

# Grounding, Essence, and the Knowledge Argument

Philip Goff, Central European University

**Forthcoming in Coleman (Ed.) *The Knowledge Argument: Then and Now*, Cambridge University Press.**

Few these days dispute that the knowledge argument demonstrates an *epistemic gap* between the physical facts and the facts about experience. I will assume it what follows that when Mary sees red for the first time she gains new propositional knowledge, knowledge which she was unable to deduce from her knowledge of the physical facts underlying red experiences. It is much more contentious whether that epistemic gap can be used to demonstrate a *metaphysical gap* of a kind that is inconsistent with physicalism. In this paper I will explore two attempts to block the inference from an epistemic gap to a metaphysical gap – the first from the *phenomenal concept strategy*, the second from *Russellian monism* – and suggest how the proponent of the knowledge argument might respond to each of these challenges. In doing so, I will draw on recent discussions of grounding and essence in the metaphysics literature.

## Part I – The Challenge from the Phenomenal Concept Strategy

### (i) What is the phenomenal concept strategy?

The phenomenal concept strategy is currently the most popular physicalist strategy for responding to the knowledge argument and related anti-physicalist arguments.<sup>1</sup> The phenomenal concept strategy concedes that post-liberation Mary is able to think about red experiences in a way she wasn't able to when she was trapped in the black and white room. She can bring to mind a red experience and think about it in terms of what it's like to have it. You just can't think about red experiences in this way if you've never before had a red experience.<sup>2</sup> Such concepts, of the kind one deploys when one thinks about a conscious state in terms of what it's like to have it, have become

---

<sup>1</sup> Loar 1990; Papineau 1998; Diaz-Leon 2010.

<sup>2</sup> It may of course be a contingent fact about humans that they get to possess a phenomenal concept of experience E only by having experience E. This does not contradict the crucial point that Mary gains new propositional knowledge, knowledge that is not derivable from knowledge of physical facts. The epistemic gap can be further supported with the Marianna thought experiment discussed below.

known as 'phenomenal concepts'.<sup>3</sup> When Mary sees red for the first time she gains a phenomenal concept of red experiences.

However, from the mere fact that post-liberation Mary has a new concept, a new way of thinking about a feature of reality, it does not follow that the feature of reality she thinks about when she employs that concept is also new. Consider the following analogy. Suppose Mary has a quite ordinary child, Frank. Mary is keen on encouraging Frank to take an interest in astronomy, and points out the Morning Star to him when he is five year's old. Frank subsequently thinks a lot about the Morning Star, and enjoys spotting it in the sky. Now suppose that when he is eight year's old Frank hears about Venus, and upon questioning his mother is told that the Morning Star is Venus. Eight year old Frank has picked up a new concept, a new way of thinking something in the world. But he did not thereby learn about some *new thing* in reality: the concept is new, but what it refers to is something Frank already knew about, namely the Morning Star.

The phenomenal concept strategist makes an analogous claim about Mary. Pre-liberation Mary knew about brain state X, the brain state that is identical with red experiencing. Post-liberation she gains a new phenomenal concept, a new way of thinking introspectively about a certain feature of reality. But her new phenomenal concept refers to something she already knew about pre-liberation, namely brain state X. 'What it's like to see red' and 'brain state X' – just like 'the Morning Star' and 'Venus' – refer to one and the same feature of reality.

Thus, Mary gains brand new full-fledged propositional knowledge when she sees red for the first time. But that knowledge consists of a different way of thinking about physical properties she already knew about: she learns that *this thing* (brain state X picked out with a phenomenal concept) is identical with *that thing* (brain state X picked out with a physical concept). The knowledge argument refutes physicalism only if Mary learns about new properties when she leaves the room: these would have to be non-physical properties, given that she already knew about all the physical properties in the room. If what Mary learns is just new ways of thinking about properties she already knew about, then physicalism is safe.

Some argue that whenever one learns a new way of thinking about something, one thereby discovers a new property of that thing.<sup>4</sup> It might be supposed, for example, that when Frank goes from thinking about the Morning Star as the Morning Star, to thinking about the Morning Star as

---

<sup>3</sup> The term 'phenomenal concepts' comes from Chalmers 1996, but use of the term has subsequently become widespread on both sides of the debate.

<sup>4</sup> See, for example, White 2007.

'Venus', he thereby learns a new property of the Morning Star: the property of being the second planet from the Sun. If this is always the case, then in gaining a new concept of red experiences, Mary would thereby learn about a new property of red experiences, and the threat to physicalism would return (For given that Mary didn't know about this property in the black and white room, we are led to the conclusion that that property is non-physical).

However, it is not obviously true that possession of a new concept always comes along with knowledge of a new property. To return to the above example, it is possible to possess the concept of Venus without knowing that it's the second planet from the Sun (I had to Google it to make sure). And it's conceptually coherent to suppose that Venus is not a planet at all. We can imagine discovering that in fact Venus is a spaceship created by aliens to spy on us. This hypothesis may not be empirically very plausible, but it is not incoherent, from which it follows that it is not a priori that Venus is a planet. It is possible that in picking up the concept 'Venus' Frank learns nothing new about the Morning Star; he simply gains a new label for it. Analogously, according to the phenomenal concepts strategy, when Mary leaves her room she simply gains a new label for a feature of the world she already knew about.

Some would still want to push the point: there must be *something* new Frank learns about the Morning Star when he learns that it's Venus. There must be *some* positive information associated with the concept 'Venus', in terms of which the concept characterises its referent. Likewise, one might be tempted to think, there must be some positive information associated with the phenomenal concept of red experiencing, in terms of which the concept characterises the experience of red. But at this point we can see that the knowledge argument relies on quite substantive claims about the workings of our concepts: that each and every concept picks out its referent by latching onto some property of it, some property which is a priori accessible to the concept user. Nothing in the story of black and white Mary justifies this highly contentious meta-semantic assumption. Indeed, if we have strong reason to accept physicalism, then the fact that dropping that meta-semantic assumption allows us to escape the knowledge argument (if indeed it does) arguably gives us good reason to do so.

## (ii) [Rescuing the knowledge argument from the phenomenal concept strategy](#)

It is plausible that properties and kinds have natures (or 'essences' – I will use these words interchangeably), in the sense that it is a factual matter what it is for the property to be instantiated. The best way to get a grip on this notion, and to make a case for it, is with reference to examples:

- *The nature/essence of sphericity*: For it to be the case that there is something spherical is for it to be the case that there is something with all points on its surface equidistant from its centre.
- *The nature/essence of party-hood*: For it to be the case that there is a party is for it to be the case that there are people revelling.
- *The nature/essence of the kind water*: For it to be the case that there is water is for it to be the case that there is something composed of H<sub>2</sub>O molecules.

We can call the above descriptions ‘metaphysical analyses’ of the properties and kinds: descriptions that capture what it is for the property or kind to be instantiated.

In terms of this notion of nature, we can distinguish between two kinds of concept:

- *Transparent concept* – A concept C of entity E is transparent just in case C reveals the nature of E (i.e. what it is for E to exist or be instantiated is a priori accessible for someone possessing C, in virtue of possessing C).
- *Opaque concept* – A concept is opaque just in case it reveals nothing about the nature of its referent.<sup>5</sup>

The concept of sphericity is a plausible example of a transparent concept. For the property of sphericity to be instantiated is for there to be something with all points on its surface equidistant from its centre; if you possess the concept of sphericity, and you’re clever enough, you can work that out from the armchair. The concept of water is a plausible example of an opaque concept. For there to be water is for there to be stuff composed of H<sub>2</sub>O molecules; but this cannot be found out from the armchair. In contrast to sphericity, you have to actually do some science to find out the essence of water.

Are phenomenal concepts transparent or opaque? When we apprehend red phenomenology in phenomenal terms, are we apprehending it in terms of its essential nature? If not, if phenomenal concepts are opaque, the knowledge argument is powerless against physicalism. For in this case, upon liberation Mary will gain new propositional knowledge, but that new propositional knowledge is no threat to physicalism. All she has gained is a new way of pointing to something in the world, and what she is pointing at may very well turn out to be one of the physical properties she knew about in her black and white room.

