, since they choose panpsychism to avoid nature's steep slopes, should embrace constitutional panpsychism. In fact, one might even say that for panpsychists it is constitutional panpsychism or nothing. But that is a corner of the dialectic we will approach in due course.

3. Panpsychism and Micro-Subjects

The panpsychist's ultimates possess phenomenal qualities. For example, the ultimates constitutive of the macro-experience of the subject just considered possessed phenomenal qualities like phenomenal coldness, tiredness and roast beef smell. But, intuitively, phenomenal qualities are those qualities that characterize conscious experience of the sort *we* (macro-subjects) enjoy; they are those qualities such that, in Nagel's phrase,²⁰ *there is something it is like 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 of experience 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 *point of view*. Something can be like something *for you* but not like anything *for me* precisely because we are distinct subjects occupying distinct experiential points of view on the world. Therefore it would appear that the phenomenal qualities under discussion cannot exist except as belonging to subjects.²¹ This, in turn, would convert the ultimates into subjects, since the panpsychist holds that ultimates possess phenomenal qualities. It is no surprise, 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; according to the panpsychist there is something that it's like to be a physical ultimate.' Strawson is equally clear on this point.²²

²⁰ Nagel 1974.

²¹ Stubenberg: 'the logic of these notions seems to require a subject: appearances appear to someone; sensations, percepts, and experiences...are always, and necessarily, owned by some subject or other' (1998:xx)

²² See e.g. his 2006.

On this picture, then, subjects like us are constituted by smaller subjects, micro-subjects, who are the first holders of the phenomenal qualities we each enjoy from atop our macroscopic perch.

Now an important question arises: What is the relationship between the *subjectivity* of these smaller subjects and the unified, single-perspective subjectivity of the human subject they compose in a given case? For simplicity we will consider only the relationship between ultimate-subjects and large-scale human subjects, without worrying whether the ultimates also compose larger subjects that compose us, for example subjects corresponding to the brain hemispheres.

We understood well enough, in rough terms, the constitution of a macro-qualitative state of feeling cold, tired and smelling roast beef by adverting to composing ultimates which, in groups, instantiated just those qualities. It was something like the way we might understand the composition of a painting by thinking about the various painted patches making up the canvas: we consider their qualities in isolation, and see how, by assembling the qualitative patches, we obtain the complete image.²³ But, I will argue, there is no way to understand, in a parallel manner, the assembly of subjects into a larger subject. Subjects do not combine. Therefore constitutional panpsychism is in big trouble. To lay the ground for understanding this difficulty, it is time we introduced into our discussion the notorious *combination problem* for panpsychism.

4. The Combination Problem

Seager, who coined the name ‘the combination problem’²⁴ describes it as ‘the problem of explaining how complex conscious states emerge from the primitive mental states ascribed [by panpsychists] to the fundamental entities of the world’.²⁵ This specification, though correct as far as it goes, is unhelpfully vague. In particular, it will be found to be ambiguous between two distinct difficulties. I

²³ The painting model meets Goff’s (2006) stricture that the experiential composite can instantiate no qualities that are not instantiated by its phenomenal parts (the ultimates).

²⁴ Seager 1995.

²⁵ MS:6.

will argue that it is only the second, the combination problem as it concerns subjects, that need concern the panpsychist. Unfortunately, this difficulty is also insuperable, and it is on this rock that the traditional doctrine of panpsychism runs aground.

Seager devised the name, but the combination problem is due to James, whose most well-known passage in this connection is the following:

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 feeling were set up, a consciousness belonging to the group as such should emerge. And this 101st feeling would be a totally new fact; 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.²⁶

The route to identifying the two variants of the combination problem begins by examining what James means by ‘feeling’ here. There are actually two possibilities, which are usually—albeit for pretty good reason—conflated. First, James may simply mean by ‘feeling’ something like ‘qualitative element’, a notion explicable as follows. Consider your overall phenomenal field at this moment—your complete experience, such that this can be taken to comprise other things which could be called experiences (such as the feeling of your bottom on your chair, or your visual experience of this page) and such that this complete experience is not itself a component of any further experience. It is a maximal property of you. Now isolate a single experiential 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. I call this visual experience of the white and black page a ‘qualitative element’—it is a qualitative element of the overall qualitative phenomenal field you enjoy, across your various senses and faculties.

²⁶ 1890:160.

