Chapter 6 – The Irreducibility of Consciousness

Contents

[6. 1The world as it is in and of itself 1](#_Toc329778621)

[6.2 Fundamentality and irreducibility 3](#_Toc329778622)

[6.3 Against deflationary forms of Russellian monism 4](#_Toc329778623)

[6.3.1 More kinds of concept: Irreducible, essentially reducible and reductively neutral 4](#_Toc329778624)

[6.3.2 Subjects of experience are irreducible 5](#_Toc329778625)

[6.4 Inflationary forms of Russellian monism 8](#_Toc329778626)

[6.4.1 Causal and non-causal grounding 8](#_Toc329778627)

[6.4.2 Nothing over and above definition of non-causal grounding 8](#_Toc329778628)

[6.4.3 Partist definition of non-causal grounding 9](#_Toc329778629)

[6.4.4 Holist definition of non-causal grounding 13](#_Toc329778630)

[6.4.5 Realisation definition of non-causal grounding 15](#_Toc329778631)

[6.4.6 Necessitarian definition of non-causal grounding 16](#_Toc329778632)

# 6. 1The world as it is in and of itself

There is the way we carve up the world, and there is the way the world is carved up in and of itself. The best way to make this distinction clear is with examples.

George Berkeley thought that the world in and of itself was such that there are minds, ideas, and nothing else.[[1]](#footnote-1) God, the all-powerful mind, causes us dependent minds to have the kind of systematic and regular ideas we know from waking life. We are also able to cause ideas from our own imagination. My ‘sensory’ ideas and my ‘imaginative’ ideas are to be distinguished by their cause: the former are caused by God, the latter are caused by me.

But Berkeley was insistent that in the world as he took it to be there were tables, rocks, planets, bodies, rivers, and all the other things we ordinarily take there to be ‘out there’. This position is coherent so long as our thought and talk of such things can be made true by facts that Berkeley took to constitute the world as it is in and of itself. If I use the sentence ‘There are tables’ in such a way that it is true if God causes table ideas, and God does indeed cause table ideas, then the sentence ‘There are tables’ is true, from which it follows trivially that there are tables. In such a case, the world is *truly describable* such that there are tables, even though the world *in and of itself* is not such that there are tables.

Another example: David Lewis believed that the world in and of itself was such that there are properties instantiated at point-sized regions of space and time.[[2]](#footnote-2) Like Berkeley, he was insistent that there were tables, rocks, planets, bodies, rivers, etc., but wanted to say that these things were ‘nothing over and above’ facts about the microscopic mosaic. Lewis seemed to think that the supervenience of macroscopic facts about tables and bodies on the microscopic facts sufficed to make tables and bodies ‘nothing over and above’ the microscopic facts. But a mere modal notion like supervenience cannot distinguish the view that macroscopic bodies are nothing over and above the microscopic facts, from emergentism about macroscopic bodies. There may be a brutely necessary law of the universe such that, when particles are arranged in such and such a way, genuinely new entities, perhaps living bodies, are brought into being. Such a view would combine supervenience and emergence.

How, then, can we make sense of macroscopic bodies being ‘nothing over and above’ the microscopic facts? On the one hand, the notion expressed by ‘nothing over and above’ talk is very intuitive. Consider the relationship between partiers and parties. It is not as though there are people dancing, drinking and having fun at Bill’s, and there is this extra thing: the party. In some sense all there is for there to be a party at Bill’s is for there to be people dancing and drinking at Bill’s. Consider the relationship between individual hooligans at a football match and a violent crowd at a football match. It is not as though there are the individual hooligans causing trouble, and then there is this extra thing, the crowd, causing trouble. Intuitively, all it is for the crowd to be causing trouble is for certain individuals to be causing trouble.

On the other hand, ‘nothing over and above’ talk can seem paradoxical. I take it that the party is not *identical* to the partiers. For identity is a relation between one thing and one thing – Clark Kent and Superman, Britain and the United Kingdom – and though the party is one, the partiers are many. But then it’s difficult to see how there could be room for a third option between (i) the view that there are only the partiers, (ii) the view that there are parties *and* (the key word is ‘and’ here) some metaphysically extra thing the party. There seems not to be logical space for the position that there are the partiers *and* the party, and yet the party is somehow ‘nothing extra’ to the partiers.

I suggest that we can capture the spirit of Lewis’s view by giving macroscopic objects the same metaphysical status they have in our interpretation of Berkeley. On our Lewis-esque view, thought and talk about macroscopic objects is made true by those facts Lewis took to constitute the world as it is in and of itself. If we use sentences ‘There are tables’ in such a way that it is true if properties instantiated at points are arranged table-wise, and properties instantiated at points are arranged table-wise, then ‘There are tables’ is true, from which it follows trivially that there are tables. Again we have the result that the world is *truly describable* such that there are tables, even though the world *in and of itself* is not such that there are tables.

We have a distinction, then, between two kinds of fact: the metaphysically heavyweight facts that constitute how the world is in and of itself, and the metaphysically lightweight facts that merely constitute how the world is truly describable (the ‘merely’ signifies that they don’t constitute the world as it is in and of itself). Actually, we don’t need to adopt an ontology of facts in order to have this distinction, we can say all we need to say with two sentential operators: ‘the world in and of itself is such that’ and ‘the world is truly describable such that’. But sometimes it will be easier to talk as though we are committed to facts.[[3]](#footnote-3)

The notion of a metaphysically lightweight fact can be analysed in terms of truthmaking and metaphysically heavyweight facts: the metaphysically lightweight facts are those which are not metaphysically heavyweight, but are entailed by propositions that are made true by the metaphysically heavyweight facts. The metaphysically heavyweight facts are those upon which truth depends, but to say this is not to give an analysis of the notion of a metaphysically heavyweight fact, as we cannot get a grip on the idea of the relevant notion of truth independently of the notion of a metaphysically heavyweight fact. A proposition demands the obtaining of a certain fact; if that fact obtains, the proposition is true. The word ‘fact’ in this explication of truth is synonymous with ‘metaphysically heavyweight fact’; the notion of a metaphysically lightweight fact is posterior to this notion of truth.

