

A DILEMMA FOR RUSSELLIAN MONISTS ABOUT CONSCIOUSNESS*

Adam Pautz
First Draft of December 15th, 2015
Brown University

[Russellian monism] captures of virtues of both
[dualism and physicalism] and the vices of neither.
---David Chalmers (2015)

Quiddity switching is what turns the lights on and off.
--John Hawthorne (2006)

Standard physicalism about consciousness faces a well-known problem. We cannot understand how soggy grey matter should necessitate technicolor phenomenology. In fact, we can easily conceive of “Zombie cases” and “altered qualia cases” where the facts about consciousness vary independently of the physical facts. Call this the *conceivability problem*. This suggests dualism. But dualism about consciousness has its own well-known problem: it is a decidedly uneconomical view of the world. Call this the *complexity problem*.

I want to look at an intriguing, non-standard form of physicalism that tries to pave a middle way between dualism and standard materialism. This non-standard form of physicalism is called *Russellian monism* because it depends crucially on Bertrand Russell’s idea that we are *ignorant* of the “intrinsic nature” of the physical world. Recently, David Chalmers (2015) has developed a strong case that Russellian monism should be considered a leading solution the mind-body problem.¹

I will explain Chalmers’ argument in detail soon, but here is the gist. Russellian monism is, as I said, a broadly physicalistic and therefore “monistic” view of the world. So Russellian monists, he thinks, avoid the complexity problem faced by dualism. Further, because Russellian monists say we are irremediably ignorant of the nature of the physical world, they have a nice response to the conceivability problem that besets more standard forms of physicalism. Indeed, we will see that their response actually allows them *agree with* dualists that the conceivability of certain “Zombie” and “altered qualia” scenarios proves their possibility, while *retaining* a form of physicalism. That is why Chalmers says that Russellian monism “captures of virtues of both [dualism and physicalism] and the vices of neither”. This is his *Hegelian synthesis argument* for Russellian monism.

* **Note to readers/browsers:** This paper is the basis of a talk I presented at the “Metaphysics at the Ranch” conference, at CUNY, and at the University of Cambridge. It is a very rough and very wordy first draft. If you are familiar with Russellian monism you can skip all of section 1. I plan to shorten it and revise it quite a bit. So if you have any comments – big or small – please let me know:
adam_pautz@brown.edu

¹ Chalmers’s discussion is indebted to Stoljar 2001. See also Maxwell 1978 and Lockwood 1989.

Unfortunately, my aim here is negative. My main claim is that the Hegelian synthesis argument for Russellian monism fails. Maybe Russellian monism can avoid one big problem: the conceivability problem. But there are other profound problems with physicalism and dualism, even if some of them haven't received as much attention. I will argue that, if we look at these other problems, we can see that Russellian monism doesn't avoid them. In fact, depending on the version, Russellian monism *combines* these problems and makes them *even more intractable*. I should say at the outset that the problems for Russellian monism I will be developing are *distinct from* the familiar ones you might be expecting. In particular, they are all different from the various versions of the "combination problem".

My criticism of the Hegelian synthesis argument will proceed by way of a dilemma. My plan will be as follows. First, I will (in §1) explain Russellian monism and its virtues. I will also introduce the distinction that will form the basis of my dilemma. This is the distinction between *reductive* and *primitivist* forms of Russellian monism. Roughly, *reductive* Russellian monism is akin to reductive physicalism of the kind defended by J. C. Smart, David Armstrong, David Lewis, Ted Sider, among others.² It just adds the Russellian idea that we are ignorant of the "intrinsic nature" of the physical. By contrast, *primitivist* Russellian monism is a special version of "primitivist physicalism", recently discussed by John Hawthorne, Terry Horgan, Gideon Rosen, and others.³ It is akin to G. E. Moore's view on goodness. On a natural way of developing Moore's view, goodness is a "simple" or "primitive" non-natural property, but its instantiation is always "grounded in" the instantiation of natural properties. Similarly, Russellian monists could hold that experience properties are "simple and irreducible", but also hold that their instantiation is always grounded in the instantiation of the "unknowable" physical properties, by way of brute, inter-level "grounding laws" – "brute" in the sense that they cannot be derived from any more basic truths.⁴

After distinguishing between these versions of Russellian monism, I will develop my dilemma. I will start with *reductive* Russellian monism (§2). As everyone knows, "reductive physicalism" in general faces various problems, even if it is in other ways very attractive and also has very resourceful adherents (Lewis, Sider, Dorr). But I will argue that, because of their unique commitments, reductive Russellian monists face *even more intractable* problems. It is an untenable position. Next (§3) I will show that that *primitivist* Russellian monism can avoid all of my problems for reductive Russellian monism. So the way to be a Russellian monist, if

² Lewis described himself as a "reductionist" (e. g. Lewis 1994) but never said what reduction amounts to. See Hall 2010 for a very helpful discussion.

