Chapter 8 – The multiple location hypothesis

Contents

[8.1 The basic idea 1](#_Toc329948113)

[8.2 The metaphysics of multiple location 2](#_Toc329948114)

[8.3 Causal manners of location 3](#_Toc329948115)

[8.4 How to look for the neural correlates of multiple location 5](#_Toc329948116)

[8.5 Phenomenological concerns with the multiple location hypothesis 7](#_Toc329948117)

[8.6 Non-compositional panpsychism and the combination problem 11](#_Toc329948118)

[8.7 Ethical implications of non-compositional panpsychism 12](#_Toc329948119)

[8.8 Causal emergentism versus non-compositional panpsychism 13](#_Toc329948120)

[Appendix – Sceptical scenarios and persistence through time – A perfect post-Galilean argument with a sad conclusion 13](#_Toc329948121)

# 8.1 The basic idea

Micro-level causal closure entails that only micro-level entities and their properties have irreducible causal impact on the happenings of the world. Flourishing in the world we take to be real requires that the mind and its properties have irreducible causal impact on the world. So why not identify the mind with a micro-level entity?

There seem to be two obvious difficulties with this proposal:

1. It seems like we’re going to have to think there’s some special privileged bit of the brain where the mind is located. Just as Descartes thought the pineal gland was the special bit of the brain where the mind interacted with the body, so we’re going to have to think there’s some special bit of the brain where the mind is located. This just doesn’t seem to be borne out by the empirical data; there doesn’t seem to be some special bit of the brain where it all comes together.
2. If my mind has a single, sub-atomic location in my brain, it doesn’t look like it’s going to have much effect on macroscopic behaviour. We want to think that my mind is causing me to write the words I’m writing now, my understanding of jokes causes me to laugh, my feelings of pain cause me to scream and run away. But if my mind has only a small, sub-atomic location, it’s not going to be able to have this kind of significant impact on macroscopic behaviour.

In fact, both of these problems follow not from the thesis that the mind has *a* very small location in the brain, but from the thesis that it has *a single* very small location in the brain. We can avoid both these problems if we can make sense of the mind having *many* small locations in the brain, of the mind being multiply located many times in the brain. Call this ‘the multiple location hypothesis.’

There are two analogies that might help here: catholic saints and time travellers. According to the Catholic faith, bi-location is a fairly common miracle amongst saintly individuals. We have stories of one individual being present in two places at the same time, giving mass in San Lucia whilst simultaneously helping the poor in Calcutta. These are alleged cases of multiple location: one individual being wholly present in two locations at once. Or suppose you go back in time to have a coffee with yourself ten years ago. Again, we have a case of a single individual wholly located at either side of the table.[[1]](#footnote-1) Even though it strikes us as bizarre, multiple location seems to be coherent. It may be counter to common sense, but common sense has no place in serious metaphysics. It is worth considering, then, the hypothesis that the mind, a single entity, is wholly present at many distinct locations in my brain.

This multiple location hypothesis gets round the two problems considered above. With regards to problem A, there need not be a special place at which the mind is located; the mind may be located many times throughout a large region of the brain. This large region is presumably the region a more conventional mind-brain identity theorist would want to locate the mind. Indeed, the proponent of the multiple location hypothesis can agree with the mind-brain identity theorist concerning the macroscopic location of the mind, whilst holding that the mind has that location derivatively, in virtue of being multiply located at many parts of that region.

Regarding problem B, although the mind cannot have much impact on behaviour if it’s only located at a single micro-level region of the brain, it can do if it’s located at many micro-level regions of the brain. Returning to the analogies, imagine a Catholic saint finds he is unable to lift up a heavy table. Solution: he multiply locates himself six times, and hence increases his lifting power by a factor of six. Or suppose you are a time traveller too weak to fight your enemy, and are having difficulty enlisting troops. Solution: travel back in time and gather an army of your former selves.

It might look from the outside like there are six people lifting the table, but in reality there is just the one saint located six times doing the lifting. Similarly, it might look from the outside like many distinct micro-level entities are acting in concert together to govern my behaviour. But on the multiple location hypothesis, it is *one* entity, located many times, that is governing my behaviour.

# 8.2 The metaphysics of multiple location

We can account for the multiple location of minds by adopting the following three theses:

*Thesis 1 –* Aristotelian realism about phenomenal qualities.

*Thesis 2 –* Substantivalism about space

*Thesis 3 –* Bundle theory of subjects of experience.

Let us take each of these in turn.

*Thesis 1*

Phenomenal qualities are in res universals. That is, a given phenomenal quality is wholly present in each region of space and time at which it is located: if the exact same shade of phenomenal red is located at L1 and at L2, then we have one thing – a specific shade of phenomenal red – wholly present at L1 and wholly present at L2. In other words, phenomenal qualities are multiply located.

*Thesis 2*

Space is a particular object in its own right, made up of (or perhaps grounding) regions of space which are themselves particular objects.

