Rough Draft

<u>How to Combine Experiential Subjects</u>

It has been said that we have no concept of a mental part-whole relation, and more specifically cannot conceive the composition of subjects out of other subjects. From this it has been concluded that panpsychism, which appears to require such compounding, must either be false, or else must posit an incomprehensible 'emergence' of larger subjects from smaller ones, and thereby forfeit any advantage it might have claimed over emergentism by avoiding such inexplicable transitions.

I wish to argue that such a part-whole relation is not as incomprehensible as it has been taken to be. Of course it is not easy to understand, but neither is the physical part-whole relation, or indeed anything involving subjects of experience. But we *can* find resources within pre-theoretical experience that might be forged into systematic accounts, and my goal in this paper is simply to outline what I take to be some of the available resources.

But first I will explore more fully the contours and context of what has become known as 'the combination problem'. This will occupy Section 1; Section 2 will be a discussion of the phenomena in pre-theoretical experience that I think provide us with ways to conceptualise part-whole relations between subjects, discussing in particular the notion of a perspective, the experience of internal conflict, and certain aspects of our experience of objects. Finally, Section 3 will indulge in some cautious speculation about how these various phenomena of subjectivity might relate to the physical world that we perceive.

Section 1:

<u>A – Subjects and Objects</u>

Bill Seager introduces the term 'combination problem', defining it as "the problem of explaining how the myriad elements of 'atomic consciousness' can be combined into a new, complex and rich

consciousness such as that we possess."[1] The problem has sometimes been taken to be that we

positively know our consciousness to be unitary, and so can definitely reject the idea that it is a

combination of many smaller consciousnesses (e.g. Lovejoy, "The cognition…which is somehow

achieved through the brain, is not the sum of…atomic, non-cognitive sentiencies"[2]), but more

commonly it is simply claimed that, since we cannot make sense of this combination except as a

brute, mysterious postulate, "panpsychism itself requires a mechanism of emergence…[so] why not

take the theoretically more economical route of letting consciousness emerge directly from the

physical basis itself rather than from a mental basis"[3]?

The problem is often taken to be especially hard when applied to the *subject* of experience. For

instance, Philip Goff says that "the problem is that subjects of experience, i.e. things which have

consciousness…just don't seem to be the kind of things that can 'sum together' to make other

subjects of experience."[4] One response of some panpsychists – usually claiming the name

'panexperientialists' – is to say that the rudimentary experiences of small, simple entities have no

associated subject: they are "experiences without an experiencer"[5]. It's not clear how successfully

this meets the objection, and moreover, some prominent panpsychists are committed to views

about the subject which preclude such a position (e.g. Strawson writes that "I cannot avoid the

difficulty in the way Coleman can…because I believe in the ultimate identity of experience and

experiencer"[6]).

Since I do not want to be accused of selecting an easy version of the problem, I will here focus on

subjects, and assume that all experience belongs to some subject[7]. But I use the terms 'subject' and

'object' without assuming any particular theory about their nature or relationship, or how they are

---

[1] Seager, 1995, p.21
[2] Quoted by Van Cleve, 1990, p.219-220
[3] Seager, 2010, p.4
[4] Goff, 2009, p.130
[5] See, e.g. Coleman, 2006
[6] Strawson, 2009, p.61
[7] Similarly, it is sometimes thought that panpsychism becomes more plausible if elementary particles are attributed 'mentality, but not consciousness'. I am not sure what this means precisely, but I will not avail myself of whatever aid such a manoeuvre would supply.

to be analytically related to terms like 'mind' and 'mental state'. I endorse only the observation that mental states, or experiences, very often appear to us with a 'bi-polar' character, permitting us to differentiate two sides or poles, that which thinks and that which it thinks ('I' and 'it', or perhaps 'I' and 'you'). This is relevant because of course we can and do make sense of part-whole relations 'in the mind' – after all, we can mentally divide something into parts. But generally, under analysis, the diversity or composition in such experiences proves to lie towards the one pole, which we call 'the object', while the relative unity that 'binds together' diverse experiences is found in the other pole, which we call 'the subject'. The subject is what seems most unified, most individual, and most removed from all division and composition, and so the test for any account of mental composition is whether it can be applied not just to mental objects but to mental subjects.[8]

That is not say that understanding mental composition as it relates to the 'object' pole, or in terms of any other aspect of the mind, is at all simple. It is not, and it demands its own independent treatment. But I set that treatment aside for now.

B: Two Aspects of the Problem

So, focusing on the composition of 'subjects' in particular, the combination problem has two aspects. Firstly there is a problem of incomprehension – we supposedly just cannot understand what meaning to attach to sentences such as "Subject A and Subject B jointly constitute Subject C". Secondly, there is a dilemmatic problem: as soon as we try to make sense of such composition, we find ourselves pushed towards one of a set of highly implausible conclusions.

Thomas Nagel provides a frequently-referenced expression of the incomprehension aspect, saying that: "We cannot at present understand how a mental event could be composed of myriad proto-mental events on the model of our understanding of how a muscle movement is composed of

---

[8] Consequently in what follows I will generally mean by 'object' something mental, or perhaps something 'insofar as it participates in mental states'. When I wish to speak of mind-independent things that are perceived or acted on, I will generally use the phrase 'external object'.

myriad physico-chemical events at the molecular level."[9] Similarly, Goff writes that although he "can find no principled reason against supposing that there is some way of being related in which a group of subjects of experience…necessitates the existence of another subject of experience", nevertheless "we are unable to *understand*…a group of subjects being related in [this] way."[10]

The 'dilemmatic' aspect of the combination problem is less often noticed, but is discussed by Gregg Rosenberg, who calls it 'the boundary problem'. He frames it in terms of "a Scylla and Charybdis for [panpsychist] theories of consciousness"[11], of either allowing the flow of causal interaction to 'merge' subjects together, or not doing so. If we allow subjects to be merged, then because "the flow of interaction in the universe is inherently unbounded, and no merely abstract pattern presents a natural condition for containing it"[12], it seems that human consciousness would be 'swallowed up' into the consciousness of vast, galactic or even universal subjects.