What we should note about this particular qualitative element, which also 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, on the commonsense view anyhow, the only subject who is 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 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.

But when James talks 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 common sense, though, one might think, a sense congenial to the panpsychist. This is a use of ‘feeling’ which James apparently intends to be synonymous with ‘subject’, since he employs the terms ‘a consciousness’ and ‘a feeling’ as equivalents here. As noted earlier, phenomenal 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 talk about these elements in terms of the different qualities they exemplify. *Nota bene*: there is, as yet, no insinuation that qualitative elements—feelings so understood—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 different ways of understanding the problem James wants to highlight. Providing his use of 'feeling' is univocal throughout the quoted passage,²⁷ it seems James either claims that:

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

or:

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 this idea, as part of the denial that the hundred feelings could form a substantial identity with the one-hundred-and-first, by saying that 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 the notion of a product that does not emerge from its ingredients ('emerge' in the bad sense of *strongly* emerge²⁸), 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 note how the Jamesian ambiguity shows up in Seager's formulation of the combination problem. When Seager talks of the 'problem of explaining how complex conscious states emerge from the primitive

²⁷ Perhaps it is not. But then I am at a loss to reconstruct the difficulty James has in mind.

²⁸ For, recall, weak emergents are in principle fully explicable.

mental states', this does not tell us whether the challenge concerns the 'emergence' (I take it Seager's is an innocuous sense of 'emergence', equivalent to 'derivation') of complex qualitative fields from simpler qualitative ingredients, or the emergence of a higher-level subject from simpler subject-components. The ambiguity is also present, albeit 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, sometimes even within the same article. Perhaps he is best understood as being of the view that both aspects of the combination problem are equally serious for the panpsychist.²⁹

I do not agree that this is the case, however—that both problems are equally serious. The main 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 microscopic instances and 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 more conceptually challenging in such combination than there is in the everyday models of qualitative combination.

We have already considered a purely additive model, on which the qualities instantiated by the composite are simply the organized set of those instantiated by the components. Our illustrative analogy for this kind of combination was the composition of a painting by qualitatively distinct, determinate paint patches. Perhaps phenomenal qualitative combination is like this. Of course, the additive

²⁹ Goff's 2006, for instance, centers on 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 seems to suggest that what blocks combination is the fact that each pain is associated with a distinct subject.

model is simplistic, even as a model of building a painting. In painting an important part of the process, something the artist spends much 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 in a different way (as common sense conceives of things): here the blending of the various qualities gives white—in painting it gives black. 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, which can be readily conceived as living on in the ongoing contribution they make to the overall quality instantiated (take the blueness away from the purple paint dab and you are left with red, thus altering the quality of the whole). There is, in short, no serious difficulty associated with the notion of a complex macroscopic qualitative state being the combination of a set of simpler quality instances. The first Jamesian combination problem, then, is no problem at all.

What of the second putative problem, concerning the combination of subjects into a 'larger' subject? This, I claim, is the real combination problem, and it is insoluble to boot.

5. *Subjects do not Combine*

What is a *combination* of things? It is the organization of ingredients into a unified product, whereby the ingredients persist within the product, though they are modified by having combined. Intuitive examples are the combination of qualitatively and spatially distinct paint patches into a painting, the combination of cooking ingredients into a meal, the combination of two hydrogen atoms and one oxygen atom into a molecule of water. Properties can combine as particulars can. The dipolar property of a water molecule is the combination of the charges of its atomic constituents. In a combination the ingredients must survive in some measure in the unity: this is what's entailed by saying the unity is their combination or union, as opposed to being merely their effect or descendant.

Goff has argued that the existence of a set of subjects, however arranged, never *a priori* entails the existence of another subject, thus the panpsychist explanatory project falls short.³⁰ Goff is correct, but the situation for the panpsychist 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 mainstream panpsychism.

Earlier we connected the notion of a subject with that of a point of view.³¹ Neither notion is easy to get a grip on, but noting one or two basic features of each will suffice to make their interrelation clear enough for present purposes. The idea of being a subject goes with being an *experiential* entity, something which is conscious of phenomenal qualities. That a given subject has a particular 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. The relationship of subjects to the phenomenal qualities they experience, then, is somewhat like the relationship of shepherds to their sheep. Each subject possesses a point of view, and thereby 'gathers in' a certain unique field of phenomenal quality instances—those it experiences at a given moment. Subjects do not have access to the 'sheep' in other flocks: each subject's phenomenal field is private to her, in the sense that no other subject experiences it. Subjects, with their points of view, are independent, experientially isolated units, each closed off from the next phenomenologically. Even if the universe contained just a single subject comprising all matter, there is

³⁰ Goff 2009.