The notion of a metaphysically heavyweight fact (a fact that constitutes the world as it is in and of itself) is a primitive notion (more precisely the sentential operator ‘the world in and of itself is such that’ is primitive). But it can be usefully clarified in terms of the epistemological notion of a metaphysically fundamental concept. The set of metaphysically fundamental concepts is the set of concepts S such that one cannot completely understand the nature of reality without possessing each member of S. Which concepts are metaphysically primitive will depend on the nature of reality.

Consider again the world as Berkeley took it to be. One cannot completely understand the nature of the Berkeleyan world without possessing the concepts mind and idea. One might have a fairly rich description of fundamental particles and their properties, but without conceiving of these things as ideas put into dependent minds by an independent mind, one will lack a complete understanding of their nature. The concepts mind and idea are the metaphysically fundamental concepts in the Berkeleyan world.

Similarly, although one might be able to get a fairly good grip on the Lewisian world via knowledge of the properties of macroscopic objects, I take it that an understanding of the properties instantiated at point sized regions is necessary for a complete understanding of that world. Spatio-temporal concepts, and transparent concepts of the specific quiddities instantiated at point sized regions, are the metaphysically primitive concepts at the Lewisian world.

It would get things the wrong way round to try to analyse the notion of a metaphysically heavyweight fact in terms of metaphysically primitive concepts. For which concepts are metaphysically primitive at a world is determined by which facts are metaphysically heavyweight. The concepts mind and idea are metaphysically primitive at the Berkelian world because the Berkeleyan world in and of itself is such that there are minds having ideas. Spatio-temporal concepts and concepts of certain quiddities instantiated at point sized regions are metaphysically primitive in the Lewisian world because the Lewisian world in and of itself is such that there are those quiddities instantiated at point sized regions.

# 6.2 Fundamentality and irreducibility

In the first half of the book, we used the word ‘fundamental’ to refer to the facts (and the entities involved in such facts) that constitute the supervenience base. There is also a natural use of the word ‘fundamental’ such that it refers to the metaphysically heavyweight facts: the facts that constitute the world as it is in and of itself (and the entities involved in those facts).[[4]](#footnote-4) Let us use the term ‘irreducible’ for the second meaning in order to avoid confusion.

Fundamentality and irreducibility need not be co-extensive. It may be that everything supervenes on the micro-level facts, but both micro-facts and certain higher-level facts are irreducible. Perhaps the world in and of itself is such that there are organisms and fundamental particles, and facts about organisms supervene on facts about particles. In this case, the world in and of itself consists of two layers of being, with facts about the lower layer necessitating facts about the higher layer.

We can distinguish, then, two forms of physicalism. Deflationary physicalists think that all facts intelligibly supervene on the physical facts, and that all and only physical facts are irreducible. According to the deflationary physicalist, the propositions of biology, or propositions about mentality, are made true by the physical facts. Inflationary physicalists think that all facts intelligibly supervene on the physical facts, but that some supervenient facts are irreducible. It may for example be that the physical facts necessitate the biological facts, which then make true biological propositions.

Cross sectioned with the distinction between pure and impure physicalism, we now have four forms of physicalism: pure inflationary physicalism, impure inflationary physicalism, pure deflationary physicalism, impure deflationary physicalism. The first half of this book constitutes an argument against all these forms of physicalism, as they all involve the thesis that the mental facts intelligibly supervene on the physical facts.

Crucially for this half latter of the book, we can make similar distinctions amongst non-physicalist theories. Consider the two forms of Russellian monism discussed in earlier chapters. We can say that *deflationary* proto-phenomenalists think that all facts supervene on the proto-phenomenal facts, and that all and only the proto-phenomenal facts are irreducible. According to the deflationary physicalist, Bill is in pain is made true by facts about the particles of Bill’s brain and their proto-phenomenal properties. In contrast, *inflationary* proto-phenomenalists think that all facts supervene on the proto-phenomenal facts, but that both the phenomenal facts and the proto-phenomenal facts are irreducible. The proto-phenomenal facts about Bill’s brain necessitate that Bill is in pain, which in turn makes it true that Bill is in pain.

Similarly, deflationary panpsychists think that all facts supervene on the micro-phenomenal facts, and that all and only the micro-phenomenal facts are irreducible. According to the deflationary panpsychist, Bill is in pain is made true by facts about the particles in Bill’s brian and their phenomenal properties. Inflationary panpsychists think that all facts supervene on the micro-phenoemnal facts, but both the macro and the micro-phenomenal facts are irreducible. The micro-phenomenal facts about Bill’s brain necessitate that Bill is in pain, which in turn makes it true that Bill is in pain.

In the next section, I will argue against both forms of deflationary Russellian monism.

# 6.3 Against deflationary forms of Russellian monism

## 6.3.1 More kinds of concept: Irreducible, essentially reducible and reductively neutral

Concepts make demands on the world as it is in and of itself (from now on ‘the World’), such that if those demands are met the concept is satisfied, and if those demands are not met the concept is not satisfied. The concept God demands that the World be such that there is an all-knowing, all-powerful, perfectly good being that created the universe. Propositions also make demands on the World, which are built up from the demands of the concepts out of which they are formed. If the demands of a proposition are met then the proposition is true; if the demands are not met then the proposition is false. The proposition God loves me demands that the World be such that there is an all-knowing, all-powerful, perfectly good being that created the universe, and that that being feel love towards me.

The most straightforward thing a concept can demand of the World is that its referent exists. This is plausibly the case with the concept God. But not all concepts are like this. The concept party does not demand that the World be such that there is a party; rather it demands (roughly) that the World be such that there are people gathered to socialise and have fun. The point of thought and talk about parties is to track facts about partiers. Of course, if the concept party is satisfied then it follows trivially that there is a party. But parties have no existence prior to truth; the world is truly describable such that there are parties in virtue of the World being such that there are partiers.