Conversely, however, if we deny that subjects can merge, the only subjects will be fundamental particles, each presumably having only very simple and basic experiences – as Rosenberg says, "feeling, feeling everywhere, but not a drop can think"[13]. For convenience, we might call the first horn of this dilemma 'monism' (or 'subject-monism') and the second 'monadism'.

However, Rosenberg does not notice a third watery monster lurking in between Scylla and Charybdis, which I will call 'mid-level profusion'. This arises as follows: to avoid monadism, we say that subjects can merge somehow; to avoid monism, we say that such 'merging' does not erase or absorb the lower-level subjects, but keeps them running in parallel with the higher level subject – in addition to my human consciousness, there are in my head several billion neuronal consciousnesses, and trillions of atomic consciousnesses, all experiencing in ignorance of each other. And if there is a

---

[9] Nagel, 1986, p.50
[10] Goff, 2009, p.133
[11] Ibid.
[12] Rosenberg, 2004, p.87-88
[13] Ibid.

consciousness for the earth, or the universe, then I can be a part of it while retaining my individuality at the same time.

The problem with this is that since the consciousness of a whole does not erase that of its parts, the consciousness of my whole brain will not erase that of "all my brain except neuron A", nor that of "all my brain except neuron B", or any of the other trillions of combinations of neurons. But many of these other consciousnesses are *just as smart as me*. Even if I lost one neuron, I could still appreciate art, experience love, and write philosophy; hence it seems that trillions of other consciousnesses in my brain can also do such things. There are, so to speak, trillions of *people* sitting here, having full-strength human experiences, in complete ignorance of each other. Indeed, statistically 'I' am almost certainly one of them, and not the subject associated with the whole brain, as are (most of) 'you'. This will strike most people are bizarrely implausible.[14]

In this paper I will not directly discuss this second aspect of the combination problem, of how to avoid – or to make acceptable – the three prospects of monism, monadism, and mid-level profusion. My focus will be on undermining the sense of total incomprehension which forms the first aspect.

That said, the two aspects intensify each other. The fact that obvious models of how subjects might compose lead to implausible results prevents complacency that, while we might not understand the composition of subjects, it must happen somehow. Conversely, without understanding how it works, we are impeded from properly engaging with what it would mean to be a part of a subject (according to monism), a collection of subjects (according to monadism), or a set of overlapping subjects (according to mid-level profusion), and from making these options appear less incredible.

---

[14] The threat of mid-level profusion, though without that name, was first recognised by writers discussing the individuation conditions of objects in general, not specifically of minds or subjects – see, e.g., Unger, 1980, Merricks, 1998, Sider, 2001. Seager, 2010, recognises its relevance to the combination problem for panpsychism, while Goff, 2011, appears to enthusiastically embrace it, defending "unrestricted phenomenal composition", where consciousness belongs to every conceivable combination of particles, no matter how Gerrymandered. It should be noted, though, that mid-level profusion does not require unrestricted composition, but only the assumption that if 'brain' is a subject, then 'brain minus 1 neurone' would be as well, and the principle that consciousness is not erased by being encompassed in a larger whole.

Rough Draft

C: Other Related Problems

There is a particular form of the combination problem which I will not try to address, namely the claim that subjects cannot compose, because strictly, *nothing* can compose: any collection of things is never more than a *collection*, and never really itself a single thing. James, who is often cited as putting forth a 'combination problem' critique of panpsychism, holds this view: he says that "all the 'combinations' of which we actually know are effects, wrought by the units said to be 'combined', upon some entity other than themselves" – it is for this reason that "the theory of mental units 'compounding with themselves'…is logically unintelligible"[15]. Leibniz is another prominent exponent of this view that no composite can possess 'real unity', and consequently he is the poster-boy of 'monadism', denying that any of 'us' are, strictly speaking, composed of parts.

I have no intention of dealing with this question, of whether or how things in general can combine; I will consider myself successful if I can show that the composition of subjects makes *as much* sense as the composition of anything else. I mention this primarily to emphasise that panpsychism merely faces more obviously and intensely a problem also faced by other naturalistic theories: of explaining the unity of consciousness in a divisible composite brain.

This has often been used historically as an a priori refutation of materialism/physicalism: a unitary mind can't be identical with something divisible, but everything physical is divisible (see for instance Descartes' 6th Meditation, where he says "Every body is by its nature divisible, but the mind cannot be divided…This one argument would be enough to show me that the mind is completely different from the body, even if I did not already know as much from other considerations"[16]).

---

[15] James, 1890, p.158. For this reason I think it is misleading to write, as Goff does, "James [is]claiming that the relationship between the 101st feeling and the 100 feelings is not like the relationship between [a] party and [its] guests" (2010, p.5) – James is rejecting combination of subjects because he rejects all forms of real combination, not because he sees combination of subjects as radically different from other cases of combination.