³¹ This conceptual connection is highly visible in Nagel's work; see for example the second half of his 1979.

a sense in which this subject would be an essentially *bounded* sphere of qualitative-experiential goings-on, with its own particular point of view—one could at least imagine a second universe ‘coming along’ or existing in parallel (in another ‘dimension’ perhaps), with a numerically different consciousness and concomitant point of view.

If each ultimate is a subject, each is a point of view, however simple, filled with its own conscious sensations. The essential difficulty for constitutional panpsychism is that points of view cannot combine.

Let us take as a template for combination the assembly of two hydrogen atoms and an oxygen atom into a molecule of water. First we’ll consider the combination in *particular*-terms, then in *property*-terms. We’ll see that these two modes operate in more or less the same way. This is because they go hand in hand: combination of particulars occurs thanks to integration of the properties of those particulars, and combination of properties occurs as the bearers of the properties are combined. When hydrogen and oxygen atoms combine into water they bond covalently, by sharing electrons. The oxygen atom completes its outer shell by borrowing one electron from each hydrogen. Thus the three atoms are *deformed*, intrinsically modified, by participating in the combination of water. Yet, importantly, all three atoms *continue to exist* once combination is achieved. Their union modifies them, but they survive in the whole, making an essential ongoing contribution to its nature. This is clear enough, in that 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.

Turning now to properties: a combination or unity acquires novel *systemic* powers. A water molecule forms a dipole, for instance: its charge is ‘two-faced’, being negative towards the oxygen side, and positive towards the hydrogen side. This feature in turn accounts for the way water molecules bond, and things like

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 the new systemic power of a unity to the powers of its isolated, pre-combination parts? The distinctive power of the combination is novel, in that none of its parts possesses (or possessed) it. It is not however an *autonomous* higher-level power, in the sense of being inexplicable with reference to the powers of the parts and their arrangement. 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 that 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 of their time around the oxygen atom, hence this side has 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 in fact manifests.³²

If micro-subjects were to combine into larger subjects, then, what we would expect to see is their coming together closely enough to modify 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 the atoms in the water molecule. But I think we can see that a set of subjects cannot combine in respect of subjectivity: it is impossible

³² James (1890:xx) mentions students of his who—erroneously, he thinks—believe they can discern in lemonade the qualitative elements of lemon and sugar. Sour lemon flavor *along with* sweetness is decidedly not what one tastes in lemonade, says James. He is right: one tastes a combination of these flavors, wherein the influence of each is discernible, but they are somewhat modified by their union. To an approximation, the sweetness of the sugar takes away from the sourness of the lemon. To see that both elements are present, it is enough to imagine taking one of them away: if one removed the lemon flavor it seems clear that lemonade would just be uniformly sweet.

both that a unified single subject (like one of us) should evolve as their product *and* that the component subjects should survive combination *as such*.

To understand why this is, let's consider a simple group of seven ultimates, and imagine different ways we might try to combine them. Our every attempt will end in failure.

First, it might be that after integrating the seven we are left with just a single subject. But this won't do. If only one subject survives combination, then at least six subjects have ceased to exist. We formerly had seven points of view in play, now we are left with one sole point of view. Thus six points of view have been extinguished, somewhere along the line. This isn't combination, it would appear, so much as a fight to the death. It doesn't help if we try to think of each of the original subjects as somehow contributing a 'quantity of point-of-view-ness' to the surviving subject. For the seven to survive in the whole *as subjects*, there would have to persist seven points of view. Yet by hypothesis we have a single point of view left over. Thus some points of view have completely ceased to exist. There's really no such thing as a 'quantity of point-of-view-ness'. Points of view are binary existents: they exist either wholly or not at all. So it might be that the resultant subject demands the contribution of *a* point of view from one of the antecedent subjects, in order for there to be a point of view around after combination. But once this contribution is made, there is nothing left for the other six original subjects to add: a point of view already exists, as required by the surviving subject. What is required of the redundant six, rather than making any further subjectival contribution, is their quiet disappearance, if a single, unified, subject is to be the upshot. After all, what could the other six points of view now add to their singleton successor except their subjective *diversity*, which would precisely be disruptive of its one-ness? Thus if we have but a single subject as upshot, combination of the original seven has not occurred.