---

<sup>5</sup> There is an option in between transparency and opacity, which I discuss in footnote 7.

But if phenomenal concepts are transparent, then the knowledge argument begins to have bite. For it follows that, when liberated Mary conceives of red experiencing under a phenomenal concept she is conceiving of it in terms of its essential nature. Now suppose physicalism is true, and red experiencing is a physical property. Assuming (for the moment, we will return to this assumption below) that physical science reveals the nature of physical properties, Mary would have learnt about the nature of that property in her black and white room. She would thus be able to recognise that the nature she is currently apprehending in phenomenal terms is identical to the nature of some physical property she knew about in the black and white room. What she had previously learnt about through book reading she would now be apprehending in the flesh, as it were.

But Mary would not be able to do that, as can be made clear with reference to Martine Nida-Rümelin's development of the knowledge argument involving the character Marianna.<sup>6</sup> Marianna has the same upbringing and education as Mary. However, upon liberation Marianna is not shown a typically red object – a tomato or a model of a fire engine – which would allow her to deduce that the colour she is currently apprehending is red. Rather she is shown a colour strip, or an arbitrarily painted object. Liberated Marianna has a red experience, which she can apprehend in phenomenal terms. But it's extremely plausible that introspective reflection on her red experience would not enable Marianna to deduce that what she is currently having is a red experience. That is to say, she would not be able to deduce that the colour she is currently apprehending is the colour of tomatoes and fire engines.

Putting all this together, an argument against physicalism emerges. If both phenomenal concepts and physical descriptions reveal the essence of the properties they denote, and red experiencing just is a physical property, then Marianna would be able to pair up the red experience she is currently having with some physical property she learnt about in her black and white room, by noting that the two have an identical nature. She would thereby be able to work out that the colour experience she is currently having is a red experience. But this is implausible.

We can thus pose the following refined knowledge argument against physicalism:

#### The Refined Knowledge Argument

1. If <physicalism is true & phenomenal transparency is true & physical science reveals the nature of the properties it denotes>, then <upon having a red experience Marianna would be able to deduce that she is having a red experience>.

---

<sup>6</sup> Nida-Rümelin 1996, 1998.

2. Phenomenal transparency is true.
3. Physical science reveals the essence of the properties it denotes.
4. Upon having a red experience, Marianna is unable to deduce that that she is having a red experience.
5. Therefore, physicalism is false.

The crucial premise of this argument, premise 2, although pretty intuitive is not uncontroversial. I have defended phenomenal transparency at length elsewhere but will not repeat these arguments here.<sup>7</sup> My claim here is simply that any defence of the knowledge argument depends on a defence of phenomenal transparency. And indeed I think this is reflected in the basic intuitions that lie behind the argument. We feel that Mary, in encountering red, learns something new *about the nature of red experiences*, and that this is hard to square with her previous education. If the 'knowledge' of liberated Mary involves no new understanding of the essence of red phenomenology, then it's hard to see how it could threaten physicalism.<sup>8</sup>

One final point. Some may want at this point to reapply the phenomenology concept strategy, by claiming that in her black and white room Mary does know the essential nature of red phenomenology but under a physical mode of presentation. When she sees red for the first time she learns about that same essential nature, but this time under a phenomenal mode of presentation (where this new mode of presentation is conceptually distinct from the physical mode of presentation in the sense that it is not possible to know a priori that these two modes of presentation pick out the same feature of reality).<sup>9</sup> However, whilst it is uncontroversial that one can *refer* to a single property under multiple conceptually distinct modes of presentation, it is a much more problematic thesis that one can *know the essence* of a single property under multiple conceptually distinct modes of presentation. There is of course still a debate to be had, which I have engaged in at greater length elsewhere.<sup>10</sup> Suffice to say that the refined knowledge argument is in much better shape than the basic knowledge argument to resist the phenomenal concept strategy.

---

<sup>7</sup> Goff 2011, 2015a, 2017.

<sup>8</sup> I am over-simplifying a little. Rather than being transparent or opaque, phenomenal concepts might be *translucent*, revealing something but not everything about the natures of their referents. Can a physicalist adopting phenomenal translucency avoid the knowledge argument? It depends on the details of the view. The physicalist would have to give a detailed account of which features of red experiencing are revealed by the corresponding phenomenal concept, and why these features do not allow Marianna to pair up red experiencing with the brain state it is identical with. In any case, the proponent of the knowledge argument is on safer ground if full phenomenal transparency can be defended, which I believe it can. I discuss these matters in much more detail in chapters 4 and 5 of Goff 2017.

<sup>9</sup> This kind of position is defended by Damjanovic 2012 and Taylor 2013.

<sup>10</sup> Goff 2015b, Goff 2017: ch. 5.

## Part II – The challenge from Russellian monism

### (i) What is Russellian monism?

We have been discussing whether or not phenomenal concepts are transparent. We turn now to consideration of the analogous issue with respect to physical concepts. Consider the following opposing theses:

*Physical Transparency* – The physical sciences reveal the nature of physical properties: a true neurophysiological description of c-fibre firing reveals the nature of c-fibre firing, physical characterisations of mass reveal the nature of mass.

*Physical Opacity* – The physical sciences reveal little or nothing of the nature of physical properties.

The physical sciences characterise properties in terms of their nomic role. Very roughly mass is characterised in terms of gravitational attraction and resisting acceleration. Brain states are characterised in terms of their role in the overall functional economy of the brain and in terms of their chemical constituents; those chemical entities are in turn characterised in terms of their causal role and physical constituents. In the light of this, there are two options regarding the meaning terms denoting physical properties:

*Causalism* – Physical terms, e.g. mass, refer to dispositional properties.

*Quidditism* – Physical predicates, e.g. mass, refer to categorical properties, but pick out those categorical properties in terms of their causal role.<sup>11</sup>

According to causalism mass is a dispositional property: roughly the disposition to resist acceleration and to attract other things with mass. According to quidditism, mass is a categorical property that we pick out in physics in terms of the causal powers it endows to its bearers (again, roughly the disposition to resist acceleration and to attract other things with mass). Causalism naturally leads to Physical Transparency: physical properties are causal role properties which are characterised in causal terms. Quidditism naturally leads to Physical Opacity: the categorical property of mass (for example) is picked out not in terms of its own intrinsic nature, but in terms of the dispositional properties it grounds. It is perhaps indeterminate which of these options reflects how physical scientists use these terms. In any case, which of these options a given philosopher is attracted to is

---

<sup>11</sup> To resolve a slight ambiguity, I intend quidditism to be the view that the use of physical terms *aims* at referring to categorical properties, which is consistent with reference failing if there are no such properties (although, see footnote 12).

likely to be determined by her or his views on the metaphysics of properties. A causal structuralist, who thinks that properties have causal essences is likely to adopt causalism. The Humean, or any philosopher who believes that properties have a categorical essence, is likely to adopt quidditism.

Suppose Physical Opacity is true because quidditism is true. It follows that physical properties have a 'hidden nature' which goes beyond what is revealed by the physical sciences.<sup>12</sup> Inspired by certain passages of Russell's *Analysis of Matter*, a numbers of philosophers have recently seen in this the potential for a promising solution to the mind-body problem.<sup>13</sup> The thought is: Perhaps it is this hidden nature of matter that explains consciousness; perhaps our difficulty making sense of consciousness as a property of matter – arguably the root of the mind-body problem – is due to our scientific ignorance of matter's intrinsic nature. This broad approach, which comes in a variety of forms, has become known as 'Russellian monism.'<sup>14</sup> We can think of Russellian monism as having a positive and a negative aspect, as follows:

- There is a 'deep' nature to basic physical properties, which goes beyond the causal-structural features in terms of which physics characterise those properties.
- The true explanation of consciousness is to be given in terms of the deep nature of physical properties.<sup>15</sup>

Is Russellian monism a form of physicalism? It depends on how we define physicalism. Some take physicalism to be the view that ideal physics exhaustively describes fundamental reality. Clearly Russellian monism is not a form of physicalism on such a definition. Others define physicalism as the view that physics is *referentially adequate*, in the sense that the fundamental facts, individuals or properties are those which are the subject matter of ideal physics, leaving it open whether or not ideal physics reveals their complete nature. For the purposes of this piece, let us think of Russellian

---

<sup>12</sup> This implication could be avoided if one held that physical terms refer to non-existent quiddities. But I am assuming that sentences of ideal physics are true, or at least that their truth is not undermined by this kind of metaphysical consideration. If there are no quiddities, then quidditism is not the correct theory of the meaning of physical terms.