[16] Descartes, 1641, p.32

Moreover, since the brain is not only divisible, but mutable, losing and gaining parts throughout its life, certain problems of personal identity (such as what to say about fission and fusion cases) can be seen as reflecting the same underlying issues. Once again, the combination problem for panpsychism is simply one version of a very general problem of understanding unified subjectivity in a material world. And that problem itself, perhaps, is just one version of the very general problem of 'the one and the many'. I think this should encourage us to hope that, if these problems admit of solutions at all, panpsychism should be able to address them at least well enough to be on a par with its rivals.

Section 2:

A – Subjects and Perspectives

The first phenomenon I want to discuss is the concept of a 'perspective', which I would roughly define as 'subjective factors, considered objectively'. The cognitive role of this concept can be best explained by a caricatured story of how we might imagine it arising in subjects who lacked it.

At first, they have perceptual experiences, and they naively assume that the various features of these experiences always convey, and correspond to, various objective features of external objects. However, certain occurrences disturb this extreme realism. Two people standing in different places describe the same object as having different shapes; one person says that two things smell the same while another says that they smell similar but distinct; one person (who happens to be a giant bee) says that a surface is visibly striped, while another (who happens to be a small human) says it is visibly solid. And of course, each individual will find that things appear different at different times, in ways that it is hard to explain as reflecting changes in the external world.

From this it could be concluded that not every feature of experience corresponds to some feature of the external world; instead, some features correspond to facts about the subject perceiving it, or to

facts about their relation. Subjects in different locations, with different discriminatory skills, or sensitive to different wavelengths of light, will see the same object in different ways.

The notion of a 'perspective' would arise when it is noticed that some of the 'subjective' features of experience can be shared among different subjects. The most systematically intelligible example is of course spatial location: not only does this object look square to subject A, standing in spot X, and trapezoid to subject B, standing in spot Y, but moreover it will look square to *any* subject standing in spot X, and trapezoid to any standing in spot Y. That is, it is square 'from the perspective of spot X', but trapezoid 'from the perspective of spot Y', and so its apparent shape is 'relative to a perspective'.

Sometimes the term 'perspective' is used to mean *only* spatial perspective, but it can also be applied much more broadly. For instance, we might analogously say that a single surface is 'striped from the perspective of bees, who can see ultraviolet light, but solid from the perspective of humans, who cannot'. And of course we often speak of 'perspectives' in cases of non-sensory 'appearance', e.g. 'from his perspective it seemed like the most reasonable thing to do', etc.[17] The core idea is that 'his perspective' both explains certain facts (why he acted a certain way) in terms of factors specific to him ('subjective'), but can itself be understood (or 'taken', or 'shared') by others ('objective'). Hence I say that the concept of a perspective is a device for speaking objectively about subjective factors.

That is the familiar idea I want to talk about; how is it relevant to the combination problem? Recall that the combination problem arises, we are told, because subjects do not compose into other subjects, although objects do. Hence it is relevant if both of the following theses are true:

1) Facts about 'subjects' can be at least partly translated into talk about 'perspectives', and

2) 'Perspectives' are susceptible of composition in the same way that objects are.

---

[17] We also speak of 'perspective' as a mass noun, as in 'gaining some perspective'. What we seem to mean here is partly a matter of 'being able to see something from multiple perspectives', and partly the resultant recognition of how certain features of a situation are in fact relative to a particular perspective.

What I mean by 1) is that what subjects do *qua* subjects seems to be primarily a matter of thinking, perceiving, or performing other mental acts, and it seems that almost all of these can be described in terms of perspectives: not only can we express differences in perception by speaking of how the same objects appear from different perspectives, but we can also go some way to expressing differences in reasoning, affect, or volition, by speaking of how the same evidence, or premises, or situation, might appear to support different conclusions, or justify different actions, when considered from different perspectives.

Thus it would seem that something close to a complete description of a subject, in terms of the totality of their mental states, their dispositions, etc., can be 'translated' into a detailed specification of a perspective. Indeed, we do often identify perspectives by reference to subjects – we recommend looking at something from X's perspective, presuming that the person we address will be able to turn what they know about X, the person, into an idea of X's 'perspective'.[18]

Given 1), point 2) says that perspectives are amenable to composition. Suppose I watch a film, and evaluate it from a political perspective; it will appear a certain way, with certain elements more salient, possessing certain properties (e.g. 'reactionary', 'othering', etc.). Then I watch it again, but evaluate it from a more purely aesthetic perspective, so that different elements stand out, and different properties attach to them (e.g. 'beautiful', 'incongruous', etc.). But I can also take the further step of actively connecting the two perspectives, looking for the ways that aesthetic properties feed into and enable political messages, and vice versa.

Is the combination of my political and aesthetic perspectives itself a perspective? It seems so. We would not consider it strange to say that a particular person had 'a distinctive perspective on films', because they routinely applied both political and aesthetic categories. Nor would it be strange to contrast this perspective with other perspectives, such as an uncritical perspective focused on what

---

[18] There are many promising avenues for spelling out this talk of perspectives more concretely, using possible worlds or relational properties. I will remain non-committal on this, since my goal is to indicate pre-theoretic resources.

was fun, awesome, or sexy. Considered more abstractly, it seems plausible that for any two ways of perceiving or thinking about objects, there could be a way of perceiving or thinking about objects which would incorporate both.

That is certainly not to say that composition is automatic; I might watch some films aesthetically, watch others politically, but it never occur to me to combine both modes of analysis. I might even, perhaps, watch a single film, analysing it in both ways, 'in parallel' so to speak, without either analysis impacting the other; it is unclear whether we would in this case say that 'my perspective' was the combination or not. Perhaps we would say that 'my perspective' was a weakly-integrated combination. But these details need not be decided here.