For combination of the seven subjects to occur, they must persist in the combined subject. Thus we must imagine at least seven points of view in play after combination. But we had seven points of view at the start, so in what sense

is the product now a *combination* of subjectivities? It seems it is not: subjects are, we said, discrete spheres of experiential goings-on. If seven subjects exist, each with its point of view, then they are still a multitude, they have not unified into a single subject. The very idea of there being seven subjects is the idea of there being seven *distinct* points of view, each with its own separate set of experiences. What goes on in one subject experientially need not be the same as what goes on in the next. We can perhaps imagine that, through entering into a particularly close association, each of the seven comes to experience just the same range of phenomenal qualities as the others. Yet we still clearly have, even in this case, *seven* different token experiencings of these qualities. This is nowhere nearer to the genuine combination of subjectivities into one subject.

We should consider one further possibility: what if some *eighth* subject comes into being as a result of the association of the original seven? Might this not represent some sort of overarching 'über-subject', comprising as constituents the seven antecedent subjects and their points of view? The original seven must survive in some form, we have said, for this to count as combination. We can imagine the experience of the eighth über-subject to be the *qualitative product* of the experiences of the other seven, according to our best model of qualitative combination. If one of the antecedent subjects experiences a unitary phenomenal blue, another a red, another a green and so on, the eighth experiences all the colours experienced by the others singly, except now in one go, or in blended form perhaps. Yet the deep difficulty remains for even this attempt to combine subjects. Since the seven antecedent subjects survive, they each still enjoy their own token-distinct experiential states. Whatever the relationship between the experience of the eighth and the experiences of the original seven as regards *phenomenal quality*, it seems that the eighth's *subjectivity*—its possession of a point of view—can never be said to be the combination of the subjectivities of the original seven, even if it is after some manner their product (by James's 'curious physical law').³³ In a way this should have been obvious, it is just what's

³³?There is one point of view suffused by red-but-not-blue-or-green, another by blue-but-not-red-or-green, and so on. The uber-subject's point of view is suffused by all the colours. So its point of view does not feature as phenomenal constituents the points of view of the antecedent subjects: for this to be the case it would have to have as a constituent a phenomenal redness

involved in saying that we have eight subjects in total here. Eight subjects means eight points of view, which means eight token-distinct and independent points of view.

As long as the seven subjects survive they will not, in respect of their subjectivity, a unified subject make. The very idea is a sterling example of a contradiction in terms. Yet if we end up with but a single subject, and point of view, then at least six points of view no longer exist, and combination of subjectivities is not achieved either. Thus subjects cannot combine in respect of subjectivity. While the *qualitative contents* of consciousness may combine, consciousnesses themselves—subjects—cannot: this is precluded by the metaphysical logic of points of view. Therefore constitutional panpsychism founders on the assertion that the ultimates are subjects.³⁴

6. From Constitutional Panpsychism to Emergentism

A simple rejoinder seems available to the panpsychist at this point: Why should she endorse ‘combination’ as we have construed it? It is obvious that subjects cannot be combined, on the present understanding of combination which requires that the components survive in the unity. Surely the panpsychist needn’t accept this picture. Instead, by ‘the combination of subjects’, she need only mean to refer to that process, *whatever* its mechanics, whereby a plurality of underlying micro-subjects ‘gives way to’ the macro-subject. Why, in other words, should the panpsychist commit to the survival of the antecedent ultimate-subjects after they have done the job of generating the macro-subject?

I think the panpsychist who replies in this way has something like Humphreys’ idea of *fusion* in mind.³⁵ Humphreys offers fusion as a model of the generation of higher-level properties by sets of lower-level properties. The lower-level properties ‘fuse’ together, producing the higher-level property, and in the

exclusive of all other qualities, ditto for the blueness, etc. But it does not have such ‘exclusive’ experiences. So its point of view does not feature the antecedent points of view.?

³⁴ Nagel spies the problem: ‘How could a single self be composed of many selves?’ (1979:194). The preceding section (and the next) can be seen as bringing Nagel’s difficulty into focus.