Let us say that a concept is ‘irreducible’ if and only if it demands that the World be such that its referent exists (e.g. God), and that a concept is ‘essentially reducible’ if and only if it does not demand that the World be such that its referent exists (e.g. party). There is a third category of concept in the middle: a reductively neutral concept has disjunctive demands, such that one of the disjuncts is the demand that the referent exists, whilst the other disjuncts are not demands that the referent exists. For example, table might turn out to demand of the Word either that it contain tables, or that particles be arranged table-wise, or that it contain table ideas regularly caused in minds by God/evil computers, etc.

It follows from the definition of an irreducible concept that its referent (if it exists) is irreducible. It follows from the definition of an essentially reducible concept that its referent (if it exists) is not metaphysically fundamental. Substantive metaphysical investigation is required to work out whether or not the referent of a reductively neutral concept is metaphysically fundamental.

## 6.3.2 Subjects of experience are irreducible

I believe that that the concept of subject of experience – the concept of a thing such that there is something that it is like to be it – is, like the concept God, irreducible. The concept subject of experience demands that its referent exists. It follows that subject of experience, like God, can only refer to something irreducible. Of course, the crucial difference between God and subject of experience is that we know for certain that the latter is satisfied. If the concept subject of experience is irreducible and satisfied, then the World must be such that subjects of experience exist.

If I am right about this, both forms of deflationary Russellian monism are false. For the deflationary proto-phenomenalist the World is not such that there is the subject of my experience. Rather, the World is such that there are particles in my head instantiating proto-phenomenal properties arranged in such and such a way, and in virtue of this the world is truly describable such that the subject of my experience exists. For the deflationary panpsychist, the World is such that there are particles in my head instantiating certain phenomenal qualities arranged in such and such a way, and this fact makes the world truly describable such that the subject of my experience exists. On both views, subjects of experience, like parties, exist only in the world as it is truly describable.

If subject of experience was radically opaque, then it would not be a priori whether it is irreducible, essentially reducible or reductively neutral, as its demands on the World would not be a priori accessible. The opacity route physicalist may claim that causal connections between deployments of subject of experience and S-wise arrangements of particles entail that subject of experience demands of the World that there be S-wise arrangements of particles. In this case, subject of experience would be essentially reducible, even though it is not a priori that this is the case. Just as party demands of the World that there be people partying (rather than demanding the existence of its referent) so subject of experience demands that particles be arranged S-wise.

However, we can quickly move from the conclusions of the first half of the book, to the conclusion that subject of experience is transparent. A particular phenomenal concept, such as pain, is transparent. If a concept of certain determinate is transparent, then the concept of the determinable of that determinate is also transparent: the concept sphericity is transparent, and hence the concept of being shaped is transparent. For one cannot understand what it is to be P-wise in a specific way, if one does not understand what it is to be P-wise. Given that we know what it is for something to be conscious in a specific way, for example, to be pained, it follows that we know what it is in general for something to be conscious: the concept consciousness is transparent.

A subject of experience is just a thing that is conscious: a thing such that there is something that it is like to be it. If we know what it is for the property of consciousness to be instantiated, then we know what it is for there to be something that is conscious. If consciousness is transparent, then subject of experience is also.

Given that subject of experience is transparent, it is a priori what it demands of the actual World. It follows that it is a priori whether it demands that the World be such that its referent exists (i.e. subject of experience is irreducible), whether it does not demand that the World be such that its referent exists (i.e. subject of experience is essentially reducible), or whether it has disjunctive demands such that one of the disjuncts is the demand that the World be such that its referent exists (i.e. subject of experience is reductively neutral). I think reflection reveals subject of experience to be irreducible, which I will try to demonstrate in what follows.

When we employ essentially reducible concepts, we are really interested in entities other than those we refer to in employing such concepts. What we talk of parties, we are really interested in partiers. When we talk of nations, we are really interested in citizens. When we talk of forests, we are really interested in trees. There couldn’t be a forest without trees because the whole purpose of the concept forest is to facilitate thought and talk about trees.

When we employ irreducible concepts, we are determinately interested in the referent of the concept. When I wonder whether God exists, my focus is very definitely on the referent of the concept. Contrast with the case of forests. When I wonder whether a certain forest exists, what I am really curious about is whether there are trees in a certain location. But my curiosity concerning whether God exists admits of no such analysis. Even if it turns out that God is formed of more basic things, call them ‘Goddons’, what I am interested in it, what directly determines the truth-value of, say, my belief in God, is whether God herself exists, not whether Goddons are arranged God-wise.

When we employ reductively neutral concepts, our metaphysical concern is neither definitely the referent of the concept, nor definitely something other than the referent. The demands of reductively neutral concepts are more easily pleased. Whether there is a chair located in the bit of the World I call my living room, or whether there are only particles arranged table-wise, so long as I am comfy watching TV all is good.

Which of these categories does thought and talk of subjects of experience fall into? Let us first consider the possibilities I want to reject. Consider the thesis that subject of experience is essentially reducible. This would entail that the real focus of thought and talk about subjects of experience is some entity or entities other than subjects of experience. The deflationary Russellian monist who holds that subject of experience is essentially reducible, might claim that the real concern of thought and talk about subjects of experience is micro-phenomenology/proto-phenomenology, just as the real concern of thought and talk about forests is trees. The obvious problem with this thesis is that people can engage in thought and talk about subjects of experience without having the slightest inkling what micro-phenomenology/proto-phenomenology is. One could not fully understand what forest talk demands of the world without knowing about trees, but one can fully understand what subject of experience talk demands of the world without knowing what micro-phenomenology/proto-phenomenology is.