Now, if it makes sense to combine two perspectives into one, it might also make sense to decompose a single perspective into its elements, which are themselves perspectives. We can easily imagine this being done by literary critics and so forth, but I think similar things can be done with very familiar examples.

For instance, I am looking at a half-empty cup of coffee. My perspective on this cup determines how I see it (it is to be drunk, it is mine, it is a cup *of* coffee, not a cup with some liquid accidentally in it, etc.) and a different subject would see it very differently (e.g. a Martian scientist, a mouse, a person desperately looking for projectiles to hurl at approaching zombies, etc.). Can we distinguish perspectives within this perspective? I think so. For instance, part of my perspective on the coffee is that of a being that consumes coffee, but not ceramics – this gives the brown liquid a special salience that the white cup lacks. Another part of my perspective is that of a being with hands – this gives the white cup a special 'affordance', i.e. a relation to certain things I could *do* with it, such as picking it up. These elements could come apart – I might be able to pick up the coffee, but not desire to drink it, or vice versa. And each element is a perspective in the sense of a general way of perceiving objects, that will apply to other objects: other things to drink, other things to pick up, etc.

Importantly, there is no question of these 'smaller' perspectives vanishing into the 'larger' perspective that incorporates them, because each perspective persists as long as the *objects* are perceived in a certain way. If I see the coffee as 'to-be-drunk' and also as 'able-to-be-picked-up', then both of those perspective persist; if I see it as 'able-to-be-picked-up-so-as-to-be-drunk' then an integrated perspective that incorporates both also exists.

Sometimes perspectives will be in conflict, as with two incompatible spatial positions, or two incompatible religious positions. It is an interesting question whether these can be combined, to form 'unstable' perspectives containing 'tension' or 'contradiction', or are simply incompatible full-stop. But once again I must decline to directly address this issue.

Note a further point about perspectives: sometimes we regard them simply as objects, and sometimes we *take* them. To put it another way, we can think *about* them, as with anything else, and we can also think *with* them, as we cannot do with, e.g., salt. The difference between recognising that another person takes a different perspective (and perhaps believing many things about that perspective, and its quality or usefulness) and actually taking their perspective, if only temporarily, is often transformative.

Moreover, we can sometimes find ourselves apprehending a perspective 'opaquely', in the sense of being able to think about it, but not being able to take it. For instance, if a friend says "Recently I've been looking at this issue from a whole new perspective, and it's really opened my eyes!", but is then cut off and tells us no more, we know many things about the perspective (who takes it, when, on what object, and with what degree of effect), but not enough to take the perspective ourselves.

A natural objection to make is that not all of our talk about subjects can be translated into talk of perspectives. In particular, subjects are *concrete* phenomena, whereas perspectives, or 'ways of seeing things', seem more like abstract phenomena, just as walking animals exist concretely, but ways of walking exist abstractly. We can of course speak of perspectives as concrete if we treat them

as instantiated universals, i.e. 'perspectives actually taken'. But then we have re-introduced a reference to subjects, as the things which take perspectives.

This might appear particularly when talking about volition: talk of perspectives can express the fact that some action 'seemed right' to someone, or even 'seemed obviously the best thing to do', but it cannot, surely, express the fact that someone therefore *chose* to take that action. I see things from a perspective, but then *I*, and not any perspective, *do* something. In general, it is natural to think that only concrete beings can exert causal powers.

I think this objection is correct; that is why I do not think that the notion of a perspective on its own could ever provide a complete resolution of the combination problem. I claim merely that it is one resource for doing so, which must be deployed along with others. The issue of volition and causal power, in particular, is relevant to my next topic.

B – Inner Conflict and Inner Multiplicity

The second phenomenon I want to discuss is the experience of inner conflict, where different elements or aspects of a single mind are 'in conflict' with each other. This description covers a variety of cases. The simplest is perhaps attentional competition: different mental contents 'compete' for scarce attentional resources, in the sense that attending to one typically means not attending to the others, and certain features of contents (intensity, salience, etc.) increase or decrease their chances of being attended to, and thereby decrease or increase others' chances.

A more complex case of competition among mental contents is found in deliberation, whether 'theoretical' deliberation about what to believe, or 'practical' deliberation about what to do. Here we can distinguish a way of speaking of (the ideas of) different options as competing with each other, and also a way of speaking of different 'factors' (desires, arguments, etc.) as competing with each other to determine the mind to believe or do what that factor favours.

For instance, suppose I am deliberating about what to do this evening: I could stay in and watch a film, stay in and work, go out with X to a pub, or go out with Y to a nightclub. On one level, my representations of these four options are competing with each other – if I choose one I cannot choose the others. On another level, the competition is between my desire for relaxation, my desire for company, my desire for professional success, my desire for fun, my desire to conserve money, etc. These factors each 'support' one or more options more than others, and 'struggle' within me to have the option which they support accepted.

It is important for certain purposes whether these 'factors' are identified with 'desires' or not. Here, though, the question can be fudged – thus I will primarily use the neutral term 'factor', but occasionally slip into speaking of 'desires' where it sounds more appropriate.

Competition between motive factors is only indirect and contingent, since two factors may both support the same option (desire to conserve money and to relax might 'ally' together to support staying in, while desire for company and desire for fun might 'ally' against them). But it is also at this level that the most interesting things happen: we experience anguish when our desires clash, whereas when our desires are in harmony but we face a collection of mutually exclusive options, we serenely choose the one our desires agree in supporting. So in what follows, references to 'inner conflict' will be to this 'deeper' level, of motive factors or 'desires' struggling over our choice among options.