³⁵ See e.g. Humphreys 1997.

process are obliterated. All that remains in existence is the fused product at the higher-level. In this vein, the panpsychist might venture a model whereby 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.³⁶ It seems this is what Seager has in mind with his concept of panpsychic 'combinatorial fusion',³⁷ and Goff apparently entertains the same model:

Perhaps the parts of my brain, before they came together to form my brain, had their own individual phenomenal lives. But when they come together to form my brain they lose their individual conscious identities, and somehow morph into o-experience had by the whole brain.³⁸

But fusion cannot help the panpsychist. The problem lies in something we have already emphasized: the essentially discrete, isolated nature of individual subjects with their independent experiential points of view. This feature of subjects means that their fusion could only provide for the *emergence* of the resultant macro-subject, implying a violation of panpsychism's core metaphysical commitments. To see why this is, we need to examine the constitution of structural properties.

Consider as a plausible case of fusion the melting down of several gold ingots into one big chunk of gold. The masses of the ingots fuse: Each ingot has its own mass before being melted down, but having been formed into one big chunk of gold, these individual masses no longer exist. Only the mass of the whole exists, produced in a logical manner by the fusion of the antecedent ingots and their respective masses.³⁹ What is important to note is this. In general, to avoid

³⁶ Humphreys allows the fusion of properties and the fusion of things.

³⁷ 'Perhaps we can regard the parts as infusing their properties into the whole and by so doing effacing themselves.' (MS:14).

³⁸ 2009:XX. Goff rejects this possibility on behalf of the panpsychist, claiming it to be tantamount to emergence. I think Goff is correct in his claim, but, puzzlingly, he offers no argument for it. I provide an argument to support this allegation of emergentism below.

³⁹ Someone might dispute this as an example of fusion—it might be thought that 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

emergence, lower-level properties must contribute to their novel product *in virtue of their metaphysical nature*, or, as we might otherwise put it, *while 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. Thus there is a requirement of metaphysical *sense-making* in the generation of a structural property: antecedents generate the novel structure by interacting according to what they *already are*, so there is a corresponding lack of surprise in the product. If we look back to the anti-emergence principle as expressed by Nagel and Strawson, in particular, we can see them both gesturing at this idea.⁴⁰

The case of fusing gold masses meets 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 one of the characteristic things that masses do. Other intuitive examples of structural properties likewise 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 dipolar charge of the molecule is simply the structural upshot of this free, self-expressive interaction. Another example is shape: one can see that in a Lego construction, 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 antecedent lower-level properties. Consider two examples Strawson employs:

developed below: to avoid emergence there needs to be an *intelligible contribution* by the fusing properties to the fused result.

⁴⁰ Recall that Nagel 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.

the emergence of the concrete from the abstract, and of the extended from the unextended. These are cases where the underlying properties cannot generate the new one structurally, because contributing *what they are*, abstraction, non-extendedness, makes no logical contribution towards what results. No amount of interaction or accumulation of different *loci* of abstraction can *by itself* yield concreteness, and likewise for unextendedness. Given what they are, abstraction and unextendedness just don't have *what it takes* to produce concreteness and extension—quite the reverse. Instead, the product positively defies the nature of the antecedents; thus concreteness and extension can only emerge, in these cases. The 'bruteness' which philosophers have always associated with emergence, is precisely this metaphysical defiance of the natures of the properties in the emergence base. We can further 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 is important to see that liquidity is a mere matter of the *dynamics* between a liquid's parts: molecules are loosely enough bonded that they can slide over one another, in effect. Thus liquidity is really just 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'⁴¹ Thus liquidity from the non-liquid is not emergence, and our account explains why.

Since subjects are *essentially* discrete,⁴² a set of pre-existing subjects can make no logical contribution to the subjectivity of a (unified) subject who exists down the line *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

⁴¹ 2006:13.

⁴² Without subjects, there are grounds for thinking that the world would best be understood as an undifferentiated single mass, i.e. that subjects are what provide for diversity *tout court*. I don't depend upon this intuition however.

subjects (at least partly) consists in. To create an über-subject from them, the predecessors' essential subjectivity—thus, discreteness—is exactly the feature to be elided, destroyed, overcome. To put the thought another way: 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. The diversity of the subject-set, of course, precisely derives from the essential oneness of any given single subject: since each subject is essentially a oneness, a set of subjects are essentially diverse, for they must be a set of onenesses. Essential one-ness from essential diversity—what would be involved in panpsychic fusion—thus looks like a case of emergence, by our criterion.