<sup>13</sup> Alter & Nagasawa 2015 is a good collection of essays on Russellian monism. Goff 2017 categorises and evaluates various forms of the view.

<sup>14</sup> Russell 1927.

<sup>15</sup> In order to distinguish Russellian monism from orthodox forms of physicalism, we can add that for the Russellian monist (i) there is an a priori entailment from the deep physical facts to the facts about consciousness (thus distinguishing it from type-B physicalism), and (ii) that a priori entailment does not obtain wholly in virtue of the causal-structural features of matter (thus distinguishing it from type-A physicalism). For more details, see Goff 2017: ch. 2.



monism as a form of physicalism, in order that we might examine whether the knowledge argument is able to refute physicalism of that form.<sup>16</sup>

### (ii) The Russellian monist response to the knowledge argument

Mary in her black and white room is described as knowing ‘all the physical facts’. However, it’s clear from the context that what this means is that she knows ‘all the facts physical science has to teach us about the physical’. For the Russellian monist, such facts do not exhaust the complete nature of the physical, as they leave out its deep nature. If Mary really did know all the physical facts, including the deep nature of the physical, then according to Russellian monism she would have a complete understanding of the grounding base of phenomenal experiencing, and would thereby be able to work out what it’s like to see red. But physical science cannot teach us this, as it picks out physical properties in terms of their dispositions rather than their deep nature; it follows that the third premise of the refined knowledge argument (that physical science reveals the nature of the properties it denotes) is false.

For the Russellian monist, Mary in her black and white room knows all the facts of physics but she doesn’t know all the physical facts. Because pre-liberation Mary’s knowledge of the relevant physical facts is incomplete, the fact that liberated Mary gains some new knowledge is no threat to physicalism.

### (iii) Rescuing the knowledge argument from Russellian monism

The Russellian monist response to the knowledge argument focuses on the grounding base of consciousness, claiming that the deep nature of the physical (not merely its causal-structural features) is required to account for consciousness. If pre-liberation Mary could somehow access the deep nature of the brain, she would be able to work out what it’s like to see red, but she can’t so she doesn’t. Now it does not seem that we currently have a positive conception of the kind of deep physical properties the Russellian monist believes underlie consciousness, properties knowledge of which would allow a congenitally blind person to work out what it’s like see red. But this fact does not in itself rule out the possibility that there are such properties. Maybe we just haven’t theorised our way to them yet, or maybe essential features of the human epistemic situation put them forever beyond our ken.<sup>17</sup>

---

<sup>16</sup> I don’t think it is helpful to categorise Russellian monism as a form of physicalism, as I argue in Goff 2017. The substantive question is whether the knowledge argument can rule it out.

<sup>17</sup> Colin McGinn (1989) argues that we are ‘cognitively closed’ to the properties that explain consciousness, whilst Derk Pereboom (2011) is a Russellian monist who expresses hope that we may one day uncover such properties. In Goff 2017: ch.7 I comparatively evaluate these possibilities.

To repeat: the focus of the Russellian monist response is upon the putative grounding base of consciousness. However, recent discussions of the nature of grounding have suggested that, in order for physicalism to be true, the nature of *consciousness itself* must play a significant role in explaining the grounding of consciousness in the physical. In what follows I will develop an argument, starting from the premises of the refined knowledge argument, that consciousness is not able to play this role. If that argument is sound, then no matter what weird and wonderful properties lie at the base of the brain, they cannot possibly ground consciousness, for consciousness itself is not suited to being so grounded.

#### (iv) Grounding via essence

Suppose Rod, Jane and Freddy are dancing, drinking and generally having fun one evening at Jane's. It follows from this supposition that there is a party at Jane's, and moreover that there is a party at Jane's *because* Rod, Jane and Freddy are dancing, drinking, etc. at Jane's. But the word 'because' here does not express a *causal* relationship; it's not as though the activities of the revellers bring into being some extra thing – the party – which then floats above their heads. Consider a further example. Suppose the rose is scarlet. It follows that the rose is red, and moreover that the rose is red *because* it is scarlet. But the scarlet colour of the rose does not secrete redness as the liver secretes bile. It seems that in both cases we have a kind of explanatory relationship which is not causal. This relationship has become known as 'grounding'.<sup>18</sup>

There is a growing consensus (in both philosophy of mind and in metaphysics) that supervenience accounts of physicalism are inadequate, and that accounts of physicalism in terms of grounding look to be the best replacement. However, Shamik Dasgupta has recently shown that defining physicalism in terms of grounding is much less straightforward than one might at first suppose.<sup>19</sup>

An obvious first attempt at defining physicalism in terms of grounding would be the following:

*Strong Physicalism* – Physicalism is the view that all non-physical facts are grounded in the physical facts.

This gives us a clear and straightforward way of understanding the view that fundamental reality is wholly physical: the many and diverse facts which make up reality are all ultimately grounded in the physical facts. Suppose that Mary is currently experiencing red. Strong Physicalism commits the physicalist, as we would expect from a definition of physicalism, to there being some physical fact

---

<sup>18</sup> Some key papers on the recent revival of grounding are Fine 2001, Schaffer 2009 and Rosen 2010. Proponents of grounding trace the idea back to an older tradition, often citing Aristotle as an influence. See Trogon 2013 for an introduction to grounding.

<sup>19</sup> Dasgupta 2014.

which grounds the fact that Mary has red phenomenology. Let us suppose that the fact that Mary has red phenomenology is grounded in the fact that she instantiates brain state B. Thus we reach the following fact:

*Red-Grounding* – The fact that Mary is in brain state B grounds the fact that Mary has red phenomenology.

Red-Grounding is a *grounding fact*, that is to say, a fact about which facts ground which. In formulating his grounding conception of physicalism, Dasgupta of course accepts that the physicalist is obliged to hold that the facts about consciousness are grounded in the physical facts, and hence is obliged to accept grounding facts like Red-Grounding. However, he denies Strong Physicalism because he does not think that the physicalist needs to hold that the *grounding facts themselves*, i.e. facts like Red-Grounding, are wholly grounded in the physical facts.

If Strong Physicalism is true, then Red-Grounding, like any other fact, is grounded in the physical facts. But Dasgupta argues that Red-Grounding cannot be wholly explained in terms of the physical; rather it must be explained at least in part in terms of the nature of red phenomenology. It is the nature of red phenomenology which explains why it is that brain state X grounds red phenomenology.<sup>20</sup>

To make this plausible, return to our party example. Just as Red-Grounding concerns the grounding of red phenomenology, so the following fact concerns the grounding of parties:

*Party-Grounding*: The fact that Rod, Jane and Freddy are revelling grounds the fact that there is a party.

Why is it the case that the fact that Rod, Jane and Freddy are revelling grounds the fact that there is a party? Intuitively this is because of the nature of a party, because of what a party is: a party is the kind of thing that exists when there are people revelling. In this way the nature of a party ‘opens itself up’ to the possibility of being grounded in specific facts concerning revelling.

---

<sup>20</sup> Dasgupta also suggests that we might explain grounding facts in terms of conceptual truths or metaphysical laws, but does not outline these proposals in detail. The former alternative would seem to lead to difficulties similar to those explored in this paper: the Phenomenal Analysis Problem would become a problem with the analysis of phenomenal concepts rather than the analysis of phenomenal properties. The latter model seems to me not very promising, as metaphysical laws are intuitively the kind of things we want to explain. Having said that, Schaffer (2017) proposes explaining all grounding, including the grounding of the mental in the physical, in terms of basic metaphysical laws. Schaffer’s view seems to me unattractive because (i) physicalism is supposed to be the view that all facts are nothing over and above the physical facts, but this would not seem to be the case if we need brute laws to get from the physical to the mental, (ii) contra Schaffer I think it’s pretty plausible that most physical facts can be grounded in something like the way Dasgupta proposes, and hence that we don’t need brute laws to mediate the grounding. I argue this in more detail in Goff 2017: ch. 2. Dasgupta’s discussion starts from a problem Sider (2012) raises with grounding theories of fundamentality, but it would be distracting to explore that here.