*Thesis 3*

Phenomenal universals are fully saturated beings, not standing in need of support from a substratum. Where a number of phenomenal universals U1, U2,U3…UN are co-located at a location L, U1, U2,U3…UN constitute a subject of experience located at L.

In conjunction these three theses entail the multiple location of subjects of experience. Suppose phenomenal qualities U1, U2,U3…UN are wholly located at L1 and wholly located at L2. At L1, U1, U2,U3…UN constitute a subject of experience, and at L2 U1, U2,U3…UN constitute a subject of experience. Given that the subject of experience at L1 just is U1, U2,U3…UN, and the subject of experience at L2 just is U1, U2,U3…UN, it follows that the subject of experience at L1 is numerically identical with the subject of experience at L2. That subject is wholly present at L1 and wholly present at L2.

To get vivid idea of the fundamental metaphysics being proposed here, it might be useful to think of space as a television screen, and universals as the pixels at various regions of the screen. At any given moment, which universals (pixels) are located at which regions (parts of the screen), is determined by the causal influence of the universals located at the previous moment. Facts about the continuity of objects over time are made true by facts about which universals (pixels) are located at which regions (screen).

# 8.3 Causal manners of location

If my micro-located mind is to have the kind of causal powers necessary to avoid sceptical scenarios, i.e. if it is to cause me to speak and act in the way we ordinarily suppose that my mind does, then it must be located trillions of times over a very large area of the brain. Let us refer to these locations at a given time T as ‘the Ls’. Call the phenomenal qualities had by my mind at T ‘the Qs’. For my mind to be currently multiply located at each of the Ls, all of the Qs must be located at each of the Ls. For we are supposing that my mind just is the Qs, hence, if there is a location L\* at which one of the Qs is not located, then my mind is not wholly present at L\*.

This leads to a prima facie problem. From the outside it will look like there are distinct micro-level objects located at each of the Ls, not because (or not only because) there are a number of distinct locations are in play, but also because different causal influence is being exerted from different Ls. It is likely that at there are least two of the Ls, L1 and L2, such that the causal influence exerted from L1 is different from the causal influence exerted from L2. Perhaps the causal influence exerted from L1 will be of a kind we associate with positive charge, whilst the causal influence exerted from L2 will be of a kind we associated with negative charge. From the outside it will look as though there is a positively charged particle located at L1 and a negatively charged particle at L2. And yet what we in reality have, according to the multiple location hypothesis, is one bundle of universals, located at L1 and L2. How can one thing have distinct causal powers at different locations?

We need to make sense of the causal powers of a given bundle of universals varying from location to location. I propose that the causal powers of a given bundle of universals are determined by the bundle’s *manner of location*. Whilst we ordinarily take location to be a two-place relation between an object (in this case a bundle of phenomenal universals) and a region of space, I suggest we can instead take it to be a three place relation, between an object, a region of space, and a *causal manner of location*.

We can think of causal manners of location somewhat analogously to the brute causal powers of dispositional essentialists. We might further suppose that they are the properties tracked by physics. Thus we can suppose there is a *negative-charge* manner of location, such that a given bundle of universals located negative-charge-ly has the causal powers we associate with negative charge. And we can suppose that there is a *N-kgs-of-mass* manner of location, such that a bundle of universals located N-kgs-of-mass-ly has the causal powers we associate with having N kgs of mass.

In virtue of being located in a certain location and in a certain causal manner, a given bundle of universals has the power to affect both the location and the manner of location of certain other bundles of universals. If my mind is located at L1 in the positive-charge manner, then at L1 my mind will exert the kind of causal influence we associated with positive charge. If my mind is located at L2 in the negative-charge manner, then at L2 my mind will exert the kind of causal influence we associate with negative charge. From the outside it will look like there are as many kinds of particles as there are causal manners. But in reality, the same universal or universals might be located in one causal manner at one location, and in another causal manner ay another location. Conversely, a different universal or universals might be located in the same causal manner at a number of distinct locations.

The upshot is that my mind exerts a different kind of causal influence at different regions in virtue of the manner of its location at those regions. By being located at trillions of locations in a number of distinct manners of location, my mind – the Qs – causes neurological changes which make laugh at jokes, run away when I’m in pain, etc.

There is no contradiction here. It is not that the mind both has negative charge type causal powers and at the same time lacks negative charge type causal powers. Rather the mind has negative charge type causal powers *at location X*, but lacks negative charge type causal powers *at location Y*. Nor does this render the Qs epiphenomenal. The Qs have causal powers, and cause things in virtue of their causal powers. It is just that that the Qs have different causal powers at different locations, and their overall causal influence on the world will vary over time.