Fusion of subjects therefore cannot yield a higher-level subject as structural product. Micro-subjects cannot *form* together, their subjectivities adding up to the subjectivity of the macro-subject, as one might think the masses of gold ingots form together into the mass of a gold lump. But then the only way to characterize the fusion-style generation of a macro-subject by micro-subjects is to say that they brutally *cause* the successor to come into being, and by doing so expend 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. As O'Connor and Wong say:

[emergent properties are] those properties whose instantiation does not even partly consist in the instantiation of distinct properties by the entity 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.⁴³

The panpsychist fusion proposal clearly fits under this rubric. Therefore, by insisting on micro-subjects the panpsychist slides from constitutional

⁴³ 2001:10-11. This is not 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

panpsychism into emergentism. Quick Recap: To avoid emergentism the panpsychist must render macro-subjectivity as a purely structural property of an assembly of ultimates. But if the ultimates are subjects they cannot combine in respect of subjectivity. Therefore the subjectivity of the composite can only be causally related to the subjectivities of the composing ultimates. In this case the macro-subject has emerged.⁴⁴

Yet if panpsychists find themselves endorsing emergence at this juncture—in order to account for the production of high-level subjects—it is hard to see why they fought so shy of it earlier on. 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 the desire to avoid emergentism, the result we face is that a panpsychism which postulates ultimate-subjects is structurally unsound. *Constitutional-panpsychism-with-micro-subjects* inevitably ends up breaking its own anti-emergentist pledge.

To the extent one holds panpsychism to be committed to micro-subjects, one will thus be forced to conclude that the view is doomed: it suffers from an internal self-contradiction.⁴⁵ 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 that the emergence of macro-phenomenality from non-phenomenal physical components (i.e. classic physicalism) cannot now consistently be dismissed as a possibility. As Seager says ‘if panpsychism itself 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.’⁴⁶ Indeed.^{47,48}

⁴⁴ 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.

⁴⁵ 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.

⁴⁶ MS:4.

7. *The Neutral Monist Alternative*

Panpsychism's internal tension is encapsulable as follows. Panpsychists hold, in effect, that all non-fundamental properties are structural: they are reducible to properties at a more basic level plus the arrangement of the bearers of those properties. This is the non-emergence principle. It drives theorists to panpsychism since they find that, short of bearing phenomenal properties, physical micro-matter is an inadequate generation base for consciousness. If physical micro-matter lacked phenomenal properties, then macro-phenomenal properties could not be structural; they would have, unacceptably, to emerge. Yet panpsychists are also committed to the idea that phenomenal properties—qualia—necessarily exist *for* someone, a subject who experiences them. Hence the ultimates are held to be subjects. Yet subjects cannot combine: a higher-level subject cannot be a structural entity with respect to lower-level subjects putatively composing it, 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.

⁴⁷ There are panpsychists whose main motivation is not an aversion to emergence. Some are moved by the Russellian 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 (see 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 thought that conventional physicalism seems to lack the conceptual resources with which to account for consciousness. Moreover the explanatory (i.e. anti-emergentist) motive has been by far the main driver for panpsychism historically-speaking, so most panpsychists are indeed vulnerable to the present argument.

⁴⁸ Someone might object that panpsychist emergence is at least emergence *within a class* of properties, i.e. the emergence of (unified) subjectivity from (plural) subjectivity, and so such emergence is in better shape than 'physicalist emergence'—not the unacceptable option I make it out to be. In reply: The relevant contrast is between *structural* and *non-structural* generation of a property. Macro-subjectivity is not a merely 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 not dualists; in the physicalist case both properties might lie within the *cognitive*, on the assumption that phenomenal consciousness figures in cognition). The point is that the property is not logically derived from its base; its instantiation is more than a mere matter of the interactive organization of the basal items. Panpsychist and physicalist versions of emergence equally violate the structurality requirement, and once the panpsychist gives this up 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'.

From our historical survey of earlier, we can take it that the panpsychist's anti-emergentism (her 'mereological rationalism', in Van Cleve's memorable phrase) runs very 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. I do not possess a full solution to panpsychism's difficulties. What I set out below is rather something of a roadmap, indicating the future direction of travel for those motivated by the panpsychist's anti-emergentist impulse. Working out the details of the theory gestured at here is something for future research.