Dasgupta proposes we explain Party-Grounding in terms of the following two facts:

*Party-Nature* – A party is essentially such that if there are people revelling then there is a party,

*Revelling* – Rod, Jane and Freddy are revelling.<sup>21</sup>

Note that the entities in the *less fundamental fact* – the fact that there is a party – are doing crucial explanatory work in the explanation of the overall grounding fact. Dasgupta argues, partly through reflection on cases, that we do not get a satisfying explanation of grounding facts from the more fundamental fact alone. For example, it would not be satisfying to answer:

“Why is it the case that the fact that Rod, Jane and Freddy are revelling grounds that fact that there is a party?”

with:

“Because Rod, Jane and Freddy are revelling”

It is only by reference to the nature of parties, to what a party is, that we get a satisfactory explanation of Party-Grounding.<sup>22</sup>

Kit Fine has previously advocated a similar kind of ‘top-down’ direction in the explanation of grounding facts:

...what explains the ball’s being red or green in virtue of its being red is something about the nature of what it is for the ball to be red or green (and about the nature of disjunction in particular) and not something about the nature of what it is for the ball to be red. It is the fact to be grounded that ‘points’ to its grounds and not the grounds that point to what they may ground.<sup>23</sup>

It is not that the less fundamental fact ‘points to’ the *specific facts* which ground it; essential truths concerning parties do not involve specific reference to Rod, Jane and Freddy. Rather the nature of constituents of the less fundamental fact F ‘point to’ some condition which is sufficient for its being the case that F.

Let us call this model of explaining grounding facts ‘grounding via essence’, or GVE. Abstracting from specific cases, we can take it to be committed to the following principle:

---

<sup>21</sup> Dasgupta’s example is in terms of conferences rather than parties, but the substance is the same.

<sup>22</sup> Karen Bennett (2011) and Louis deRosset (2013) try to ground the grounding facts in the fundamental facts. I am persuaded by Dasgupta’s arguments against this strategy, some of which I have outlined above, and which are given in more detail in section VI of his 2014.

<sup>23</sup> Fine 2012.

*Grounding via Essence (GVE)* – For any grounding fact F in which f2 is grounded in f1, F is grounded in the fact that there is a condition C such that (i) a constituent of F2 is essentially such that if C is satisfied F2 obtains, and (ii) F1 logically entails that C is satisfied.<sup>24</sup>

GVE is in tension with Strong Physicalism, as according to GVE the grounding facts are partly grounded in facts about the nature of higher-level entities, rather than being wholly grounded in the fundamental physical facts. Pain-Grounding, for example, is partly grounded in Pain-Nature. The ‘top-down’ direction of explanation means that chains of grounding explanation don’t always move in a downward direction. For this reason Dasgupta rejects Strong Physicalism, and adopts a definition of physicalism according to which certain facts are ‘exempt’ from needing to be grounded in the physical, even if physicalism is true.

Obviously there must be some limit on which facts physicalism ‘allows’ not to be grounded in the physical; it is inconsistent with physicalism, for example, to deny that the facts about consciousness are grounded in the physical. Dasgupta’s view is that physicalism does not require that *facts about essences* are grounded in the physical. This is because, according to Dasgupta, facts about essences are *autonomous*, or not *apt to be grounded*; that is to say they are not the kind of fact for which the question of grounding arises.

I do not have space here to give Dasgupta’s complete defence of the autonomy of facts about essences, which stretches beyond his paper on the definition of physicalism, but I will briefly refer to an analogy he offers to help clarify and motivate the idea.<sup>25</sup> The analogy is between facts which are not apt for grounding and facts which are not apt for causal explanation. The fact that  $2+2=4$  lacks a causal explanation, but not in the sense that the big bang may lack a causal explanation; the fact that  $2+2=4$  is not the kind of fact which requires or admits of causal explanation. By analogy there may be a category of fact which neither requires nor admits of grounding explanation, and essential truths are a plausible candidate. According to Dasgupta, the question ‘What explains the fact that a party is the kind of thing that exists when there are people revelling?’ is ill-posed in something like the way ‘What caused 2 and 2 to equal 4?’ is ill-posed. Nobody who knows what a party is should be troubled by this question.

Having defended the autonomy of facts about essence, Dasgupta offers the following improvement on Strong Physicalism:

---

<sup>24</sup> Goff 2017. GVE is not an analysis of grounding, but a view about how the grounding facts are grounded. Dasgupta does not give an utterly precise definition of what the model of explanation involves in general, but this principle seems to be suggested by his examples.

<sup>25</sup> The argument for the autonomy of facts about essence is continued in Dasgupta 2016.

*Weak Physicalism* – Physicalism is the thesis that all non-physical facts which are *substantive*, i.e. apt to be grounded, are grounded in facts which are either physical or autonomous.<sup>26</sup>

Dasgupta admits that the GVE model of grounding makes it harder for those wishing to defend physicalism about the mind:

...physicalism [given GVE] requires that there are essential connections between mind and body. This will be disappointing to physicalists who hoped that formulating physicalism in terms of ground would rescue them from having to offer tight connections of essence or analysis between mind and body. On the current picture, this hope is dashed.<sup>27</sup>

Assuming the GVE model of grounding, the physicalist is obliged to construe the essences of conscious states in such a way that they are able to explain the grounding of consciousness in the physical. In the next three sections I will develop this challenge, and show how the proponent of the knowledge argument can use it to respond to the Russellian monist.

#### (v) The Phenomenal Analysis Problem

The phenomenal facts are the facts about which individuals have which phenomenal properties. However, I want to focus on consciousness properties themselves rather than the individuals that have them. Therefore, I will abstract from phenomenal facts involving specific individuals and focus on facts involving only quantificational structure and phenomenal properties, e.g. the fact that there is something that feels pain. Call such facts the ‘pure phenomenal facts’. I take it that the pure phenomenal facts are grounded in phenomenal facts involving specific individuals, e.g. the fact that someone is in pain is grounded in the fact that Bill is in pain.<sup>28</sup>

The adoption of GVE obliges the physicalist to hold that conscious states are in a certain sense *analysable*. In this section I will outline the sense of analysis I have in mind, and why the physicalist is obliged to think that phenomenal properties are indeed analysable in this sense. In section VI I will canvas physicalist theories according to which phenomenal properties can be so analysed; and in section VII I will construct an argument for the claim that phenomenal properties are not analysable in the required sense.

---

<sup>26</sup> This is not the final definition Dasgupta ends up with, but it will serve for the purposes of this paper. His final definition is a little stronger, and hence will inherit any difficulties the Phenomenal Analysis Problem raises for Weak Physicalism. I offer my own grounding definition of physicalism in chapter 2 of Goff 2017.

<sup>27</sup> Dasgupta 2016: 586.

<sup>28</sup> For a given fact F, there may be many different facts which each individually ground F: the fact that someone is in pain is grounded in the fact that Bill is in pain, the fact that Sarah is in pain, etc.

As discussed above, we can think of the analysis of a property as a matter of giving a description of its nature, of what it is for something to have the property. Rosen (2010, section 10) gives the general form of an analysis as:

For all  $x$ , for it to be the case that  $Fx$  just is for it to be the case that  $\phi x$ .

In the case of knowledge, one candidate analysis is:

For all  $x$ , for it to be the case that  $x$  knows just is for it to be the case that it is not accidental that  $x$  is right that  $p$  is the case (for some proposition  $p$ ).<sup>29</sup>

It is crucial to note two things about the notion of analysis in play here. Firstly, it is metaphysical rather than linguistic: we are concerned with the definition of properties not the definition of words. Secondly, and perhaps relatedly, it is not to be assumed that an analysis is accessible *a priori*. It is plausible for example that the metaphysical analysis of water is known only empirically.