There are many metaphysical views according to which the causal powers of the Qs vary from one possible world to another, dependent on the contingent laws of nature that happen to obtain in a given world. Nobody thinks that the fact that the causal powers of the Qs varies from world to world renders the Qs causally impotent. And so it seems coherent to suppose that the causal powers of the Qs might vary within a world without this rendering them causally impotent. I suspect it is a doctrine of common sense that a given property cannot vary its causal powers with a world, but as I have already noted, common sense has no place in serious metaphysics.

# 8.4 How to look for the neural correlates of multiple location

We have in broad brush strokes a picture of a world in which multiple location is possible. To fill in the details we have to work out what the laws of nature must be like such that, in the right neurological conditions, consciousness of the right kind comes to be located in enough micro-level regions to govern the behaviour we pre-theoretically take to be governed by consciousness. This is a largely empirical project, working out which processes in the brain are responsible for consciousness-governed behaviour. But it shaped by a non-empirical constraint: in order to avoid sceptical scenarios, it must be the case that our pre-theoretical beliefs about the causal impact of consciousness are largely correct.

Of course, there is some scope for revision. There is a plenty of evidence that consciousness has a more limited role that we might imagine. But the need to avoid sceptical scenarios puts a limit on the possibility of revision. It cannot be the case that my agony yesterday had no causal role in my going to the doctors today, or that my thoughts never cause my words. Indeed, the very possibility of scientifically investigating consciousness relies on such anti-sceptical assumptions. We can learn about the consciousness of another only by assuming that the connection between their consciousness and their behaviour is more or less how we pre-theoretically take it to be. Let us call this the ‘causal constraint’.

There is a further constraint, which David Chalmers calls the ‘principle of structural coherence.’ Chalmers’ has argued that there is a systematic correlation between the structure of consciousness and the structure of *awareness*, where the latter is a purely functional notion, defined as ‘the contents of awareness are to be understood as those information contents that are accessible to central systems, and brought to bear in a widespread way in the control of behavior’ (p. 17):[[2]](#footnote-2)

It is a central fact about experience that it has a complex structure. The visual field has a complex geometry, for instance. There are also relations of similarity and difference between experiences, and relations in such things as relative intensity. Every subject’s experience can be at least partly characterized and decomposed in terms of these structural properties: similarity and difference relations, perceived location, relative intensity, geometric structure, and so on. It is also a central fact that to each of these structural features, there is a corresponding feature in the information processing structure of awareness.

Take color sensations as an example. For every distinction between color experiences, there is a corresponding distinction in processing. The different phenomenal colors that we experience form a complex three-dimensional space, varying in hue, saturation, and intensity. The properties of this space can be recovered from information-processing considerations: examination of the visual systems shows that waveforms of light are discriminated and analyzed along three different axes, and it is this three-dimensional information that is relevant to later processing. The three-dimensional structure of phenomenal color space therefore corresponds directly to the three dimensional structure of visual awareness. This is precisely what we would expect. After all, every color distinction corresponds to some reportable information, and therefore to a distinction that is represented in the structure of processing (p. 18).

Arguably the principle of structural coherence is also an essential commitment for avoiding sceptical scenarios. If an individual is able to report that their experience has such and such a structure, then information (in a purely causal sense of information) about such structure must be available for verbal report. And in so far as we accept that an individual’s reports of the structure of his/her experience are accurate, that information about structure which is available for verbal report must correspond to the structure of the individual’s consciousness.

Let us think about what is required for the principle of structural coherence to be respected given the multiple location hypothesis. A given individual’s consciousness, by being multiply located at very many distinct micro-level locations, constitutes that individual’s awareness. The structure of the consciousness which is located at each of those micro-level locations mirrors the structure of awareness at the macro-level. This correlation requires explanation: either the structure of consciousness at each micro-level location is somehow grounding the structure of awareness at the macro-level, or the structure of awareness at the macro-level is somehow grounding the structure of consciousness at each micro-level location.

It must, it seems, be the structure of awareness that is impacting on the structure of consciousness, rather than the other way round. Although the individual’s conscious mind, in virtue of its various locations and its various causal powers at those locations, constitutes the individual’s awareness, the mind’s causal powers are determined by its causal manners of location rather than its intrinsic categorical nature. It is the intrinsic categorical nature that is *doing the causing*, in virtue of its causal powers. But it has the causal powers it has in virtue of its causal manners of location. It cannot be, then, that awareness has the structure it does because consciousness has the structure it does. The explanation must go the other way round: consciousness has the structure it does because awareness has the structure it does.

Somehow, therefore, (i) a large number of micro-level components of the brain (those constituting awareness) come to have indiscernible conscious experience, (ii) the structure of that experience is determined by structure of the higher-level state they constitute (i.e. awareness). The proponent of the multiple location hypothesis has the semi-speculative/semi-empirical task of working how the laws of nature must be such that this comes to be the case in those brains that we must take to instantiate consciousness in order to avoid sceptical scenarios. What is being outlined here, then, is not so much a theory as a research project.