So this is the second familiar phenomenon I want to discuss. As the above example illustrates, we commonly personify the factors involved to some extent: in explaining what is going on, we find it convenient to attribute them agency, because in some sense they do appear to exert causal power, and because in some sense they exert this power in a 'teleological' way, guided towards some goal they value. We are more likely to adopt such a 'personifying' approach when the conflict is particularly sharp and 'acrimonious', i.e. where the triumph of one factor may be ruinous from the perspective of others, and where such a ruinous triumph seems likely. In these cases people may

even express feelings like fear or helplessness towards factors within themselves, and speak about being 'taken over' by something, or 'unable to control' their own actions.

But is this really of any metaphysical import? After all, we personify many things, from ships to nations. Usually we regard such talk as merely metaphorical: there are not 'five people' at war in one head, but one person with five incompatible desires.

(Before proceeding we might note a third way to describe these cases: a single person intractably taking five incompatible perspectives. My desire for fun and my desire for company would then be described as 'me, taking the perspective that fun is the most important thing' and 'me, taking the perspective that company is the most important thing'. Cases of bitter inner conflict might then be described as "me, struggling against 'myself from one perspective', from another perspective".)

But let us return to the question: is there any good reason to see the language of inner multiplicity, deployed to explain inner conflict, as having any metaphysical significance at all? It is worth pointing out that some fairly prominent philosophers have thought so. Plato stands as perhaps the most august proponent of the idea: in Book IV of the Republic he argued that consideration of the different sorts of internal conflicts humans display warrants the positing of three distinct elements of the soul, whose proper relations are then the subject of the remainder of the book. Freud is a more modern influential proponent of similar ideas, for whom the division (into id, ego, and superego) is explained as a process of developmental differentiation rather than a permanent split of supernatural origin. A more recent account centred around internal multiplicity and conflict is Daniel Dennett's 'Pandemonium' model in 'Consciousness Explained'(1991).

Nevertheless, there are good *prima facie* objections to this line of thought, in particular the following two. Firstly, isn't the claim of inner multiplicity in fact denied by our practices of holding people responsible for their actions, even actions taken during periods of inner crisis? Indeed, isn't it denied just as much by the agent's own description of events, when they claim to be 'overwhelmed'

or 'taken over' precisely by forces they regard as external; by speaking of *themselves* being thus overwhelmed they appear to affirm their unity even while bewailing their division.

Secondly, if deliberation is to be construed as a matter of various 'factors' battling with each other, what are we to say of the 'battle-ground' itself? The very fact that these 'rivals' can 'compete' with each other seems to be explained by reference to an underlying unity which connects them, namely the unity of sharing the same mind, belonging to the same subject.

I think these objections can be answered – not so as to conclusively demonstrate inner multiplicity as the only defensible interpretation of inner conflict, but sufficiently to demonstrate that it is *one* defensible interpretation. We can always interpret cases of inner conflict so as to accord with the complete unity of a single subject, but we can also interpret them in terms of inner multiplicity, and neither method is more intellectually suspect than the other.

On the first point, that our normal practice with respect to people (including ourselves) who feel 'divided' is to continue to treat them as a single person, I would suggest that this practice is overdetermined, and so would be adopted regardless of its fit with the metaphysical facts. For a start, there are numerous social reasons to continue to regard and treat people as unitary – for continuity of property and obligations, for criminal and civil accountability, and because doing so may be a good way to encourage them to seek to resolve or prevent inner conflicts.

Moreover, even the most divided people genuinely are unitary in numerous respects – physically, for a start, and in their history, in the information they have access to, and so forth. We are dealing with both *ambiguities* over what sense of 'unity' or 'subject' is at issue, and with matters of *degree*. If we are inclined to think of subjects and their unity in digital terms, as all-or-nothing, then we will naturally tend to explain away evidence of partial disunity as merely superficial, at least up until the point at which it becomes so severe that we regard the person as psychotic or otherwise pathological.

The second objection was that interpreting inner conflict in terms of inner multiplicity is manifestly inadequate because it fails to account for the very fact that these 'conflicting' forces are all present to one another in such a way as to allow direct conflict (as opposed to the indirect conflict, via. words or weapons, that occurs when they are in separate people).

This objection flows from a number of considerations, at least some of which can be set aside quite easily. Firstly, it will partly be prompted by the remarkable *degree* of interaction between forces 'in a single mind', the speed, complexity, and strength of their influences on each other. But this is easily explained by pointing out that they are communicating via a set of neural connections that enable fast, complex, strong interactions in perfectly unmysterious ways (myelination, etc.)

Secondly, part of the objection will be that the *type* of interaction is quite different – it is directly mental-on-mental, rather than (as in all other cases) mental-on-physical-on-mental. But if we are considering the prospects for panpsychism, this begs the question: according to panpsychism, *all* interactions are mental-on-mental. Even if those mental things have physical properties, the interaction of these two sets of properties in causation is so disputed and little-understood that the opponent of panpsychism cannot mount a non-question-begging objection along these lines.

What remains of the objection when these issues are set aside is something like this: when two of my desires struggle against each other, they do not struggle against each other as a subject struggles against an object, because their interactions are not subject-object structured in the way that my interactions with a fire, a dog, or another human are. That is precisely why they are interactions within one subject. This objection is important, and to address it I will need to move on to my third topic.

Rough Draft

C – Phenomenological Distance and Fusion

When we perceive external objects we usually experience them as being at a certain, greater or lesser, 'distance' from us. Obviously one sense in which this is true is that we perceive things in space, as having spatial relations to us and to other things. The first point I want to defend in this third discussion is that this is not the only, or the most important, sense of 'distance' we experience.