Putting the general form as Rosen does is in a certain sense limiting, as it restricts us to accounting for the essence of the property in terms of facts about the bearer of that property. Let us call such analyses, in which a property  $P$  is defined in terms of other properties of the bearer of  $P$ , ‘intra-substance analyses’. There is a more radical kind of analysis, in which a property  $P$  is defined in terms of facts concerning individuals which are or at least could be distinct from the bearer of  $P$ . Returning to parties, for it to be the case that there is something that is a party just is for it to be the case that there are people revelling; note that we are not here accounting for the essence of partyhood in terms of some property had by the bearer of partyhood (persons are not parties). We can call analyses of this more radical form ‘inter-substance analyses’, the general form of which is as follows:

For it to be the case that there is an  $x$  such that  $Fx$  just is for fact  $P$  to obtain,  
Where  $x$  is not a constituent of  $P$ .

Orthogonal to the distinction between intra-substance and inter-substance analyses, we can distinguish analyses between properties of the same kind – call these ‘intra-categorical analyses – from analyses between properties which are or could be of different kinds – call these ‘inter-categorical analyses’. The analysis of less determinate colours into more determinate colours is an

---

<sup>29</sup> This analysis is taken from Unger 1968.

example of an intra-categorical analysis; the analysis of chemical kinds into physical kinds is an example of an inter-categorical analysis.<sup>30</sup>

Cross-sectioning these two distinctions we potentially have four distinct categories of analysis, ranging from those which do the least metaphysical bridging – intra-categorical and intra-substance – to those which do the most metaphysical bridging – inter-categorical and inter-substance. Arguably some properties are unanalysable, in which case there will be no non-trivial account of the property's nature.<sup>31</sup> Existence is a plausible candidate for an unanalysable property: plausibly there is no non-trivial way of accounting for what it is for something to exist. There are also mainstream views according to which modal properties, normative properties, and the causal relation, are unanalysable.

It is sometimes claimed that phenomenal properties are unanalysable. However, it is plausible that at least some phenomenal properties can be analysed into other phenomenal properties of the same subject, i.e. that there are phenomenal property analyses of the least metaphysically bridging kind: intra-categorical and intra-substance. For example, for it to be the case that there is a subject S having a colour experience just is for it to be the case that S has some property F such that F is a specific colour experience; for it to be the case that there is a subject S having the disjunctive property feeling pain or pleasure just is for it to be the case either that S feels pain or that S feels pleasure. And there may well be a variety of more subtle analyses, perhaps pain is analysable into an effective component and a qualitative component.

However, armchair reflection does not seem to reveal analyses of phenomenal properties which are either inter-categorical or inter-substance. Prima facie, it's hard to see how what it is for there to be a subject S feeling pain could be analysed into a fact not involving phenomenal properties, or a fact not involving S (or both). Contrast with the case of partyhood. For it to be the case that there is something which is a party just is for it to be the case that there are certain people  $X_1, X_2...X_n$ , such that  $X_1, X_2...X_n$  are revelling; the fact in terms of which partyhood is analysed involves people not parties. No analogous analysis of subject-hood into facts about non-subjects suggests itself, at least on first reflection.

One might take this to be prima facie reason to adopt the following thesis:

---

<sup>30</sup> The distinction between inter-categorical and intra-categorical presumably admits of a certain degree of contextual flexibility, i.e. there are cases of grounding which are correctly classed as inter-categorical in some contexts and intra-categorical in others. This won't matter for the claims I want to make here.

<sup>31</sup> For any property F, a trivial of F would be: For it to be the case that there is an x such that Fx just is for there to be an x such that Fx.



*Minimal Phenomenal Analysis* – Phenomenal properties admit of neither inter-substance nor inter-categorical analysis.

What I have offered so far given is at best prima facie grounds for accepting Minimal Phenomenal Analysis; in the next two sections I will give more sustained attention to the question of whether it is true. But for the moment let us explore the trouble Minimal Phenomenal Analysis, if true, causes for physicalism.

According to the GVE model, in the case of any given grounding fact F, the constituents of the less fundamental fact play a crucial role in explaining F. To indulge in metaphor, the grounding facts ‘reach out’ to, or in Fine’s phrase ‘point to’, the facts that ground them. Partyhood points to its ground in virtue of the fact that it’s in the nature of partyhood that if there are people revelling then there is a party. But if Minimal Phenomenal Analysis is true, it’s hard to see how they could do much reaching or pointing. The phenomenal properties of a given subject can reach out to other phenomenal properties of that subject. But they could not it seems reach out to non-conscious fundamental physical individuals and properties, in such a way to explain the ultimate grounding of the facts about consciousness in facts concerning such individuals and properties.

To try to make this a little more precise, Weak Physicalism in conjunction with GVE entails the following thesis:

*Phenomenal Deflation* – For any actually obtaining pure phenomenal fact Q, there is a condition C, such that (i) there is some constituent E of Q such that it is in the nature of E that if C is satisfied then Q obtains, (ii) the fundamental physical facts logically entail that C is satisfied.

If Minimal Phenomenal Analysis is true, then for any actually obtaining pure phenomenal fact Q of the form <there exists a subject S such that S has phenomenal property P>, the only condition sufficient for Q’s instantiation which could be extracted from the nature of P would be of the form <S has phenomenal property P\*>, in which S is the same subject which has P, but in which P\* may perhaps be a phenomenal property not identical with P. It’s hard to see how such a condition could be logically entailed by the fundamental physical facts. How on earth could a fact about an enormous number of non-conscious fundamental physical entities related in extremely complex ways, logically entail the existence of a macro-level conscious subject?

One way of bolstering this concern is to point out that mereological nihilism seems to be logically coherent. There is no contradiction in the thesis that fundamental particles never compose composite objects, i.e. that there are particles arranged table-wise, planet-wise, etc., but no tables,

planets, etc. Assuming that conscious subjects (at least the conscious subjects we are pre-theoretically committed to) are macro-level entities, the logical coherence of mereological nihilism entails that no fact exclusively concerning micro-level entities entails the existence of a conscious subject.

Thus, assuming that (i) the fundamental physical facts are micro-physical facts and (ii) human conscious subjects are macro-level entities, we can pose the following argument:

### The Incompatibility of Minimal Phenomenal Analysis and Physicalism

*Premise 1* – If physicalism is true, then *Phenomenal Deflation* is true: For any actually obtaining pure phenomenal fact Q, there is a condition C, such that (i) there is some constituent E of Q such that it is in the nature of E that if C is satisfied then Q obtains, (ii) the micro-physical facts logically entail that C is satisfied.

*Premise 2* – If Minimal Phenomenal Analysis is true, then for any actually obtaining pure phenomenal fact Q of the form <there exists a subject S such that S has phenomenal property P>, the only condition sufficient for Q's instantiation which could be extracted from the analysis of P would be of the form <S has phenomenal property P\*>, in which S is a macro-level entity.

*Conclusion 1* – Therefore, if Minimal Phenomenal Analysis is true, then physicalism is true only if the micro-physical facts logically entail the existence of a macro-level entity.

*Premise 3* – It's not the case that the micro-physical facts logically entail the existence of a macro-level entity.

*Conclusion 2* – Therefore, if Minimal Phenomenal Analysis is true, then physicalism is false.

Why does it not follow from the coherence of mereological nihilism that facts about micro-level entities *never* ground the existence of macro-level entities? The proponent of GVE will likely hold that in general the essences of macro-level entities are rich enough to account for their grounding in the micro-level facts. I have argued, for example, that is in the nature of table-hood that facts about tables (or at least table-shaped macroscopic objects) can be analysed in terms of *patterns of penetration resistance among regions of space*, such that it's plausible that the existence of the relevant patterns could be logically entailed by the micro-level facts.<sup>32</sup> Thus, even though the micro-level facts in and of themselves do not logically entail the existence of tables, the micro-level facts

---

<sup>32</sup> Goff 2017: ch. 2.

logically entail a certain condition C, such that the kind table is essentially such that if C is satisfied then there are tables.