³ For discussion of this kind of "primitivist physicalism", see Hawthorne (2006, p. 206), Horgan (2010), Rosen (2010). Terminology can be confusing here. On this type of view, experience-properties are "primitive" in the sense that *they don't have bi-conditional real definitions in other terms*; but in another sense they are *not* "primitive", since their instantiation is always *grounded in the instantiation of other properties*. (See Rosen 2010 n. 2 and sect. 11 for a similar distinction.) I will use "primitive" in the former sense. This view may be strange but it is not incoherent. In fact, there is a well-known parallel view of color, called "color primitivism", according to which colors are primitive properties but their instantiation by things is grounded in the instantiation by those things of complex reflectance properties (or whatever). See Chalmers 2006b, p. 67; and Byrne and Hilbert 2007 for discussion.

Hawthorne and Horgan take this general type of picture "emergentist". But since "emergence" has some many meanings (see Barnes (2012), Chalmers (2006a), and Wilson (1999)), it seems best to avoid this term.

⁴ On a Finean view, the grounding laws might be derived from more basic essentialist truths. I will criticize this essence-based version of primitivist Russellian monism at pp. 36-37.

at all, is to be an primitivist Russellian monist. But primitivist Russellian monism faces a new set of problems. In fact, I will argue that it is just as objectionable as dualism. Where dualism requires special, brute psychophysical laws, primitivist Russellian monism requires special, brute “grounding laws”. Despite the recent enthusiasm for grounding in metaphysics, these brute “grounding laws” are no better than the brute psychophysical laws of dualism. Finally, I will argue that, in any version, Russellian monism shares a strange and overlooked puzzle for dualists about “psychophysical luck” (§4).

The conclusion I will draw is that there is no version of Russellian monism that is clearly superior to the standard options, contrary to Chalmers’s Hegelian synthesis argument. The point of the discussion is not to show that Russellian monism is *false* but that there is no strong argument for it.

1. What is Russellian Monism about Consciousness?

1.1 The basic idea

My first order of business is to explain Russellian monism in greater detail. Afterwards (§1.2) I will make the distinction between reductive and primitivist forms of Russellian monism, which will be the basis of my dilemma.

I will introduce Russellian monism with a fanciful analogy. Imagine a “pixel” world containing little *yellow pixel people*. These yellow pixel people cannot observe individual pixels that make them up: they can only track “holistic patterns” in the pixels. But they are smart. They come up with a “physics” of their world which posits unobservable pixels. They don’t know all that much about these hypothetical pixels. They know that



the pixels stand in spatial relations and that they have a few monadic properties, X, Y and Z, that obey certain strict laws (perhaps akin to the laws in some version of Conway’s *Game of Life*). So, they only know this kind of purely “structural description” of the micro-pixels that are at the foundation of their world. They have no idea about the “intrinsic character” of the pixels.

This creates for them a kind of explanatory gap. What is the relationship between these structural facts and their own *glowing, yellow pixel faces*? As long as they stick to the purely structural facts, they are mystified. How is it that from these color-less structural facts comes vibrant yellow? To them this is just as unaccountable as the appearance of the Djinn when Aladdin rubbed his lamp. Some of them run a conceivability argument. The purely structural facts about the pixels couldn’t possibly necessitate the appearance of their own yellow pixel facts, they insist, since it’s conceivable that the purely structural pixel facts should be the same and yet there be no *yellow pixel faces*.

However, some of the yellow pixel people – the *Russellian pixelists* - realize that there may be a solution. Maybe there is more to the unobservable pixels they have posited than the “structural facts” about them. Maybe there are further facts about the identities of the properties that figure in the relevant simple laws. In particular, maybe one of them is *micro-yellowness*. That is, maybe chromatic quality is at the micro level as well as the macro level. If so, there is in fact no explanatory gap, and the conceivability argument is founded in ignorance. If only the pixel people knew that one of the pixel properties was *micro-yellowness*, then they would “see” that

the micro pixel properties, *in combination* with the right spatial-structural facts, *must* yield their own *yellow* pixel faces at the macro level.

Russellian monists diagnose our actual situation in much the same way. Let me explain its main tenets.