The most obvious case of a non-spatial meaning of 'distance' is emotional distance, as when we describe our relationships with someone as 'close', or say that someone is acting 'distant'. Our social interactions, in fact, are often concerned primarily with manipulating our level of 'distance', either trying to draw someone 'closer', or to 'withdraw' from them, or to maintain an appropriate distance.

What is the analogy between this case and the spatial case? The obvious connection is that in both cases, what is 'closer to us' is both easier for us to act on, and more likely to act on us. Typically, any form of physical influence is stronger, and less likely to be impeded, between nearer objects; analogously, people we are 'close to' can easily affect our feelings, because we care more about them, and because they know more about us; conversely we can more easily affect them, whereas someone will seem 'distant' when they seem concerned with something unrelated, and thus less likely to be emotionally affected by our utterances or actions. This entails a difference in epistemic accessibility: what can act on us more easily is typically also what we can more easily perceive and learn about, and what we can more easily affect will typically be more able to perceive us.

So in our emotional relations we perceive people in terms of their 'distance', in this expanded causal sense; do we find the same sort of perception in other, more 'physical' cases? I think we do. We often experience desires to 'withdraw from' repulsive or unpleasant things, or 'approach' attractive things, which generate not only spatial movements but other actions, such as placing a barrier between ourselves and a thing in order to 'distance' ourselves from it by impeding our interaction with it. We can also take purely mental actions to adjust our experience of closeness with

something: we can both mentally 'withdraw' by trying to become less interested in, and so less vulnerable to, a given object, or we can 'approach' it by focusing our attention on it, to the exclusion of other things, so that small changes in it will produce greater changes in our phenomenology.[19]

I will refer to this as 'phenomenological distance', defining it as a distinctive phenomenological dimension reflecting the perceived ease with which a thing can affect us, and vice versa. It seems to me that this dimension plays a significant role in our everyday experience of objects.

The second claim I want to make about this 'phenomenological distance' is that at its theoretical minimum (i.e. maximum closeness) it would constitute identity, and that although we probably never in fact experience such an absolute minimum with any external objects, the prospect of it – of an object approaching us ever more closely until it 'merged' or 'fused' with us – is implicit in what we do experience, and exerts a significant influence on our affective lives.

I have in mind a number of phenomena: sexual and romantic desire, and the fears and anxieties it provokes; feelings of disgust, and their special relationship to boundaries (such as living/dead and internal/external); possibly mystical and religious yearnings for fusion with divinity; humanity's characteristic concern for privacy and concealment, etc. The role played by the prospect of 'fusion' is different in each case – sometimes desired, sometimes feared, sometimes abstracted from immediate physicality, sometimes immersed in it, etc. But it seems to me that analysing such phenomena in terms of a prospect of fusion is at least one natural way to describe them.

It might be objected that, although we may occasionally speak in such terms, they can easily be analysed away – for instance, even retaining talk of phenomenological distance, it might be said that we are never afraid of, nor desirous of, 'fusion' with an object, but only of a certain finite degree of

---

[19] The internet phenomenon of 'shockers', thankfully now on the wane, relied on this: it first showed people a detailed image and told them that some game or test required them to identify some small feature (e.g. a difference between two pictures of a scene), inducing them to mentally 'move closer' to it. Then, after a certain time elapsed, it played a shrieking noise while showing a hideous face expand rapidly on the screen (as though 'coming towards' the viewer). Because they had become 'close' to the computer screen, this sound and image had an exaggerated and hilarious effect, whereas if they had been forewarned or otherwise able to maintain a sense of 'distance' it would have much less effect.

closeness that exceeds our present degree. As before, I happily accept that such redescriptions are possible, and defensible. I need only deny that they are necessary, and I see no reason to think that they are decisively preferable as descriptions of our emotional phenomenology than are descriptions in terms of anticipated fusion.

Of course people will typically, if asked, deny that they wish to be 'fused' with anything, even a lover. But this could easily be said to be a product of factual knowledge, and not a reflection of the underlying facts about the structure of their emotions. That is, people know that, given the facts of human biology and technology, any literal 'fusion' of two human beings (or of a human being with any object except perhaps food and air) would be highly dangerous and would lead to a marked loss of ability to pursue their other projects. Or they might simply be employing foresight to recognise that any literal fusion would likely be permanent, and that any desire for fusion they might have would probably not be so permanent. So I do not think this objection is decisive.

Now, how does this relate to the combination problem? The point arose in the previous discussion that interaction between two states of a single subject seemed to be qualitatively different from interaction between two subjects, each taking the other as an object, regardless of the degree of speed or complexity in that interaction. As a result, it was said, experiences of inner conflict cannot be a model for any process of composition or decomposition of subjects.

These reflections about phenomenological distance and phenomenological fusion cast doubt on that. They suggest that, on at least one plausible analysis of our experience, there is a difference of *degree* between identity, as we feel it between 'me thinking X' and 'me thinking Y', and distinctness, as we feel it between ourselves and external objects. This difference of degree corresponds to differences in the extent of causal integration between two things; where there is perfect causal integration, then there is identity (for we are then dealing with a single thing, at least if we define 'things' in causal terms), and where there is relative causal independence, then each perceives the other as an 'object' distinct from itself.