# 8.5 Phenomenological concerns with the multiple location hypothesis

In its commitment to bundle theory, the multiple location hypothesis is potentially subject to challenge on phenomenological grounds. Recall our re-interpretation of C. B. Martin’s anti-bundle theory intuition from chapter 5:

A particular phenomenal quality has to be *of* something…what is referred to as the ‘feeling of pain’ *cannot be* thought of under some other description as an object that could have existed without need of being the painful feeling of anything but as an object existing in its own right.

In so far as one’s careful reflections on the nature of consciousness lead one to the conclusion that phenomenal qualities are unsaturated entities, this looks to count against the multiple location hypothesis as I have described it, according to which phenomenal qualities are bundles of saturated universals.

In my experience, many philosophers I would be inclined to count as honory post-Galileans are in some sense inclined to identify subjects and phenomenal qualities. Galen Strawson, for example, has a wonderful piece of guided meditation to the conclusion that experience, subject of experience, and content of experience are identical.[[3]](#footnote-3) For my own part, I’m rather inclined to the above Martin-esque intuition, although I’m open to persuasion.

Perhaps there is a way of circumnavigating this potential worry. Even if phenomenal qualities stand in need of support, perhaps what supports them is itself a universal, perhaps some kind of substratum-like universal which is wholly present in each experience supporting the phenomenal qualities of that experience. Acquaintance with a given phenomenal quality will reveal whether or not it stands in need of support, but not necessarily whether what supports it is a particular or a universal. A substratum that is itself a universal is no bar to the possibility of multiple location: if the Qs (the bundle of phenomenal qualities associated by my mind at a particular time) instantiated by the substratum universal are located at L1, and the Qs instantiated by the substratum universal are located at L2, then the subject of my experience will be wholly located at both L1 and L2.

Whether this modification of the view is a genuine possibility will depend on one’s view concerning the object of acquaintance. If I am acquainted only with phenomenal qualities themselves, then there seems no reason to think phenomenological reflection will reveal anything about the nature of the putative bearers of those qualities. But if I am acquainted with *the state of affairs of* *the subject of my experience* *having phenomenal qualities*, then it may be introspectively apparent that the substratum involved in that state of affairs is a particular. A post-Galilean who feels enough conviction that introspection reveals one’s phenomenal qualities to be instantiated by an unrepeatable substratum, will not be able to analyse multiple location in the way I have suggested above. Similarly, if one’s phenomenological reflections incline one to either a trope theoretic view of phenomenal qualities, or an austere nominalist conception of subjects of experience, then the account of multiple location we have been working with must be rejected.

Perhaps one could just take it to be a brute fact that the particular object that is the subject of my experience is multiply located. Call this the ‘brute multiple location hypothesis.’ After all, Catholic saints and endurantist time travellers seem to be particular things that are located at distinct spatial locations, and stories of such things seem to be coherent.[[4]](#footnote-4) However, as I shall to explain below, I’m not sure the details of the brute multiple location hypothesis can be worked out without it resulting in a horribly complex and inelegant theory.

According to the brute multiple location hypothesis, in systems realising awareness, there is one particular object (a mind), multiply located, constituting the system (all of it, or enough of it to satisfy the causal constraint). Suppose a mischievous neuro scientist takes your mind at a given location in your brain, and my mind at a given location in my brain, and swaps them round?[[5]](#footnote-5) My mind will now be involved in constituting my awareness, by being multiply located in my brain, as well as being involved in constituting your awareness, by being singly located in your brain. Conversely, your mind will now be involved in constituting your awareness, by being multiply located in your brain, as well as being involved in constituting my awareness, by being singly located in my brain (remember ‘awareness’ is a purely functional notion).

We argued in 8.4 that, in a given individual, the structure of its awareness determines the structure of the consciousness of the micro-component that – multiply located – constitutes that awareness. Hence, one might be inclined to think that the ‘instance of my mind’ that constitutes your awareness will have a consciousness structurally isomorphic with your awareness, whilst the ‘instances of my mind’ that constitute my awareness will have a consciousness structurally isomorphic with my awareness. The trouble is that that the notion of ‘an instance of my mind’ does not make sense in this context; *ex hypothesi* the mind is a particular. Perhaps a particular object can be multiply located, but it cannot be the case that a particular object has contradictory properties, which would be the case if my mind both did and didn’t have conscious experience structurally isomorphic with the your awareness.

Enduranists, those philosophers who think that a particular object is wholly present at each moment of time at which it is located, face a similar problem with regards to the contradictory properties of a single particular at different points in time. In the context of the metaphysics of temporal persistence, this is known as ‘the problem of temporary intrinsics.’ Suppose a poker was straight yesterday and bent today. Assuming eternalism, the view that all times are equally real, the poker has contradictory properties: it is both straight and bent. The endurantist solution is to index shape properties to times. The stick does not have the property of being *bent-simpliciter* or *straight-simpliciter*, rather it has the property of being *bent-yesterday* and the property of being *straight-today*. Thus the threat of contradiction is avoided.