Perhaps some could counter that we don’t understand what subject of experience talk demands of the World, but once we have accepted that subject of experience is transparent, this ceases to be an option. Once we accept phenomenal transparency, if we want to hold that subject of experience is essentially reducible, we are obliged to give some account of what the real focus of thought and talk about subjects is. The only way I can think of doing this is by giving some kind of causal analysis of such talk. The analytic functionalist could claim that subject of experience talk demands of the World not that there be subjects of experience, but merely that things be arranged such and such a way causally. On such a view, the fact that things in the World are arranged such and such a way causally makes the world truly describable such that there are subjects of experience. But of course, we rejected such causal analyses of the phenomenal facts in chapter 3. Once we accept phenomenal transparency and reject analytic functionalism (/behaviourism/information theoretic representationalism), there is no way of making sense of the view that subject of experience is essentially reducible.

Let us next consider the thesis that subject of experience is reductively neutral. This would entail that the demands of subject of experience are disjunctive: either (A) the World is such that there are subjects of experience, or (B) the world is such that so and so is the case, or (C) the World is such that such and such is the case, etc. The deflationary Russellian monist will then claim that one of the latter disjuncts is met by phenomenal/proto-phenomenal facts about the World.

The trouble is the only way I can think to fill in the latter disjuncts is, again, to give some causal analysis of subject thought and talk. An analytic functionalist could say that subject of experience demands either (A) that the world is such that there is a single thing playing the subject of experience role, or (B) the World is such that there are particles jointly playing the subject of experience role, (C) the World is such that the cosmos instantiates certain qualities in a certain spatio-temporal manner in a way that realises the subject of experience role. Once we have rejected analytic functionalism, which we did in chapter 3, I can see no other way to cash out such putative disjunctive demands of subject talk.

I think that the reason deflationary proto-phenomenalism seems to many to be coherent is that it generally involves the claim that we have no positive conception of the hidden nature of matter, which inclines us to think: “Given that I don’t any grip on the hidden nature of matter, how can I rule out that it has some weird and whacky nature that turns out to somehow satisfy subject of experience talk? Perhaps if I had a positive conception of matter as particles as they are in and of themselves, it would just be obvious that it satisfied the demands of subject talk”. But even if we have no positive conception of the real nature of particles, we do have access to the demands of subject talk (given phenomenal transparency). And if those demands are disjunctive, it ought to be apparent to us that this is so (given phenomenal transparency). But deep reflection does not reveal the demands of subject talk to be disjunctive, and so they cannot be so. It follows that subject of experience is not reductively neutral.

We are left with the thesis that subject of experience is irreducible. Let me finish this section by giving some positive support for this thesis. When I am concerned that there are currently many subject of experience feeling pain in Syria, there is no way of analysing my concern such that it could be about some other Worldly state of affairs than subjects of experience feeling pain. If the bit of the World I am focusing on turns out not to contain subjects feeling pain, then my sympathy is misplaced. If I wonder whether the woman I love reciprocates my feelings, there is no way of analysing my yearning such that it could be satisfied by the World containing something other than a subject of experience. If the bit of the World I am focused on turns out not to contain a subject of experience, then my romantic yearnings are deeply delusional. Concern about mentality is direct concern about the World.

When I am thinking about forests, I am really interested in trees. When I am interesting in armchairs, I am interesting in whatever is required to make me comfy watching TV. But there is no plausible way of analysing my concern for subjects of experience such that it is anything other than a concern for subjects of experience. The concept subject of experience is irreducible. It follows that if subjects of experience exist, which of course they do, they are irreducible. The world in and of itself is carved up into subjects of experience.

# 6.4 Inflationary forms of Russellian monism

## 6.4.1 Causal and non-causal grounding

Subject of experience are irreducible, but they may not be fundamental. The macro-phenomenal facts may be in some way *grounded* in the micro-phenomenal or the proto-phenomenal facts, which is to say that the macro-phenomenal facts may obtain *in virtue of* the micro-phenomenal/proto-phenomenal facts. On such a picture, the world in and of itself is carved up into distinct layers of being.

Let us make the following stipulative distinction between emergentism and Russellian monism. The emergentist holds that the macro-phenomenal facts are grounded in more fundamental facts in the sense that they are caused by more fundamental facts; X being caused by Y is one way in which X can be grounded in Y, i.e. one way in which X can obtain in virtue of Y. The Russellian monist takes the macro-phenomenal facts to be in some way *non-causally* grounded in more fundamental facts, specifically micro/proto-phenomenal facts.

But what exactly is the difference between fact X being causally grounded in fact Y, and fact X being non-causally grounded in fact Y? In what follows I will consider a number of distinct ways of understanding ‘non-causal grounding’, and hence a number of distinct ways of understanding Russellian monism. On most understandings of non-causal grounding, I will reject the thesis that the phenomenal facts are so grounded in the micro/proto-phenomenal facts. The options for the Russellian monist end up being rather limited.

## 6.4.2 Nothing over and above definition of non-causal grounding

Fact X is non-causally grounded in fact Y if fact X is nothing over and above fact Y.

Thus we might say that the fact that there is a heap of sand in location R is grounded in the fact that there are grains of sand arranged heap-wise in location R, and mean by this that the fact that there is a heap at R is nothing over and above the fact that there are grains arranged heap-wise at R. On my account of nothing over and above-ness outlined above, this entails that the World is not such that there is a heap at R. I can make no sense of the claim that the heap is distinct from and yet ‘nothing extra to’ the grains, unless it is understood as the claim that the grain facts make true the heap facts, such that the heap exists only in the world as it is truly describable (and hence does not constitute an addition in metaphysically significant being).

Hence, on the nothing over and above definition of grounding, grounded entities do not exist in the world as it is in and of itself, but only in the world as it is truly describable. This is no good for trying to make sense of the non-causal grounding of subjects of experience, as such things exist in the world as it is in and of itself.

## 6.4.3 Constitutional definition of non-causal grounding

Fact X is non-causally grounded in fact Y if X obtains in virtue of Y, and the entities involved in Y are the constitutive proper parts of the entity(s) involved in X.