If we applied this to the relations among our own 'desires' or 'motive factors' or other components, we would probably have to say that they are exceptionally causally integrated, but may diverge just slightly from complete identity, precisely in those cases where they causally 'disrupt', 'interfere' or 'compete' with each other. This may not just be because they are 'right next to each other' inside the skull, connected by a billion little wires, but also a matter of 'sharing a substrate' – that is, if these factors are not physical 'bits' of the brain, but relatively persistent 'patterns' of neural activity, then they will often be instantiated in the same regions of matter.

Either way, their causal integration far outstrips anything that a human being experiences with anything outside itself. If we hypothesised that they were in some sense 'each' a subject, they would experience each other as more 'phenomenologically close' than in any sort of experience we can imagine. The closest analogue to their perceptions of each other might well be some sort of intense narcotic or religious experience.

An objection has probably been brewing in the reader's mind. Surely I'm pulling a fast one, by directing our attention to *experiences of identity and distinctness*, when what is really at issue is *identity and distinctness of experiences*. If two subjects both *feel* as though they are 'merged' with each other, that does not change the fact that they are *two subjects*, which manifests itself as the possibility of their independently ceasing to feel thus 'merged' – that is, you may feel as though we have 'become one', but your error is proven by the fact that it need not follow from this that I *also* feel that way. To lend credence to such mystical or romantic fantasies is to be tender-headed.

Certainly, I must concede that people very often do have such fantasies – people often misjudge the real closeness, or fusion, between themselves and some other. But people are also often wrong in their factual estimations, or feel misplaced fear, misplaced trust, misplaced indignation, and this does not show that factual estimations, fear, trust, or indignation are *in general* invalid or misleading. Attitudes like fear or indignation contain their own standards, by which particular instances can be judged erroneous – e.g. an arachnophobe's fear is misplaced because spiders are

not really dangerous, and a bigot's indignation is misplaced because miscegenation is not really obscene. Similarly, an erotomaniac's inflated sense of closeness with the object of their affections is not real closeness because the two are not in fact causally integrated to any great degree: the feelings or actions of one are neither very well known by, nor very able to affect, the other.

Once we accept that feelings of closeness etc. may in particular cases be misguided, it is just question-begging to say that experiences of identity and distinctness are *never* relevant to the identity and distinctness of experiences. It may of course be true, but we are not compelled to say it in order to be hard-headed. My point is not that we should always avoid overruling people's pre-theoretical experiences and what they seem to signify: it is that we should only overrule them where we have good theoretical reason to do so.

If we had good evidence that selves were metaphysically inviolate, wholly distinct and never subject to gain or loss of parts, then such overruling would be justified in this case. But in fact theoretical considerations seem to tell in precisely the opposite direction: not just panpsychists, but all naturalists, have a theoretical problem of fitting unified subjects into the divisibility and continuity of the physical realm. If some aspect of pre-theoretical experience seems to be relevant to the ways this unity breaks down (or expands), then we should take an interest in it regardless of its association with kooks and fantasists.

Section 3:

I suspect the above three topics (perspectives, inner conflict, and phenomenological closeness) do not exhaust the phenomenological data that could be adduced to help us make sense of the composition of subjects (especially we turn to consider unusual cases like split-brain patients). But I think they are enough to show that we do have data to work with. In cases of inner conflict, we seem to experience a unitary subject moving closer to decomposition; in cases of great

phenomenological closeness, we seem to experience distinct subjects moving closer to fusion. And in the notion of a perspective we have a way to translate at least much of our talk of subjects into a format that does admit, unmysteriously, of composition and decomposition.

The remarks I wish to add in this last section are extremely speculative. They are not intended as the putting-forward of a theory, but as explorations of what a theory might look like, if it sought to make use of the kinds of phenomena discussed above. In particular, they are attempts to find points at which to attach mental notions to physical notions, not reductively, but so as to give us some guidance as to where in the physical world different sorts of experiences, and different sorts of experiential subjects, are to be found.

One very basic question is whether to identify subjects with persisting physical *things* (fundamental particles, or something a bit larger), or with physical *processes*. That is, should a human subject be identified with 'a human brain', or with 'human brain activity'? This will determine what sort of 'parts' we look for in it – parts corresponding to 'parts of the brain', such as the left frontal lobe, or parts corresponding to 'patterns within the larger pattern'. My inclination is to suspect that the latter is more appropriate, if only to account for the apparently very basic *temporality* of consciousness. But I am also inclined to think the dichotomy may to some extent be a false one: the brain as a whole, and all of its cellular components, rapidly recycle most of their *matter*, but appear to retain some sort of identity nevertheless, and so they themselves might be best called 'patterns'.

A second question is about the quantitative aspect of consciousness: someone feeling a pleasure and a displeasure at once might be overall either pleased or displeased, according to the relative 'strengths' or 'intensities' of the two states. It seems that the overall nature of our conscious state is somewhat, but not exactly, the sum of its parts – that is, someone feeling two pains will be, by and large, 'more pained' overall than someone feeling only one of them, but the phrase 'by and large' covers up a wide range of more complex interactions the pains might have – one might serve to

distract from the other, or else the two together might produce feelings of helplessness or confusion that intensified both.

One hypothesis we might then entertain is that when two subjects of experience 'merge' the resultant subject's experiences will be roughly as intense as their respective experiences 'put together', but also greater or lesser insofar as those two sets of experiences 'interfered with' or 'reinforced' each other, perhaps conceived in some very rough analogy with the sort of constructive and destructive interference displayed by waves super-posed on each other.