David Lewis argued against this strategy for rescuing endurantism on the grounds that it takes shapes to be relations, rather than intrinsic properties:

 ‘[Lewis starts by describing the view he is rejecting]….contrary to what we might think, shapes are not genuine intrinsic properties. They are disguised relations, which an enduring thing may bear to times. One and the same enduing thing may bear the bent-shape relation to some times, and the straight-shape relation to others. In itself, considered apart from its relations to other things, it has no shape at all….This is simply incredible….If we know what shape is, we know it is a property [i.e. an intrinsic property] not a relation.’ (Lewis 1986: 204)

Notice Lewis’s appeal to a pre-theoretical intuition: shape is not a relation. For the post-Galilean this is not a legitimate appeal to intuition, as it at best tells us something about *Lewis’s* concept of shape, or the folk concept of shape in so far as Lewis has the same concept of shape as most other people. But why think that our folk concept of shape is satisfied? Maybe our folk concept of shape is not satisfied, but a relationalist concept of shape is satisfied. Arguably our folk concept of solidity – filling all regions of space with certain boundaries – is not satisfied, although a close enough concept – resisting penetration – is satisfied. Perhaps something similar is true with regards to the concept of shape (compare the discussion of intuitions in 5.2).

Recall in 5.2 how we reconfigured the intuitions of Armstrong and Martin to make them fit the post-Galilean party line, by making them intuitions about consciousness. If we can do the same with Lewis’s intuition about shape, then we can perhaps turn it into a legitimate source of metaphysical data. Here’s what a post-Galilean translation of Lewis would look like:

‘[again starting off describing the view under attack]….contrary to what we might think, phenomenal qualities are not genuine intrinsic properties. They are disguised relations, which an enduring thing may bear to times. One and the same enduring thing may bear the feeling-pain relation to some times, and the feeling-pleasure relation to others. In itself, considered apart from its relations to other things, a subject of experience has no phenomenal qualities at all….This is simply incredible….If we know what a phenomenal quality is, we know it is a property [i.e. an intrinsic property] not a relation.’

It is very plausible that a strong intuition about the essential nature of consciousness is expressed in this passage: phenomenal qualities are genuinely intrinsic properties of a subject, rather than relational properties a subject bears to times. The post-Galilean can use this intuition to reject eternalist endurantism as an account of the persistence through time of subjects of experience. It is contradictory to suppose that a subject feels pain yesterday and doesn’t feel pain today (something can’t feel pain and not feel pain), and we cannot avoid the contradiction by taking pain to be a relation between a subject and a time (so that the subject instantiated *pain-yesterday* but not *pain-today*) given the intrinsicality of pain.

How then do we answer the problem of temporary intrinsics? How do we avoid the contradiction that one subject is both feeling pain and not feeling pain? One option would be to follow Lewis’s solution, which is to adopt perdurantism: a subject persists through time in virtue of having distinct temporal parts at different times. One temporal part of the subject (the part that exists yesterday) feels pain; a distinct temporal part of the subject (the part that exists today) does not feel pain. Alternately, we can adopt presentism: only the present moment exists. If the state of affairs of the subject feeling pain no longer exists, then there is no threat of its contradicting the presently existing state of affairs of the subject not feeling pain.

This has an extremely interesting upshot: there is no middle way for the post-Galilean between endurantist presentism and eternalist (or growing block) perdurantism. Eternalist (or growing block) endurantism is ruled out by the intrinsicality of phenomenal qualities. In the appendix to this chapter, I will discuss certain difficulties with both of remaining options.[[6]](#footnote-6)

Returning to the brute multiple location hypothesis, recall that we seemed to end up (when my mind at one location in my brain is swapped for your mind at one location in your brain) with the contradiction that my mind located in your head has experience structurally isomorphic with your awareness, whilst my mind located in my head has experience structurally isomorphic with my awareness. We have now learned that we cannot remove this contradiction by supposing that each experience is a relation to location. And of course we cannot remove the contradiction analogously to the presentist strategy for answering the problem of temporary intrinsics (as my brain and your brain co-exist), or analogously to the perduarantist strategy (as *ex hypothesi* the mind is wholly present in each of its micro-level locations).

Therefore, in order to avoid ending of up with a contradictory theory, the proponent of the brute multiple location hypothesis must somehow ensure that it is not physically possible to take my mind at one location in my head and put it into your head. Perhaps my mind will simply lose a location once attempts are made to take it from a given location away from a system realising my awareness. The laws of nature needed to ensure this are going to result in a deeply inelegant theory.

The non-brute multiple location hypothesis avoids these difficulties because, at a fundamental level, we don’t have particular objects that can be moved about. Rather we have facts about which universals are instantiated at which locations. Recall the pixel/screen analogy from the end of 8.2. The pixels on a TV screen are not particular objects that can picked up and moved to a different bit of the screen.