To take an example, we might say that there is an organism at location L in virtue of the facts about certain organic parts located at sub-regions of L, and that this is non-causal grounding as the organic parts at L are constitute parts of the organism at L: the organic parts at L *constitute* the organism at L.

This definition will straightforwardly distinguish causal grounding of *objects* from non-causal grounding of *objects*. Thus, an emergentist about subjects of experience will take those subjects to have no micro-level parts, whereas an inflationary physicalist or Russellian monist may take subjects of experience to be constituted of the micro-parts that ground them. But if we want to distinguish causal from non-causal grounding of *properties*, then we need to think of the grounded properties as constituted of the grounding properties. Hence, the Russellian monist will need to suppose not only that the subject of my experience (an object) is constituted of micro-level parts, but that my experience itself (a property) is constituted of the experiential/proto-experiential properties of micro-level parts.

Unfortunately, this understanding of Russellian monism cannot be reconciled with the acquaintance view. When I attend to my conscious experience, its real nature is directly revealed to me. If my conscious experience were constituted of micro-phenomenal or proto-phenomenal properties, then these properties would be apparent to me in this direct revelation. But such alleged micro components of my conscious experience are not apparent in introspection. It follows that there are no such things.

I put forward this argument in earlier work, as a way of articulating the so called ‘combination problem’ for standard forms of panpsychism. A number of panpsychists and proto-phenomenalists have tried to respond to this challenge. Sam Coleman diagnoses what he calls the ‘Goff problem’ as routed in two mistaken assumptions concerning how panpsychism is supposed to work:

1. Phenomenal ultimates are themselves subjects of experience.
2. Phenomenal assembly can only be aggregative.[[5]](#footnote-5)

Coleman defends a kind of middle way between panpsychism and proto-phenomenalism. The fundamental constituents of matter have ‘phenomenal qualities’, by which Coleman means the qualities we are aware of in introspecting our experience. And yet, outside of the subjectivity of organisms, the micro-level phenomenal qualities exist in the absence of consciousness; they are ‘unfelt qualia’. The fundamental constituents of matter, then, are not subjects of experience, although they do instantiate phenomenal qualities. This enables him to jettison the first assumption and move towards a kind of phenomenal combination he thinks avoids my concerns:

Given the ‘transparent access’ thesis [i.e. the acquaintance view], Goff’s central observation is that we do not experience an aggregate of phenomenal ultimates – a collection of separate *loci* of phenomenality. Since we do not experience such an aggregate, he reasons, our consciousness is not composed of a multitude of phenomenal ultimates. This inference reveals that Goff believes the only possible manner of assembling phenomenal ultimates to be aggregative. For if he allowed that there might be other ways of arranging a phenomenal multitude, he could not move from the phenomenological claim that we do not experience an aggregate to the conclusion that our consciousness is not composed of a phenomenal multitude. Thus we see that Goff endorses the Second Assumption. But why does he think that phenomenal ultimates could only assembled aggregatively? The plausible answer is that he thinks this because he takes phenomenal ultimates to be subjects of experience, and it is *a priori* that subjects, like minds, are discrete, inviolable spheres of mentality. You can stack them, but you cannot pool them. So Goff, too, endorses the First Assumption: in fact it drives the argument (p. 147).

Dropping the idea that the phenomenal ultimates are subjects, thinks Coleman, frees us up to a non-aggregative conception of their unity, in which ‘the phenomenal ultimates mutually condition one another, as they phenomenally fuse’ (158). He suggests that we can understand such phenomenal fusion by analogy to the blending of paint in a Goya (as opposed to a Seurat), or the way in which the flavours red wine and roast beef ‘pleasingly interpenetrate’ (157-8).

Coleman is not entirely explicit, but the analogies and the talk of ‘phenomenal fusion’, suggest that we are not dealing with a layered picture of reality, where phenomenal ultimates co-exist with the phenomenal wholes they fuse into. We rather have a picture of phenomenal ultimates losing their individual identities as they morph into a unified whole. Coleman is thus contrasting two kinds of combination: aggregation and phenomenal fusion.

However, I don’t think that aggregation and fusion exhaust the combinatorial options. It is natural to think that aggregates are not real unities at all, that is, not genuine additions in being to their parts. A heap of sand, for example, is nothing over and above its parts.[[6]](#footnote-6) In fusion, the converse is the case. The parts cease to have individual existence as they melt into the whole. On the fusion model, the part of the brain which is identical with the mind no longer has parts; they have fused into a unified whole.

It is natural to seek a middle way between these two options. Let us say that in cases of ‘constitution’, as opposed to aggregation or fusion, both parts and whole exist simultaneously as distinct features of the world as it is in and of itself, and it is a fact about the world in and of itself that the nature of the whole is formed from the nature of the parts. It is the thesis that experiences are *constituted* by more fundamental entities – call this ‘phenomenal constitution’ – not the thesis that experiences are fused from more fundamental entities – call this ‘phenomenal fusion’ – that I take to be incoherent (I give empirical objections to phenomenal fusion in 6.4.4). For if trillions of micro-level phenomenal/proto-phenomenal elements constituted, as opposed to fused into, my conscious experience, then those trillions of elements would still be present in the constitution of my conscious experience post-combination. Given the acquaintance view, I would be aware of these trillions of phenomenal/proto-phenomenal elements in introspection. But I’m not.

Regarding Coleman’s denial of the first assumption, like many I find it difficult to make sense of the idea that phenomenal qualities might exist independently of subjects. Upon careful armchair reflection, it seems to me that the concept of a phenomenal quality just is the concept of a quality that characterises some experience: *how pain feels* is a quality that can only be instantiated by some subject feeling pain. This is a delicate matter, requiring the kind of calm and careful reflection I recommend in chapter 5, and Coleman has written much to try to influence our intuitions in this regard, writings which I would like to retrospectively class as fine work in the post-Galilean tradition.[[7]](#footnote-7)

However, even if we allow that micro-phenomenal qualities exist independently of subjects, I don’t think this helps us make sense of phenomenal constitution, as opposed to phenomenal fusion, for the reasons I give above. Even if there are not trillions of *subjects* constituting my conscious experience, in phenomenal constitution (as opposed to phenomenal fusion) there would be trillions of micro-level phenomenal/proto-phenomenal *properties* constituting my conscious experience. Given that I am not aware of a vast number of micro-level phenomenal/proto phenomenal qualities constituting my experience, I can infer given the acquaintance view that they do not exist.