The problem in the case of conscious subjects is that, if Minimal Phenomenal Analysis is true, the natures of phenomenal properties are not rich enough to yield a condition C which (i) is logically entailed by the micro-physical facts, and (ii) is sufficient for the obtaining of macro-level facts concerning conscious subjects. Assuming Minimal Phenomenal Analysis, the only condition extractable from the nature of phenomenal properties which is sufficient for the instantiation of those properties concerns the bearers of phenomenal properties. Assuming the bearers of phenomenal properties are macro-level entities, that condition is not logically entailed by micro-level facts.

Rosen has discussed, in more general terms, something like this difficulty for the physicalist, and Fine has suggested two responses.<sup>33</sup> Firstly he suggests that the kind of grounding which obtains between the mental and the physical is *natural* rather than metaphysical. Secondly, he suggests that even if it is not in the nature of a given mental property to ground a connection with a *specific* physical property, it may nonetheless be in the nature of each mental property that it has *some* physical ground.

Given the standard understanding of physicalism, and assuming that natural grounding goes along with natural necessity, Fine's first response would not lead to physicalism as it is normally understood but to property dualism. Although definitions of physicalism wholly in terms of supervenience are out of favour, it is generally agreed that the supervenience of all facts on the physical facts – with the strength of metaphysical necessity – is a necessary condition for physicalism. And there are arguably good reasons for taking it so. One of the main arguments for physicalism is its capacity to reconcile mental causation with the causal closure of the physical in a way that does not lead to problematic overdetermination. There is not space here to fully defend this claim, but it is broadly agreed that physicalism can only do this if it is taken to be the thesis that all facts are metaphysically, and not just naturally, grounded in the physical facts.<sup>34</sup>

Turning to Fine's second response, it is not obviously true that it is in the nature of phenomenal properties that they require some physical ground. Disembodied subjects of experience seem perfectly conceivable, which gives us at least *prima facie* grounds for thinking that the essence of phenomenal properties is compatible with their existing ungrounded.<sup>35</sup> But even if Fine were right

---

<sup>33</sup> Rosen 2010: section X; Fine 2012.

<sup>34</sup> See for example Pereboom 2002 and Bennett 2003. I argue for this in Goff 2017.

<sup>35</sup> I explore the challenge the conceivability of disembodied subjects poses to physicalism in Goff 2010.

that it is part of the nature of phenomenal properties that they have some physical ground, this in itself would hardly be sufficient to explain a grounding connection with the physical. Presumably it is not in the nature of pain that just any old physical state is sufficient for there to be pain. If physicalism is true, there are a limited range of physical states which are capable of grounding pain. If grounding facts are to be explained in terms of the nature of the grounded items, there must be something in the nature of pain which determines which of all possible physical states are the ones capable of grounding pain.

I conclude at this stage that the physicalist (or at least the physicalist who accepts the GVE model of grounding) has to deny Minimal Phenomenal Analysis; she must argue that, contrary to initial appearances, there is a way of analysing phenomenal properties which facilitates the bridging of distinct substances and property kinds. Let us call this the 'Phenomenal Analysis Problem' for physicalism. In the next section I will consider ways in which physicalists are likely to respond to the Phenomenal Analysis Problem, before going on to consider how proponent of the knowledge argument could turn the Phenomenal Analysis Problem into an argument against physicalism.

#### (vi) [Physicalist responses to the Phenomenal Analysis Problem](#)

How might physicalists respond to the Phenomenal Analysis Problem? Following David Chalmers, we can distinguish two forms of physicalism:

- *Type-A physicalism* – Type-A physicalists deny that there is an epistemic gap between the truths of physics and phenomenal truths: Mary in her black and white room would be able to work out all truths about the nature of experience.
- *Type-B physicalism* – Type-B physicalists accept that there is an epistemic gap between the truths of physics and the phenomenal truths: When she is liberated from her black and white room Mary learns a truth (although that truth is just a new way of thinking about properties she already knew about, i.e., type-B physicalists adopt the phenomenal concept strategy discussed earlier in Part I).

Type-A physicalists tend to be analytic functionalists, that is to say, they take mental states to be *causally defined*: to have a red experience, by definition, is to have an inner state which 'plays the red experience role,' i.e. mediates between sensory inputs, behavioural outputs and other mental states in the way that is distinctive of red experiencing.<sup>36</sup> Clearly to accept analytic functionalism is

---

<sup>36</sup> Armstrong 1968, Lewis 1994. Actually, the views of Lewis and Armstrong are slightly more complicated than this. They defend forms of the identity theory according to which we pick out mental states in virtue of their functional role. So strictly speaking they are analytic functionalists not about pain but about the property of being pain. It is an a posteriori matter, on their view, what pain itself is, and hence their response to the

to deny Minimal Phenomenal Analysis. Analytic functionalism entails that, for any phenomenal property, that property can be analysed in terms of causal roles which could be realised by non-conscious individuals which are distinct from the individual instantiating that phenomenal property, e.g. micro-level physical parts of the individual acting in concert. In other words, analytic functionalism is committed to inter-categorical and inter-substance analysis of phenomenal properties. Just as it is a priori that if there are people revelling then there is a party, so according to analytic functionalism it is a priori that if certain entities play the red experiencing role then there is red experiencing.

Moreover, it is plausible that such an essential nature could explain the grounding of the phenomenal in the physical. For the analytic functionalist pain is essentially such that there is something experiencing red if the following condition is met: <there is an entity or entities playing the red experiencing role>, and it is plausible that the physical facts logically entail that this condition is met. A similar story can be told about each phenomenal property, according to analytic functionalism, thus accounting for the truth of Phenomenal Deflation, and thereby paving the way for a GVE-kosher account of the grounding of consciousness in the physical.

Type-B physicalists reject analytic functionalism, adopting instead the phenomenal concept strategy response to the knowledge argument (discussed in Part I). Hence, they would give a very different kind of response to the Phenomenal Analysis Challenge. Such physicalists tend to hold that phenomenal concepts are opaque: that they reveal nothing about the nature of phenomenal states. David Papineau is quite explicit about this:

No doubt there are ways of thinking of things that make certain essential properties a priori knowable. But I take such a priori knowledge to derive from (possibly implicit) compositionality in the relevant modes of thinking, and so not to be associated with the most basic ways in which thought makes contact with reality...When it comes to these basic points of contact, I find it hard to take seriously any alternative to the assumption that our atomic concepts are related to reality by facts external to our a priori grasp, such as causal or historical facts.... I don't recognise any way in which the mind 'captures' something, apart from simply referring to it.<sup>37</sup>

Given that Papineau thinks that phenomenal concepts are atomic concepts, it is clear that he thinks that phenomenal concepts are opaque. Brian McLaughlin is even more explicit:

---

Phenomenal Analysis Problem regarding pain itself would be the same as that offered by the type-B physicalist.

<sup>37</sup> Papineau 2006: 102-6.

Phenomenal concepts....do not conceptually reveal anything about the essential nature of phenomenal properties: they simply name or demonstrate them.<sup>38</sup>

Whilst other type-B physicalists may not be explicit about their commitment to the opacity of phenomenal concepts, it is often implicit in their favoured theories of phenomenal concepts as demonstratives,<sup>39</sup> indexicals,<sup>40</sup> recognitional concepts,<sup>41</sup> or concepts which refer in virtue of facts about teleology or causal connections.<sup>42</sup> If my phenomenal concept of pain is merely a demonstrative, or refers in virtue of facts outside of my a priori grasp, then it is hard to see how it could yield any insight into the essential nature of the state it tracks.

The type-B physicalist, therefore, will not take the analysis of a conscious state to be a priori, as phenomenal concepts reveal nothing of the nature of the states they denote. However, they need not deny that conscious states admit of analysis (recall that metaphysical analyses need not be available a priori). It is an empirical question what the right metaphysical analysis is of any given conscious state, and hence there can be no a priori grounds for denying that conscious states have highly complex physical, or indeed functional, natures suited to explaining the grounding of the phenomenal in the physical.

If the phenomenal concept strategist identifies phenomenal properties with functional properties, then she can give an account of the grounding of the mental in the physical similar to that offered by the analytic functionalist above. The only difference is that the functionally defined essential nature of the phenomenal which explains the grounding connection is not a priori accessible. Alternately, she can identify phenomenal properties with physical properties themselves, leaving no remaining challenge to a physicalist account of consciousness.<sup>43</sup> In either case, the result would be a kind of phenomenal concept strategy for responding to the Phenomenal Analysis Problem.