Panpsychists may well be right in holding that qualia—the qualitative properties we find in experience—are irreducible, and so (given the claim that all higher-level properties are structural) fundamental. Where they go wrong is in attaching subjectivity essentially to qualia; the notion that there has to be 'someone' around to experience a given quale. Making the ultimates into little subjects chokes them off from being able to constitute a macro-subject. The moral of this failure is that qualia must be split off from subjectivity—the experiential awareness of qualia by subjects. Macro-qualia considered merely qualitatively, like patches of redness 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 relational, reductive treatment. That's the only way to render macro-subjectivity as a structural property.

A position taking qualities as irreducible features of matter, but which has subjective awareness of these qualities be a relational (and so reducible) affair, is *neutral monism*, of the sort the later James inclines towards.⁴⁹ On neutral monism, unexperienced qualities permeate basic matter. Certain portions of matter 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

⁴⁹ See James 1912/2003.

conception of qualia whereby they can exist unexperienced, then a relational conception of subjectivity. Overall, the proposed movement of thought takes us some of the way back from panpsychism towards conventional physicalism, so it won't be surprising if physicalists find the end position significantly less off-putting than panpsychism. The difficulty is to persuade panpsychists that embracing neutral monism is the way to get what they want most, as well as most of what they want.

Recall that when discussing Jamesian 'feelings', we encountered the conceptual possibility of qualitative elements of experience existing without subjects to whom they belonged. This is a conceptual possibility to which we must now extend metaphysical seriousness. Of course, the notion of unexperienced qualia is not by any means anathema to mainstream physicalist positions. 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 appears to be an expanse of blue quality of which the subject is aware. This blueness is a quale instantiation. Now, direct realists who are also realists about colors explicitly externalize our qualia, holding that the blue quale apprehended in seeing the sky belongs to the sky (whatever *that* is) and not to the mind of the seer.⁵⁰ These theorists do not consider the blueness in question to evaporate as soon as nobody is looking at the sky, for they are decidedly *not* Berkeleians. This then commits them to the thesis that qualia can endure when unexperienced. That is just the proposal we need to make sense of here, although panpsychists are more likely to think of the qualia we are aware of as inhering in the first place in the matter of the brain. Panpsychists tend to be indirect realists about perception, holding that one is perceptually aware of external objects and properties by first being aware of internally instantiated qualia.

The neutral monist thesis—granted the irreducibility of qualities—is that all matter possesses qualia, which need not be experienced in order to exist: Just as the blueness of the sky or the redness of a fire engine persist when no-one is

⁵⁰ See e.g. Tye 2009. Lycan also puts forward this view, though he appears less of a realist about colors than Tye.

around to see them, on a commonsense or naïve realist perspective, so the qualities of the ultimates—perhaps qualities quite alien to our acquaintance—persist whether or not any subject is aware of them. A particular ultimate, then, might be phenomenally blue in the way a direct realist (and color realist) believes the sky is blue.⁵¹ 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.⁵²

Strawson is representative of the panpsychist mainstream in maintaining that 'there cannot be experience without an experiencer'⁵³. This is a tautology, very plausibly, but argumentatively a weak one: the neutral monist thesis that the qualities featured in experience can exist outside of being experienced is not the claim that any experience can exist unexperienced, which would clearly be absurd. It is the claim that some of the 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 the underlying Strawsonian claim, alternatively, is that such qualities are *ipso facto* experiential, then this is simply the claim neutral monists deny, as well as the claim standing in the way of 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, which has it that colors exist without perceivers, and that pains can wake us up in the night. If pains—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

⁵¹ C.f. the particulate imaginings of Unger 2005. There is a school of thought for which the subatomic entities postulated by physics stand in need of an intrinsic qualitative 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.

⁵² A neutral monist, as much as 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 physical irreducibility of qualia.

⁵³ 2006:XX.

structural makes subjectless qualia compulsory, as long as it is also held that qualia are irreducible.