Galen Strawson tries to avoid these difficulties by avoiding a commitment to the acquaintance view. He contrasts the following two theses:

*The Full Revelation Thesis*

In the case of any particular experience, I am acquainted with the whole essential nature of the experience just in having it.

*The Partial Revelation Thesis*

In the case of any particular experience, I am acquainted with the essential nature of the experience in certain respects, at least, just in having it.[[8]](#footnote-8)

The former thesis, properly understood, is entailed by the acquaintance view, and Strawson readily admits that a commitment to this thesis renders phenomenal constitution incoherent, for the reasons I give. But he takes this to be a ‘devastating refutation’ (255) of the acquaintance view, rather than of the view that my conscious experience is composite. Strawson adopts instead that partial revelation thesis, which is entailed by what I called in chapter 4 ‘semi-acquaintance’.

In chapter 4 I argued against semi- acquaintance on the grounds that it cannot fully account for the rational certainty I have, say, that I feel pain right now. The acquaintance view explains this rational certainty in terms of the fact that the nature of my pain is directly revealed to me. But on the semi-acquaintance view, only an aspect A of the nature of pain is directly revealed to me. The result would be rational certainty, not that I feel pain, but that I instantiate A.

Moreover, on such a view, although we are only semi-acquainted with pain, we are fully acquainted with A. We can, then, raise the same difficulties regarding the constitution of A. If A were constituted of micro/proto-phenomenal qualities, then this would be apparent in our conception of A. Given that such constitution isn’t apparent in our conception of A, we can infer that A isn’t so constituted. It seems that we have got nowhere.

In conversation, Russellian monists tend to express sympathy with the semi-acquaintance view, and follow Strawson in thinking that this is the route to making sense of phenomenal constitution. But I have never come across a Russellian monist who has cashed out semi-acquaintance by explaining exactly which aspects of consciousness are revealed to us in a phenomenal conception, and which aspects are merely opaquely denoted. How do we carve up *how pain feels* into window and screen? It is difficult to see how semi-acquaintance could be properly cashed out, and until we have a positive proposal, there doesn’t seem to me much force to the idea that phenomenal constitution may be saved by appeal to semi-acquaintance, especially given the problems I raise above.

I would like at this point to clear up a potential confusion concerning the acquaintance view which I suspect is common, and may account for some of the attraction of the semi-acquaintance view. Coleman correctly interprets me as signing up to a full revelation view, but initially describes this view as ‘heavyweight and controversial’. He suggests a more modest and palatable thesis to be the view that ‘one has, in the introspection of a given phenomenal element of which one is the subject, a direct and complete access to *how that element feels*, its phenomenal quality’ (p. 143). But this ‘modest thesis’ *just is* the full revelation view/acquaintance view, as I understand it.

The acquaintance view is the view that the nature of a phenomenal quality is directly revealed to the subject in a direct conception. This is entirely compatible with the thesis that phenomenal qualities might themselves be one aspect of some greater property, perhaps encompassing phenomenal and non-phenomenal elements. But even if, say, the anxiety I am currently feeling is part of some greater quality, I can still ask whether the feeling of anxiety itself, considered in isolation from this greater property, is constituted of more fundamental elements. Given the acquaintance view, I can infer that it is not.

## 6.4.4 Holist definition of non-causal grounding

Fact X is non-causally grounded in fact Y if X obtains in virtue of Y, and the entity(s) involved in X are parts of the entity involved in Y.

Jonathan Schaffer has recently championed *priority monism*: the view that facts about parts of the world are grounded in facts about the world as a whole.[[9]](#footnote-9) We ordinarily think of priority as going in other direction, from parts to wholes: the table is the way it is in virtue of the properties of its parts and the way they are arranged. The cosmos as a whole, assuming there is such a thing, is the way it is in virtue of facts about stars and planets contained within it. The priority monist reverses this order. The parts of the table are the way they are because of how the table is as a whole. The one fundamental object is the cosmos as a whole; stars, planets, table and people are the way they are because of how the one fundamental object is. The cosmos is the ground of all being.

There are a number of ways we might think of the properties of the universe. We could think of them as *distributional properties*. The notion of a distributional property comes from Josh Parsons; some examples are the property a surface has of *being polka dotted*, or the property a poker has of *being hot and one end and cold at the other*.[[10]](#footnote-10) Intuitively distributional properties are ways of ‘filling in an object’, as Parsons puts it. A physicalist could take the distributional properties of the cosmos to be *having such and such a distribution of mass*, *having such and such a distribution of charge*, etc. Imagine a nicely ordered possible world which has its mass arranged throughout spacetime in a perfect polka dotted distribution. If this world is an object, then it has the distributional property of having a polka dotted distribution of mass.[[11]](#footnote-11) For the priority monist thinking along these lines, the facts about the distributional properties of the cosmos ground facts about parts of the cosmos. For example there might be massy particles arranged in a polka dotted pattern throughout spacetime, in virtue of the fact that the cosmos as a whole instantiates a polka dotted arrangement of mass.

Alternately, we can follow Horgan and Potrc in thinking of the universe as instantiating certain properties in spatio-temporally local manners.[[12]](#footnote-12) Suppose there is a hard, brown table located in region R. A priority monist following the Horgan and Potrc analysis of the properties of the universe might ground this state of affairs in the fact that the universe instantiates solidity and brown-ness R-wise.