### (vii) The Phenomenal Analysis Argument

The knowledge argument provides strong support for the thesis that there is an epistemic gap between the physical and the experiential: Mary cannot deduce from the physical facts what it's like to see red. This epistemic gap rules out the analytic functionalist response to the Phenomenal Analysis Problem. And Phenomenal Transparency, a premise in the refined knowledge argument, rules out the phenomenal concept strategist's response to the Phenomenal Analysis Problem.

---

<sup>38</sup> McLaughlin 2001: 324.

<sup>39</sup> Papineau 1993, Perry 2001.

<sup>40</sup> Tye 1995: ch. 6, Lycan 1996, sect. 3.3.

<sup>41</sup> Loar 1990/1997, 2003, Tye 2000, chap. 2, Carruthers 2000, 2004, Perry 2001, Levin 2007a, b.

<sup>42</sup> Papineau 2002, 2007.

<sup>43</sup> At least no challenges arising from the nature of consciousness.

Furthermore, these two premises, if true, render Minimal Phenomenal Analysis extremely difficult to deny, as I will now explain.

If Phenomenal Transparency is true, then for any phenomenal property F the nature of F is a priori accessible (for anyone possessing a phenomenal concept of F, in virtue of possessing a phenomenal concept of F). If there is some inter-substance or inter-categorical analysis of what it is for something to have F, then that analysis will be a priori accessible. But the only a priori accessible analyses of the essential nature of phenomenal properties ever proposed in the history of philosophy are either intra-substance or intra-categorical (usually both) or those offered by proponents of analytic functionalism.<sup>44</sup> The former kind of account does not help with the Phenomenal Analysis Problem and the epistemic gap rules out accounts of the latter kind.

To put it another way, to avoid Phenomenal Analysis Problem the physicalist would seem to require some kind of inter-substance and inter-categorical analysis of phenomenal properties. If Phenomenal Transparency is true, that analysis must be available a priori. If there is an epistemic gap, then we have good reason to think that there are no a priori accessible analyses of phenomenal properties which are either inter-substance or inter-categorical, as all extant proposals are inconsistent with the epistemic gap. Of course it could be that nobody has yet uncovered the correct inter-substance and inter-categorical analysis of phenomenology, but unless the Russellian monist can at least gesture at such an analysis, the hope that there is one seems like an unwarranted leap of faith.

Thus, we reach the following argument against physicalism (in which a ‘non-minimal analysis’ is one which is either inter-categorical or inter-substance):

### The Phenomenal Analysis Argument

*Premise 1* – If there is an epistemic gap, then either Minimal Phenomenal Analysis is true or there is a non-minimal analysis of phenomenal properties which is not a priori accessible.

*Premise 2* – If Phenomenal Transparency is true, then it’s not the case that there is a non-minimal analysis of phenomenal properties which is not a priori accessible.

*Premise 3* – Phenomenal Transparency is true

---

<sup>44</sup> There is one exception. Panqualityists (Coleman 2014, 2016) offer an inter-categorical analysis of phenomenal properties in terms of awareness of qualities. In its reductive forms, it also offers an inter-substance analysis of phenomenal properties by involving a functionalist analysis of awareness. Whereas the knowledge argument defeats standard forms of analytic functionalist, the panqualityist analytic functionalist can maintain there is an epistemic gap as pre-liberation Mary will lack an acquaintance with red qualities. Thus, panqualityism would seem to be one way around this argument and would need to be rejected on independent grounds. I try to do this in Goff 2017: ch. 6.

*Premise 4* – There is an epistemic gap (supported by the Mary/Marianna thought experiments).

*Conclusion 1* – Therefore, Minimal Phenomenal Analysis is true.

*Premise 5* – If Minimal Phenomenal Analysis is true, then physicalism is false (derived from the argument for the incompatibility of physicalism and Minimal Phenomenal Analysis defended in section V).

*Conclusion 2* – Therefore, physicalism is false.<sup>45</sup>

This argument has the potential to rule out not only standard forms of physicalism, but also Russellian monist forms. This is because Minimal Phenomenal Analysis is just as problematic for the Russellian monist as it is for the physicalist and for exactly the same reasons. Just like the standard physicalist, the Russellian monist – at least on a standard understanding of it – tries to ground the phenomenal facts in complex micro-level physical facts. It's just that for the Russellian monist, the complex physical facts to which consciousness is being reduced have a deep nature, without which they would be unable to ground consciousness. But none of the difficulties I outlined in describing the Phenomenal Analysis Problem go away if physical facts have a deep nature, as these problems arise from the nature of phenomenal rather than physical properties. Specifically, the problem of trying to make sense of how unanalysable phenomenal properties 'reach out' to complex physical facts is made no easier by the supposition that those complex physical facts have a deep nature.<sup>46</sup>

I happen to think, or at least hope, that there is a way for the Russellian monist to resist this argument. I have been assuming a 'bottom-up' form of Russellian monism according to which all facts are grounded in facts at the micro-level. However, there are 'priority monist' forms of

---

<sup>45</sup> This argument is closely related to the argument I present in Goff 2015b and the 'subject irreducibility problem' I present in Goff 2017: ch. 8. It might be thought objectionable that I am giving two names – 'the subject-irreducibility problem' and 'the phenomenal analysis problem' – to essentially the same difficulty. But I'm not sure of how best to formulate the problem, and so I think it may be useful to retain the two slightly different forms (with distinct names) presented here and in Goff 2017. The one presented in this paper was formulated earlier than the one in my book. But as I finish this paper as my book goes to publication, I'm inclined to think this is the better version.

<sup>46</sup> I am here targeting only 'constitutive' forms of Russellian monism (Chalmers 2015) according to which facts about human and animal consciousness are grounded in facts about more fundamental forms of consciousness. Non-constitutive forms of panpsychism avoid these worries, on account of not needing to make sense of the grounding of human and animal consciousness. Panpsychist forms of Russellian monism (Strawson 2006, Roelofs 2015, 2016, Goff 2017) would not need an inter-categorical analysis of phenomenal properties, as on this view ordinary consciousness is grounded in more basic states of consciousness. However, if the panpsychist thinks that facts about human consciousness are grounded in micro-level phenomenal facts, then she is obliged to offer an inter-substance analysis of phenomenal properties. I consider some possibilities in Goff 2017: ch 8, arguing that none is adequate. As I say in the main text, I ultimately think the Russellian monist can avoid this difficulty by adopting priority monism.



Russellian monism according to which all facts are grounded in facts concerning the whole universe. In Goff 2017 and Goff forthcoming I argue that the specific form of grounding that obtains in the priority monist universe, which I call ‘grounding by subsumption’, does not fit the GVE model, and hence that the priority monist avoids the need to offer an analysis of phenomenal properties. However, this paper is already too long so I won’t explore this possibility here.

## Conclusion

The knowledge argument makes a strong philosophical case for there being an epistemic gap between the physical facts and the facts about experience. That epistemic gap does not, in and of itself, entail the falsity of physicalism. However, if we accept certain plausible theses concerning phenomenal concepts (that they are essence-revealing) and grounding (that it fits the GVE model), then we can build on that epistemic gap and create a formidable challenge both to conventional physicalism and to Russellian monism.<sup>47</sup>

## References

- Alter, T. & Nagasawa, N. (Eds.) 2015. *Consciousness and the Physical World*, Oxford University Press.
- Armstrong, D. M. 1968. *A Materialist Theory of Mind*, Routledge and Kegan Paul.
- Alter, T. & Walter, S. 2007. (Eds.) *Phenomenal Concepts and Phenomenal Knowledge*, Oxford/New York: Oxford University Press.
- Bennett, K. 2003. ‘Why the exclusion problem seems intractable, and how, just maybe, to tract it’, *Noûs* 37, 471–97.
- Bennett, K. 2011. ‘By our bootstraps,’ *Philosophical Perspectives* 25, 27–41.
- Brüntrup, G. & Jaskolla, L. (Eds.), 2016, *Panpsychism*, New York: Oxford University Press.
- Carruthers, P. 2000. *Phenomenal Consciousness: A Naturalistic Theory*, Cambridge: Cambridge University Press.
- Carruthers, P. 2004. ‘Phenomenal Concepts and Higher-Order Experiences.’ *Philosophy and Phenomenological Research* 68: 2, 316–36.