The neutral monist considers that qualities do not require subjects around to experience them—the whole material world is en-qualified. What though 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 on this front, by turning conscious awareness into a relation. James has it 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.⁵⁴

For neutral monists, rather than being a kind of magical bubble or intrinsic phenomenal glow, consciousness is simply that relation whereby qualities are brought into awareness for their bearer: it is the curling of the material world back upon itself, to apprehend its own character. What is the nature of the relation in question? This is an area ripe for future innovation, for those sympathetic to the present argument. James is somewhat vague about the relation he envisages: he has it 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 such qualia, or the knowers. Such knowledge is the consciousness-constituting relationship. James also talks suggestively of *representation* playing a crucial 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, considered along with its two differing contexts respectively, by a new retrospective experience (1912/2003:12)

⁵⁴ 1912/2003: 13.

This is difficult. If we make the move to neutral monism, we will need to develop detailed models of what the relevant representational mechanism might be. For now, noting once more in passing—for the sake of the dialectic—that in order to render subjectivity a structural property (now) former panpsychists will be forced into some or other relational construal of conscious awareness, I restrict myself to offering one tentative concrete suggestion along these lines. Other proposals are most welcome.

Many find extant versions of the *higher-order thought theory* of consciousness (HOTT) unsatisfactory. On this theory, it is when a sensory state in the brain (e.g. a visual percept as of a blue sky) is appropriately represented by a thought that this state is brought to conscious awareness. There is not space to go into great detail about the perceived shortcomings of HOTT, but one of these is certainly its apparent struggle to capture the phenomenology: internal representations to the effect that some sensory state is present seem rather dry things, inadequate to providing for immediate, technicolor phenomenology. Yet neutral monists are placed to make a couple of helpful modifications to HOTT, which may well end up making it more appealing as an analysis of consciousness. For one thing, 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 seems to struggle to generate, are already built in to our theory from the outset. The other side of the difficulty is the apparent 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 if the sensory state it represented were more intimately connected with it: specifically, we might go as far as *embedding* qualia-carrying sensory states within the thoughts that assert them to be present. This would be something similar to Papineau’s quotational model for phenomenal concepts,⁵⁵ except deployed, in this case, to capture subjectival awareness of qualities. A relevant 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

⁵⁵ Papineau 2002.

of getting qualia right into the mechanism that provides for awareness of them, and rendering such awareness suitably 'immediate'.

I've just sketched one relationist neutral monist attempt to capture that aspect of subjectivity which is the sheer awareness of qualities—a.k.a. consciousness. What about the other aspect of subjectivity I have made much of, 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 (of some variety) of qualia carried by certain material portions of the brain, the brain is to be viewed as a device with two key tasks as regards consciousness: It firstly marshals micro-qualia into macroscopic instances useful to the organism for survival, such as macroscopic pains and visual percepts, and it enables awareness of these qualia for the organism. The range of qualia of which a given brain can be aware is limited to those qualitative instances within the 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 out of direct representational reach. Brains are effectively a breed of self-directed microscopes, which reveal their own inner qualitative character.⁵⁶

The neutral monist HOTT is an example of the sort of relational theory of subjectivity the neutral monist might produce. Many other possibilities exist, ripe for 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 that one considers such a model to be impossible, one must be dubious about the prospects of 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.

⁵⁶ Thanks to Giovanni Merlo for the microscope metaphor.

8. Conclusion

I have argued, overall, that to remain true to their anti-emergentism, 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. We can be to some degree optimistic about the prospects for 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.'⁵⁷

References

- Armstrong D. M. (1978) *Universals and Scientific Realism. Vols. 1-2*. London: Macmillan.
- 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)
- (1996) *The Conscious Mind* (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.
- Howell, R. (2009) The Ontology of Subjective Physicalism. In *Nous* 43 (2):315-345.
- Humphreys (1997) How Properties Emerge. In *Philosophy of Science* 64: 1-17.
- James, W. (1912/2003) *Essays in Radical Empiricism* (Dover).
- (1909/2005) *A Pluralistic Universe*. (Stilwell, KS: Digireads.com).
- (1890) *The Principles of Psychology*.
- Loar, B. (1990) Phenomenal States. In *Philosophical Perspectives* 4:81-108.
- Nagel, T. (1979) Panpsychism. In *Mortal Questions* (Cambridge: Cambridge University Press).

⁵⁷ 1979:194.

- O'Connor, T. and Wong, H. (2005) The Metaphysics of Emergence. In *Nous* 39: 658–678.
- Papineau (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
 —(1992) *Metaphysics of Consciousness* (London & New York: Routledge and Kegan Paul).
- Strawson, G. (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 (1998) *Consciousness and Qualia* (Amsterdam: John Benjamins).
- 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.