Hence, the post-Galilean who is inclined to believe that properties are themselves particulars, or are instantiated by particular substrata, is going to have difficulty adopting the multiple location hypothesis. As I remarked in chapter 5, I suspect that few post-Galileans will be inclined to austere nominalism. The immediate objects of introspection seem to be concrete qualities, and hence the disjunction of trope theory and Aristotelian realism about universals seems to be the default position. I argued in chapter 5 that it must be possible in principle to choose between trope theory and Aristotelian realism about universals on phenomenological grounds, but it is very difficult to see how this might be done. A post-Galilean who feels enough conviction that introspection reveals phenomenal qualities to be tropes, or instantiated by unrepeatable substrata, may be unable to adopt the multiple location hypothesis.

So far we have discussed phenomenological concerns regarding the claim that a subject is a bundle of *universals*, but one might also have phenomenological concerns regarding the claim that the subject is a *bundle* of universals. Intuitively, the phenomenal qualities in my experience have some deeper kind of unity than mere co-location. More specifically, intuitively my overall experience grounds specific aspects of my experience, rather than vice versa.

The theory is easily modified to accommodate this intuition. The intuition stands in opposition to the view that, at a fundamental level, at any occupied region of space, there are a number of phenomenal qualities that just happened to all be located there. The intuition calls for some metaphysical glue binding the individual qualities into a unity. We can provide this glue by offering the following more fundamental story: at any occupied region of space, there is a single determinate state of consciousness, such that that single determinate state of consciousness grounds a number of phenomenal qualities. The distribution of determinate states of consciousness across space at any given moment will be determined by the distribution of determinate states of consciousness across space at the immediately prior moment. In turn, the presence of a given determinate state of consciousness at a given location will ground the presence of certain phenomenal qualities at that location.

It seems to me, then, that the only significant phenomenological concern with the multiple location hypothesis is that post-Galileans whose reflections lead them to think that phenomenal qualities are tropes, or universals instantiated by an unrepeatable substrata, will have to suffer the inevitable inelegance of the brute multiple location hypothesis. I have so far not been able to find in introspection a way of decide between trope theory and realism about universals, or between the view that we are acquainted only with phenomenal qualities and the view that we are acquainted with the state of affairs of my mind (or my mind’s substratum) instantiating phenomenal qualities, although my overall method commits me to thinking there must be some way to decide these issues through introspection. I leave it to future work to decide whether introspection is for or against the multiple location hypothesis.

# 8.6 Non-compositional panpsychism and the combination problem

Panpsychists believe that the fundamental nature of reality is mental. Most forms of panpsychism in the literature construe it as a layered view of the world, with particles with very simple consciousness constituting brains with much more complex consciousness. These standard forms of panpsychism face what is commonly known as ‘the combination problem’. There are different interpretations of what exactly the combination problem amounts do, but the basic issue is that it is hard to make sense of ‘little’ conscious things coming together to make ‘big’ conscious things. The inspiration for the combination problem is the following passage from William James:

Take a hundred of them [feelings], shuffle them and pack them as close together as you can (whatever that may mean); still each remains the same feelings it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean. There would be a hundred-and-first-feeling there, if, when a group or series of such feelings where set up, a consciousness *belonging to the group as such* should emerge. And this 101st feeling would be a totally new fact; the 100 feelings might, by a curious physical law, be a signal for its *creation*, when they came together; but they would have no substantial identity with it, not it with them, and one could never deduce the one from the others, nor (in any intelligible sense) say that they *evolved* it (James 1983, 162).

If, as I argued in chapter 6, consciousness is an irreducible, simple, saturated, form of being, then the combination problem looks to be insoluble.

The multiple location hypothesis is clearly not a standard form of panpsychism, as it is not a layered picture of the world: minds are micro-level entities. Still, if the only categorical universals involved in this picture of the world are phenomenal universals, then the multiple location hypothesis will count as a form of panpsychism. This form of panpsychism would avoid the combination problem, as there is no combination. Let us call a panpsychist form of the multiple location hypothesis ‘non-compositional panpsychism.’

Non-compositional panpsychism also has the advantage of offering us a completely transparent conception of the world, except in so far as there are phenomenal qualities that we don’t understand because they don’t constitute human conscious experience. Suppose my arguments from chapter 6 were wrong, and phenomenal, or proto-phenomenal qualities, do somehow combine to constitute our experience. Still, the prospects for our forming a transparent conception of such combination look pretty bleak. One is immediately acquainted only with one’s own conscious experience, which presents itself to introspection as a simple, underived quality. If phenomenal/proto-phenomenal combination does occur, it’s likely that perceiving it, or even conceiving it, is physically impossible for a human being.