There might then be both subjects which were, quantitatively, 'huge' in the number of other subjects they incorporated, but nevertheless had only very weak experiences because their multitudinous parts were uncoordinated and so their experiences mainly 'cancelled out'. Conversely, other subjects, much 'smaller' in absolute terms, might have much more intense experiences because they coordinated their parts in the right way, and moreover as a result their visible behaviour better reflected the structure of consciousness, rather than being the outcome of a 'mechanical' adding of simple units. This would vindicate our intuitive assumption that animals with brains, despite their relative smallness, might have more intense states of consciousness than things like mountains or moons.

When a pair of experiences 'reinforces', to become 'more than the sum of its parts', how is that different, phenomenologically, from cases where they 'interfere' or 'cancel out', and become 'less than the sum of its parts'? Here we might have recourse to (some version of) the notion of 'identification', which connects to my previous discussion in two ways.

Firstly, experiences of increasing phenomenological closeness may have either positive or negative valence (and so may be fiercely resisted or ardently pursued), and identification seems to often play a role in determining which: closeness with something (or someone) we do not identify with is likely to be frightening, whereas closeness with something (or someone) we do identify with is more

prone to be accepted or welcomed. Secondly, identification seems to be linked with perspective taking: when we identify with someone, that usually means we are prone to 'take their perspective' in some sense. At the extreme, someone whose perspective we are aware of in a wholly 'opaque' way, as a thing that exists but not as a perspective we can take, would seem to be paradigmatically 'alien' and 'other'.

To develop these thoughts further would require its own discussion, in particular to clarify the various different senses of 'identify with'. My suggestion here is just that, in some vague sense of that word, subjects which identify with each other might be expected to combine 'harmoniously', into something 'more than the sum of its parts', while those which do not might be expected to combine 'violently', diminishing each. But as I said, these thoughts are extremely speculative.

Conclusions:

My conclusion is that the combination problem is not as hard as the 'hard problem' of consciousness in Chalmers' sense[20]: it is a problem that is hard to solve but which we have every reason to anticipate solving, rather than one which we cannot even imagine solving. We have conceptual resources that can be applied, and I am optimistic that with sufficient intellectual labour, it can approach as near to resolution as philosophical problems ever do.

Of course I have not attempted to solve it here, and it is still quite possible that it will be solved only at some cost to common-sense assumptions about subjectivity. But then, of course, panpsychism is already a serious blow to common-sense. It will be important to distinguish as clearly as possible between the sort of common-sense that we ought to hang tenaciously onto, and the sort that we might justly regard as mere prejudice to be swept away by intellectual progress.

---

[20] See Chalmers, 1995

Rough Draft

References:

Chalmers, D., 1995, 'Facing Up to the Problem of Consciousness', Journal of Consciousness Studies 2 (3), 1995, pp. 200-219

Coleman, A., 2006, 'Being Realistic: Why Physicalism May Imply Panexperientialism', in 'Consciousness and its Place in Nature; Does Physicalism Entail Panpsychism?', ed. Freeman, A., Imprint Academic

Dennett, D. C, 1991, 'Consciousness Explained', Boston and New York: Little, Brown

Descartes, R., 1641, 'Meditations on First Philosophy', trans. Bennett, J., 2004,

Goff, P., 2006, 'Experience Don't Sum', in 'Consciousness and its Place in Nature; Does Physicalism Entail Panpsychism?', ed. Freeman, A., Imprint Academic

    2009, 'Can the Panpsychist Get Around the Combination Problem?', in 'Mind that Abides: Panpsychism in the New Millenium' ed. Skrbina, D., John Benjamins Publishing Company

    2010, 'There is No Combination Problem',
http://herts.academia.edu/PhilipGoff/Papers/327429/There_is_No_Combination_Problem

    2011, 'Property Dualists Should be Panpsychists', Online Consciousness Conference

http://consciousnessonline.wordpress.com/2011/02/16/property-dualists-should-be-panpsychists/

Freud, Sigmund, 1920, Beyond the Pleasure Principle; Trans. by C. J. M. Hubback. London, Vienna: International Psycho-Analytical, 1922; Bartleby.com, 2010. www.bartleby.com/276

James, W., 1890, 'The Principles of Psychology', vol. 1, Henry Holt and Co. Reprinted by Dover books, 1950.

Leibniz, G., 1686-7, Correspondence with Arnauld, A., trans. Bennett, J., 2004

Lovejoy, A., 1927, 'The Meanings of 'Emergence' and Its Modes', quoted in Van Cleve, J., 1990, 'Mind-Dust or Magic? Panpsychism Versus Emergence', Philosophical Perspectives, pp.215-226

Merricks, T., 1998, 'Against the doctrine of Microphysical Supervenience', Mind 107, 59-71.

Nagel, 1986, 'The View From Nowhere', Oxford University Press.

Plato, "The Republic", trans Lee, H. D. P., (1955, 1974), Penguin Books Ltd.

Rosenberg, G., 2004, 'A Place for Consciousness: Probing the Deep Structure of the Natural World', Oxford : Oxford University Press

Seager, W., 1995, 'Consciousness, Information, and Panpsychism', Journal of Consciousness Studies, 2, pp. 272-88.

    2010, 'Panpsychism, Aggregation, and Combinatorial Infusion', Mind & Matter Vol. 8(2), pp. 167-184

Sider, T., 2001, 'Maximality and Intrinsic Properties', Philosophy and Phenomenological Research 63, 357364.

Strawson, G., 2009, 'Realistic Monism: Why Physicalism Entails Panpsychism' in 'Mind that Abides', ed. Skrbina D., John Benjamins Publishing Company

Unger, P., 1980, 'The Problem of the Many', Midwest Studies in Philosophy V5, pp.411-467