How might we make sense of the phenomenal facts being grounded in facts about the cosmos? Suppose my subject of experience is located at location L1, and the subject of your experience is located at location L2. Following the Horgan and Potrc model, we could say that the cosmos instantiates my conscious experience L1-wise, and your conscious experience L2-wise. It is a little difficult to make sense of what it means to instantiate a state of consciousness in a given spatio-temporal manner, but perhaps this is just a reflection of our ignorance regarding the fundamental nature of the world.

A more pressing concern with this strategy is that it doesn’t seem to allow us to make sense of there being more than one subject. If your consciousness and my consciousness and both instantiated by the cosmos, then what we have are not two distinct subjects, but one subject that has both your consciousness and my consciousness. Although Horgan has written much on consciousness, Horgan and Potrc don’t for some reason focus on consciousness in their writings on monism. However, they would, I think, try to get round this problem by claiming that propositions concerning distinct subjects of experience are made true by facts about the single cosmos as a whole.[[13]](#footnote-13) This is not an option once we have accepted, as we have, that the World is such that there are multiple subjects of experience.

What about Parsonian distribution properties? We could say that the cosmos instantiates such and such a distribution of subjects across space, and the fact that the cosmos instantiates this ‘subject distribution’ grounds the existence and nature of individual subjects. I don’t think we are able to conceive of subjects being distributed across space such that the fact that the universe instantiates such and such a subject distribution is prior to the facts about individual subjects. Inevitably, we end up conceiving of the individual subjects, and then thinking of the cosmos as a mere aggregate of these subjects (and anything else located in space). But again, we can put this down to our ignorance. Each of us has direct insight into the nature of only one subject; we don’t have direct insight into some larger bit of the World which might potentially ground that one subject.

It may be then, that the cosmos instantiates some unified global property which somehow includes my consciousness and your consciousness as aspects, in such a way that the existence and nature of all subjects is grounded in the existence and nature of the cosmos. We might get some kind of grip on this by analogy to aspects of one’s own unified conscious experience. At the present moment I instantiate a determinate state of consciousness. But that state of consciousness has aspects: the throbbing pain as in my left knee, the phenomenal red as of the duvet covering me, the auditory experience as of birds outside. Intuitively, the state of consciousness as a whole grounds these individual phenomenal qualities by containing them as aspects. Similarly, perhaps the cosmos grounds your subject of experience and my subject of experience by containing them as aspects.

The analogy with the cosmos is not perfect, as the individual aspects of my consciousness experience are not objects in their own right, whereas the subject of my experience and the subject of your experience are objects their own right. Whilst the subject of my experience grounds only properties, the cosmos – if it is the one fundamental object – grounds distinct objects. Nonetheless, this analogy provides some kind of insight into the possibility that the cosmos is the ground of all phenomenology.

Once we have accepted the acquaintance view, this picture will deviate from standard versions of priority monism. On the most theoretically satisfying versions of priority monism, the existence and nature of the cosmos grounds the existence of certain smaller states of affairs, which in turn ground the existence of certain smaller states of affairs, and so on right down to the micro-level. On such a view, my consciousness would ground trillions of micro-level properties, perhaps directly, or perhaps indirectly by grounding properties which ground properties which ground properties…eventually bottoming out at micro-level properties.

However, given the acquaintance view, we know that my consciousness, assuming that it is a macroscopic property of a reasonably sized bit of my brain, does not ground trillions of micro-level properties. For if it did, I would be able to deduce the existence of these micro-level properties given my complete understanding of the nature of my consciousness. The grounding of smaller states of affairs in larger states of affairs, then, must bottom out in this case, not at the micro-level, but at the quite high macro-level at which my consciousness exists. The subject of my experience will turn out to be a *large simple*: a macroscopic object which does not have parts.

We end up with a view somewhat similar to the phenomenal fusion view discussed in 6.4.3. For on the phenomenal fusion view, too, a subject of experience is a large simple, as the smaller entities that formed it melted away in its formation (whether or not Coleman intends to commit to phenomenal fusion, the view is explicitly defended by Bill Seager[[14]](#footnote-14)). On the phenomenal fusion view, a subject of experience, once formed, is a fundamental entity. On the priority monist view currently under consideration, a subject of experience is grounded in the cosmos which contains it as an aspect. But in both cases, a subject of experience has no parts.

I don’t think that the thesis that subject of experience are big simples is incoherent, but it is on the face of it empirically implausible. For it seems that any large part of my brain *does* have parts. In the next chapter, we will consider whether a functioning brain has *emergent causal powers*, that is, causal powers over and above the causal powers of its parts. But whether or not the brain has emergent causal powers, it seems that we have a rich understanding of the functioning brain in terms of the functioning of its parts. We understand the basic structure of neurons, and how they communicate by means of neuro-transmitters. We have good grip on the division of labour between various areas of the brain, and of how overall functioning can be altered by changing parts of the brain, for example how long term potentiation can strengthen a given neural pathway and increase the likelihood that a single neuron can cause an action potential. It is hard to make sense of this empirical knowledge on the assumption that the brain, or a large part of it, simply lacks parts.

Compare to the Eintein-Bose condensates. [SORT THIS]

At the very least, I think the Russellian monist wanting to take this route must do a lot more work to make the thesis that brains are big simples consistent with contemporary brain science. For the time being, I will take it that we have still not found a sense in which subjects might be non-causally grounded in more fundamental facts.

## 6.4.5 Realisation definition of non-causal grounding

Fact X is non-causally grounded in fact Y if Y realises X.

We can get at the notion of realisation through examples. The hardware of my laptop realises its software. The neurological workings of the brain realise its functional states. The mechanism of the watch realises its time-telling function. In all of these cases, one can have an understanding of the realised property without understanding its realiser. I have a complete understand of the higher-level functioning of my watch, but have no clue about the underlying mechanism.