Thus, non-composition panpsychism is an attractive view, even if we are not forced in its direction by worries about causal closure. The only properties we know to exist in the world, the only properties we really understand the nature of, are phenomenal qualities. I can see no reason to postulate strange, unknowable features of the world, if the phenomenal qualities we know and love can serve as the categorical bases of the dispositional properties we know and love from physics. The more I think about, the more I’m inclined to think non-composition panpsychism has a real shot at being true.

# 8.7 Ethical implications of non-compositional panpsychism

If non-compositional panpsychism is true, we overlap. If you and I have a qualitatively indiscernible pain, then a bit of me is numerically identical with a bit of you. The more our thoughts and feelings come to resemble, the less it will be true to say that we are distinct individuals. If my conscious experience should one day become indiscernible with your conscious experience, we would be literally one and the same subject of experience.

This potentially had immense ethical implications. The supposed rational basis for my acting selfishly, for acting on maxims that value the subject of experience located in my head above the subjects of experience located in the heads of others, is premised on the belief that the subject of experience located in my head is wholly distinct from the subjects of experience located in the heads of others. If non-compositional panpsychism is true, this premise is false, and hence the rational basis for selfishness undermined.

# 8.8 Causal emergentism versus non-compositional panpsychism

I finish, then, with two kinds of view: non-compositional panpsychism and causal emergentism. If micro-level causal closure is true then we ought to go for non-compositional panpsychism. If it turns out that there are emergent causes which might plausibly be identified with mental events, then we should work towards a causal emergentist view. As has already been noted, these are not so much two complete theories as two research projects.

The emergentist project is more empirically driven, as we are beholden to the empirical facts concerning emergent causal forces. The non-compositional panpsychist project is more speculative, given that causal closure ensures that mental events are in a certain sense not empirically distinguished. For the non-compositional panpsychist, where we decide to locate consciousness in the brain will be determined by pre-theoretical commitments concerning the relationship between consciousness and behaviour, together with the empirical facts concerning which neural events govern the behaviour we pre-theoretically take to be governed by consciousness.

I leave it to empirically inclined philosophers and philosophically inclined scientists to settle the question of causal closure, and hence decide which of these two projects is to be pursued.

# Appendix – Sceptical scenarios and persistence through time – A perfect post-Galilean argument with a sad conclusion

We decided in 8.5 that the post-Galilean must choose between presentist enduranism and eternalist (or growing block) perdurantism. In this appendix I will raise concerns with both of these options.

Let us first consider the perdurantist option. In 8.5 I described the perdurantist as believing that subjects persist through time in virtue of having temporal parts located at different times. Actually, how we describe the view depends on whether we define subjects as being the direct bearers of consciousness, or whether we define subjects as bearing consciousness derivatively. Call subjects on the first definition ‘direct subjects’ and subjects on the second definition ‘derivative subjects.’

On the perdurantist view, it is only derivative subjects that persist through time, in virtue of having direct subjects as temporal parts. A derivative subject is ‘conscious’ only in the derivative sense that it has temporal parts which are direct subjects. If a direct subject persisted though time, or least through time long enough to directly bear contradictory phenomenal qualities, then we would be back to the problem of temporary intrinsics, with no way of resolving the problem. We avoid the problem as the temporary intrinsics for derivative subjects, as they don’t directly instantiate contradictory phenomenal qualities, e.g. feeling pain and not feeling pain, rather they instantiate consistent properties defined in terms of the phenomenal qualities of its parts, e.g. having a temporal part that feels pain and a temporal part that doesn’t feel pain.

This is arguably a sceptical scenario. The thing that directly instantiates my consciousness, or the thing that directly instantiates my partner’s consciousness, does not persist through time. Of course the perdurantist will take it to be true that ‘My partner existed yesterday and existed today’, made true by the existence of an entity that persists through time in virtue of being composed of temporal parts that are direct subjects. However, as with the argument at 7.3.4, I am asking about what is required of the World in order for this to be true in a way that preserves what is important. Doing this requires deep and serious reflection on the reality of human interaction. You need to put the book for a bit.

Suppose you partner felt anger yesterday and feels joy today. Reflect on your relief that your partner doesn’t feel anger anymore. The point is that, on the perduarantist view, the thing that directly felt anger yesterday doesn’t exist anymore. There is something here today that has some peculiar property of being composed of something that directly feels anger yesterday. But this peculiar property is surely not one that one has any relevance in the psychological ascriptions we apply to others, and in terms of which we understand the relationships which constitute the value of life.

Similarly, on the perduarantist view, the direct subject of my experience will not exist in the future. Of course, the perduarantist has an understanding of the sentence ‘I will exist in an hour’, such that it comes out true. By the ‘I’ in that sentence does not refer to the subject of my experience, the thing that non-derivatively has my conscious experience. Rather it refers to some strange entity that is composed of things that non-derivatively have my conscious experience. Is it any consolation that that strange entity will exist in the future? It still seems a terrifying thought that the direct subject of my experience, the thing I am certain of the existence of, will not exist then. I think therefore I am, but not for long.