---

<sup>47</sup> In Goff 2017 I argue that the knowledge argument on its own is unable to defeat physicalism. There is no real difference in substance to what I say here. What the knowledge argument establishes is an epistemic gap between the physical and the phenomenal, and as I hope to have shown a lot more needs to be done to build that into a metaphysical gap inconsistent with the truth of physicalism.

- Chalmers, D. J. 2015. 'Panpsychism and panprotopsychism,' in Alter & Nagasawa 2015.
- Coleman, S. 2014, "The Real Combination Problem: Panpsychism, Micro-Subjects, and Emergence", *Erkenntnis*, 79: 1, 19–44.
- Coleman, S. 2016. 'Panpsychism and Neutral Monism: How to Make Up One's Mind,' in Brüntrup & Jaskolla 2016, 249–282.
- Damnjanovic, N. 2012. 'Revelation and Physicalism,' *Dialectica* 66 1: 69-91.
- Dasgupta, S. 2014. 'The possibility of physicalism,' *The Journal of Philosophy* 111: 9, 557-592.
- Dasgupta, S. 2016. 'Metaphysical rationalism,' *Nous* 50: 2, 379-418.
- deRosset, L. 2013. 'Grounding Explanations,' *Philosophers' Imprint*, 1–26.
- Diaz-Leon, E., 2010. 'Can Phenomenal Concepts Explain the Epistemic Gap?' *Mind* 119: 476, 933–51.
- Fine, K. 2001. 'The Question of Realism,' *Philosophers' Imprint*, 1, 1–30.
- Fine, K. 2012. 'Guide to Ground,' in F. Correia and B. Schnieder (Eds.) *Metaphysical Grounding: Understanding the Structure of Reality* Cambridge: Cambridge University Press, 37–80.
- Goff, P. 2010. 'Ghosts and sparse properties: Why the physicalist has more to fear from ghosts than zombies,' *Philosophy and Phenomenological Research* 81: 119-39.
- Goff, P. 2011. 'A posteriori physicalists get our phenomenal concepts wrong,' *Australasian Journal of Philosophy*, 89: 2, 191-209.
- Goff, P. 2015a. 'Real acquaintance and physicalism,' in P. Coates & S. Coleman (Eds.) *Phenomenal Qualities: Sense, Perception, and Consciousness*, Coates, P. and Coleman, S. (Eds.), Oxford University Press.
- Goff, P. 2015b. 'Against Constitutive Russellian monism,' in Alter & Nagasawa 2015b, 370–400.
- Goff, P. 2017. *Consciousness and Fundamental Reality*, New York: Oxford University Press.
- Goff, P. forthcoming. 'Cosmopsychism, micropsychism, and the grounding relation,' in W. Seager (Ed.) *The Routledge Handbook of Panpsychism*, London: Routledge.
- Levin, J. 2007a. 'Nagel vs. Nagel on the Nature of Phenomenal Concepts.' *Ratio* 20: 3, 293–307.
- Levin, J. 2007b. 'What is a Phenomenal Concept?' in Alter & Walter 2007.

- Lewis, D. 1994. 'Reduction of Mind,' in S. Guttenplan (Ed.) *Companion to the Philosophy of Mind*, Blackwell.
- Loar, Brian. 1990/1997. 'Phenomenal States.' Originally published in J. Tomberlin (Ed.) *Philosophical Perspectives 4: Action Theory and Philosophy of Mind*, Atascadero, CA: Ridgeview, 81–108. Reprinted in substantially revised form in N. Block, O. Flanagan & Güüven Güüzeldere (Eds.) *The Nature of Consciousness: Philosophical Debates*, Cambridge, MA: MIT Press, 597–616.
- Loar, B. 2003. 'Qualia, properties, modality,' *Philosophical Issues* 13/1: 113-29.
- Lycan, W. G. 1996. *Consciousness and Experience*, Cambridge, MA: MIT Press.
- McGinn, C. 1989. 'Can we solve the mind-body problem?' *Mind* 98: 391, 349-66.
- McLaughlin, B. P. 2001. 'In defence of new wave materialism: A response to Horgan and Tienson,' in C. Gillett & B. Loewer (Eds.) *Physicalism and its Discontents*, Cambridge, New York: Cambridge University Press.
- Nida-Rümelin, M. 1996. 'What Mary couldn't know.' in T. Metzinger (Ed.) *Phenomenal Consciousness*, Schoenigh: Paderborn.
- Nida-Rümelin, M. 1998. 'On Belief About Experiences: An Epistemological Distinction Applied to the Knowledge Argument,' *Philosophy and Phenomenological Research* 58: 1, 51–73.
- Nida-Rümelin, M. 2007. Grasping Phenomenal Properties, in Phenomenal Concepts and Phenomenal Knowledge, in Alter & Walter 2007, 307–36.
- Papineau, D. 1993. *Philosophical Naturalism*, Oxford: Blackwell.
- Papineau, David, 1998. 'Mind the Gap,' *Philosophical Perspectives*, 12: S12, 373–89.
- Papineau, 2002. *Thinking about Consciousness*, Clarendon Press: Oxford.
- Papineau, D. 2006. 'Comments on Galen Strawson,' in A. Freeman (Ed.) *Consciousness and its place in Nature*, Academic Imprint.
- Papineau, D. 2007. 'Phenomenal concepts and perceptual concepts,' in Alter & Walter 2007.
- Pereboom, D. 2002. 'Robust Nonreductive Materialism,' *Journal of Philosophy* 99, 499–531.
- Pereboom, D. 2011. *Consciousness and the Prospects for Physicalism*, New York: Oxford University Press.
- Perry, H. 2001. *Knowledge, Possibility and Consciousness*, Cambridge, MA: MIT Press.

- Roelofs, L. 2015. *Combining Minds: A Defence of the Possibility of Experiential Combination*, Ph.D. thesis, University of Toronto. [Roelofs 2015 available online]
- Roelofs, L. 2016, "The unity of consciousness, within subjects and between subjects", *Philosophical Studies*, 173(12): 3199–3221.
- Rosen, G. 2010. 'Metaphysical Dependence: Grounding and Reduction,' in B. Hale & A. Hoffmann (Eds.) *Modality: Metaphysics, Logic, and Epistemology*, Oxford: Oxford University Press.
- Russell, B. 1927. *The Analysis of Matter*, Kegan Paul.
- Schaffer, J. 2009. 'On What Grounds What,' in D. Chalmers, D. Manley & R. Wasserman (Eds.) *Metametaphysics: New Essays on the Foundations of Ontology*, Oxford: Clarendon Press, 2009, 347–83.
- Schaffer, J. 2017. 'The ground between the gaps,' *Philosophers' Imprint* 17: 11.
- Sider, T. 2012. *Writing the Book of the World*, Oxford: Oxford University Press.
- Strawson, G. 2006. 'Realistic materialism: Why physicalism entails panpsychism,' 13: 10-11.
- Taylor, J. H. 2013. 'Physicalism and phenomenal concepts: Bringing ontology and philosophy of mind together,' *Philosophia* 41: 1283-97.
- Trodden, K. 2013. 'An introduction to grounding,' in M Hoeltje, B Schenieder & A. Steinberg *Varieties of Dependence*, Basic Philosophical Concepts, Philosophia Verlag, 97-122.
- Tye, M. 1995. *Ten Problems of Consciousness*, Cambridge, MA: MIT Press.
- Tye, M. 2000. *Consciousness, Color, and Content*, Cambridge, MA: MIT Press.
- Unger, P. 1968. 'An analysis of factual knowledge,' *Journal of Philosophy* 65: 6, 157-70.
- White, S. L. 2007. 'Property dualism, phenomenal concepts, and the semantic Premise,' in Alter & Walter 2007, 210-48.