This suggests that if we can understand the relationship between macro-mentality and micro/proto-mentality on the model of realisation, then we can reconcile Russellian monism with the acquaintance view. It could be claimed that, just as I completely understand the higher-level function of my watch, but have no clue about its underlying mechanism, so I completely understand the nature of my own experience, whilst being in the dark about its underlying micro/proto-phenomenal realiser.

The problem is that in all cases of realisation, we form the description of the realised property by a process of abstraction. To speak loosely, we take a concrete state of affairs, and then consider it ‘without all the details filled it’. The concrete reality I am now typing on is a complicated physical mechanism. When I consider it as a laptop, I abstract away from the details of the physical mechanism; to describe the realised is to give a partial description of the realiser.

There are two consequences of this. Firstly, a realised entity can exist independently of its actual realiser, for in some alternative situations the partial description that characterises the realised entity might be filled in differently. My laptop might have its internal mechanisms replaced by quite different mechanisms in such a way as to preserve higher-level functioning. But at the same time, a realised entity cannot exist independently of *any* realiser, for to describe some realised entity is just to give a partial description of some realiser.

However, it is possible for consciousness to be instantiated in the absence of any other property. We can reach this conclusion by conjoining the conceivability of ghosts (argued for in chapter 3), the new transparency thesis (defended in chapter 4) and phenomenal transparency:

1. *Conceivability of ghosts –*There are ghosts is conceivably true.
2. *New Transparency thesis –* If there are ghosts is conceivable, then it is true at some possible world considered as actual.
3. *Phenomenal transparency*: Phenomenal concepts are transparent.
4. Therefore, there are ghosts is true at some world considered as actual (from 1-3).
5. Therefore, there are ghosts is true at some world considered as counterfactual (from 3&4).

The counterfactual world where ghosts exist just is a world where consciousness exists without being grounded in any other property. The fact that such a world is possible entails that consciousness is able to exist without being realised, and hence our conception of consciousness is not formed by abstracting from some more ‘filled in’ state.

Again, this doesn’t mean that my conscious experience is not part of some bigger state of my brain, perhaps involved non-experiential, or micro-experiential aspects. But in this case, the aspect of that state that is my conscious experience is a fully formed property in its own right, not an abstraction from some fully formed property.

## 6.4.6 Necessitarian definition of non-causal grounding

Fact X is non-causally grounded in fact Y if X obtains in virtue of Y, and the obtaining of Y necessitates the obtaining of X.

Perhaps one might think that the difference between causal and non-causal grounding is that the former is contingent – dependent on the contingent laws of nature – whilst the latter is necessary. Indeed, it used to be commonplace to suppose that the laws of nature are contingent, whilst ‘metaphysical laws’, such as laws of mereology, are necessary.

However, these days there are philosophers that deny each of these. Causal essentialists take the laws of nature to be necessary: if the essence of mass is to endow objects with the disposition to attract other things with mass, then there is no possible world in which massy objects don’t attract each other. And some argue that both mereological nihilism (parts never form wholes) and merelogical universalism (any collection of parts form a whole) are possible.[[15]](#footnote-15)

Perhaps it is harder to make sense of the laws of nature being necessary if properties have a categorical, rather than a causal real nature, and it is natural to take proto-phenomenal properties to have a categorical nature. But this difficulty may just be a matter of our epistemic limits. We have only one small window onto the categorical nature of the world: our own conscious experience. Perhaps if we really understood the categorical nature of mass, it would just be obvious that massive objects must attract each other. In the same way, perhaps if we really understood the categorical nature of proto-phenomenal properties, it would be just obvious that that they must give rise to consciousness. In this case, we could not distinguish these two cases of grounding in terms of their modal status.

Of course, if we suppose that diachronic grounding relations between distinct events in a single layer of the World are contingent, and that synchronic grounding relations between distinct layers of the World are necessary, then we could take advantage of this difference to define ‘causal’ relations as contingent. We would then get the result that the grounding between match striking and match lighting is ‘causal’, whilst the grounding between micro/proto-phenomenal facts and macro-phenomenal facts (assuming there is such a thing) is ‘non-causal’. This is a terminological choice one is entitled to make. However, we must bear in mind that these ‘non-causal’ relations are indiscernible from relations we would be happy to call ‘causal’ if the world turned out to be as described in the previous paragraph.

The substantive point is that my determinate state of consciousness is a fully-formed, simple property that is instantiated in the world as it is in and of itself. If may be that it is brought into being by more fundamental fully-formed, simple properties, and to that degree my consciousness may turn out to be non-fundamental. But it is not dependent on more fundamental properties in the more obviously non-causal of being *constituted of* or *realised by* more fundamental properties.

1. Reference [↑](#footnote-ref-1)
2. Reference [↑](#footnote-ref-2)
3. Reference [↑](#footnote-ref-3)
4. Reference Sider and Williams. [↑](#footnote-ref-4)
5. Reference [↑](#footnote-ref-5)
6. Even if there is an object in the world as it is in and of itself that is composed of the grains of sand, that object is not the heap, as heap an essentially reducible concept (the point of heap thought and talk is to track grains). In the same way, if the World is such that there is some weird object composed of the temporal parts of the partiers during the time they are partying, that objects is not the party. [↑](#footnote-ref-6)
7. Reference [↑](#footnote-ref-7)
8. [↑](#footnote-ref-8)
9. Reference [↑](#footnote-ref-9)
10. Reference [↑](#footnote-ref-10)
11. Jonathan Schaffer (reference) suggests that we could also use the notion of a *regionalised properties* or the *regionalised instantiation* of a property to make sense of the properties of the universe as a whole, but I will just focus on distributional properties for the sake of simplicity. [↑](#footnote-ref-11)
12. Reference [↑](#footnote-ref-12)
13. In fact, they defend not *priority monism*, but *existence monism*, the view that the World is such that there is only one object, hence this seems to be the only strategy available to them. [↑](#footnote-ref-13)
14. Reference [↑](#footnote-ref-14)
15. Reference [↑](#footnote-ref-15)