A final example: I am currently not feeling pleasure, but the knowledge that I am about to be tortured is terrifying me. My concern is surely that the direct subject of my experience, *that thing that is currently non-derivatively feeling pleasure*, will soon be non-derivatively feeling intense pain (people are not ordinarily concerned about the fate of spacetime worms composed of direct subjects). But on the perdurantist view this is just false: the thing that is currently non-derivatively feeling pleasure won’t exist in a moment.

I believe on the basis of these examples that perdurantism amounts to a sceptical scenario, and that this is a strong reason not to believe it. In much the same way, I have strong reason to believe that I’m a brain in a vat.

On the other hand, there are arguably empirical strong empirical grounds for doubting presentism: it seems to be inconsistent with special relativity. Call the set of events which are in the present ‘the present set’. All members of the present set are simultaneous with each other, given that they all obtain at the same time, i.e. the present time. If presentism is true, all and only events that exist are members of the present set. Assuming it is a non-relative fact which events exist, it must also be a non-relative fact which events are members of the present set, and hence a non-relative fact that certain events, i.e. the members of the present set, are simultaneous with each other. Presentism, therefore, seems to imply that there are certain events such that it is a non-relative fact that those events are simultaneous. But special relativity denies this: which events are in the present set, and in general which events are simultaneous, is relative to a frame of reference.

Some presentists have denied special relativity in order to allow for the truth of presentism.[[7]](#footnote-7) This is not as crazy as it sounds, given that the Lorenzian theory which preceded special relativity, and which allows for an objective present, is empirically equivalent to special relativity. But Einstein’s interpretation of the empirical datum in question, that the speed of light appears to be the same in all frames of reference, is much more simple and elegant that Lorenz’s. Whereas Lorenz gives an ad hoc explanation of why the speed of light appears to be the same in all frames of reference, Einstein’s explanation of why the speed to light appears to be the same in all frame’s of reference is beautifully simple: the speed of light appears to be the same in all frames of reference because it is the same in all frames of reference.

When one lacks evidence to the contrary, I think it reasonable to assume that sceptical scenarios are false. But if we do have evidence to the contrary, as we seem to in this case, then the epistemic judge will not forgive us for indulging our desire for the world to be as we want to it to be rather than the way we have reason to think it is. I am inclined to think, then, that perdurantism amounts to a sceptical scenario, but it is nonetheless one which we have the misfortune of having reason to believe.

Matters would be different if we could adopt eternalist endurantism about subjects of experience. Eternalism, in not requiring an objective present, is consistent with special relativity. Endurantism about subjects allows us to avoid the sceptical scenario in which subjects do not persist through time. But, as we discussed in 8.5, the combination of eternalism and endurantism about subjects leads to subjects having contradictory properties, and the contradiction cannot be avoided by taking phenomenal qualities to be relations to times, given the intrinsicality of phenomenal qualities.

We have here a perfect example of a post-Galilean argument, combining an empirical consideration (the best explanation of the fact that the speed of light appears to be the same in all frames of reference is that it is the same in all frames of reference) and an intuition concerning the essential nature of consciousness (phenomenal qualities are intrinsic properties). The conclusion of this argument makes me very sad, but I am learning to live with it.

1. In the time traveller case, we only have multiple location, in the sense of one entity being *wholly* located in two distinct locations, only on an endurantist model of temporal persistence. On a perdurantist model, one temporal part of yourself is sat having coffee opposite another temporal part of yourself. We will return to these issues in 8.5. [↑](#footnote-ref-1)
2. Reference [↑](#footnote-ref-2)
3. Reference [↑](#footnote-ref-3)
4. What is the difference between a trope that is multiply located and a universal? Universals are essentially non-particular. If phenomenal qualities turn out to be universals, then there is no possible world in which there are two distinct but qualitatively indiscernible pains. If phenomenal qualities turn out to be tropes, then there are worlds in which there are two distinct but qualitatively indiscernible tropes, but there may also be worlds in which a single particular trope is, as a matter of brute fact, multiply located. [↑](#footnote-ref-4)
5. I am grateful to Emma Bullock for making me think in detail about this possibility. [↑](#footnote-ref-5)
6. Another interesting upshot is that Catholic saints do end up being of dubious coherence after all. For it looks like the saint’s having one kind of conscious experience at one location is inconsistent with her having another kind of conscious experience at another location. And given that phenomenal qualities are intrinsic, we cannot avoid the contradiction by taking each conscious experience to be a relation to a location. Perhaps we can say that the saint has a single conscious experience that incorporates her experience at both locations. [↑](#footnote-ref-6)
7. I had an amusing conversation with John Bigelow, in which he said of special relativity, ‘I can’t wait for them to get rid of that theory’. [↑](#footnote-ref-7)