The first commitment of Russellian monism is *physicalism*. Russellian monism, as I will be understanding it here, is a form of physicalism – albeit a somewhat unusual form. Roughly, physicalism says that there is a certain set of “basic” properties and relations. They include the fundamental physical properties and relations, as well as certain topic-neutral properties and relations. The pattern of instantiation of all other properties and relations is wholly determined by the pattern of instantiation of these “base properties”. Russellian monists accept this basic idea. Compare how, in the pixel world, the pattern of instantiation of all properties and relations is determined by the pixel-colors and their spatial arrangement.

Next, *quidditism*. Roughly, *quiddities* are properties play certain fundamental nomic roles, where other properties might have played these same nomic roles. In the pixel world, *being black* and *being yellow* are the micro-quiddities. In other pixel worlds, other micro-colors could have played exactly the same nomic roles that they play. The thesis of quidditism about our world is exactly analogous. The idea is that the properties that in our world play the fundamental nomic roles specified by the fundamental physical laws (“the mass-role”, “the charge-role”, etc.) are *quiddities*. They have the following property: they play fundamental nomic roles that could have been played by other properties.

If you think mass is what plays the mass-role, and charge is what plays the charge-role, and so on, then masses and charges and so on *just are themselves the ultimate quiddities*. And, if physics is on the right track, there is only a *handful of types* of fundamental quiddities.

Following Chalmers (2015), we can explain quidditism a bit more exactly by introducing a helpful bit of terminology. Let the *structural terms* be defined by just a list: they are logical, mathematical, nomic, and spatiotemporal terms. (The distinction between structural terms and the rest is very important for Russellian monism; different versions result from different ways of drawing the drawing the line. See Stoljar and Soames for important discussions.) Chalmers says that, if one uses a Ramsey sentence to characterize fundamental physics, it will only contain structural terms. It will be a super complex *structural description* along the following lines:

(∃X1)(∃X2)(∃XN)(so-and-so fundamental individuals (particles, fields, whatever) instantiate X1, so-and-so fundamental individuals instantiate X2, . . . so-and-so stand in such-and-such spatial-temporal relations, and X1 and X2 figure in such-and-such fundamental laws)

The thesis of quidditism is that this existential generalization has certain witnesses in our world and different witnesses in other worlds. The witnesses in our world are the fundamental quiddities of our world.

The next tenet of Russellian monism is *quidditistic ignorance*. This thesis is hard to start clearly. But roughly it is the thesis that the quiddities are not just numerically distinct from one another. Each has a “substantial character” of some sort. The idea of quidditistic ignorance is that in *our* world we cannot know truths

of a certain kind: we cannot know facts about what are the “substantial characters” of the properties that play the fundamental nomic roles – the mass-role, the charge-role, and so on. For instance, maybe the charge-role is played by yellowness and the mass-role is played by blueness. Or maybe these roles are played by little experiences, so that micro-objects have the gift of sentience! We will never know *truths like this*. As Russell wrote in *The Analysis of Matter*, “we know nothing of the intrinsic quality of the physical world, . . . we know the laws of the physical world, in so far as these are mathematical, pretty well, but we know nothing else about it”.

I count the quiddities, if such there be, as “physical”. After all, their evolution obeys precisely mathematical laws of physics, and they are the same for the conscious and unconscious parts of nature. In fact, on one way of thinking, the quiddities *just are* the charges, masses, and so on. That is why I count Russellian monism a form of physicalism. (I take this to be an uninteresting verbal issue. If you disagree, you can take this as a *stipulation* about how *I* will use ‘physical’ in this paper.) It just a version of physicalism that holds that we are irremediably ignorant of the complete truth about the physical world. We can grasp and even know the *structural* physical facts, but not the *quidditistic* physical facts.

Now some standard physicalists *accept* quidditism and quidditistic ignorance, for instance David Lewis in his paper “Rameseyan Humility”. So these doctrines are not really what set apart Russellian monists from standard materialists like Lewis. What does is a commitment to the final main thesis of Russellian, the thesis of *consciousness as quiddity-involving*.

What does this mean? Let’s begin with the opposite idea of a property being *quiddity-neutral*. The Russellian monist will say that as a rule *almost* all high-level properties that we humans talk about are *quiddity-neutral*. For instance, *being a mountain* is quiddity-neutral. To see this, consider a world that has an identical structural description to our world but where the quiddities are different. Call this a *structural duplicate*. Surely in this world there are still mountains! After all, the quidditistic difference would be *undetected by humans*, even with the aid of our most powerful microscopes. The fact that there are mountains in this world is not even partially grounded in what the quiddities are; it is grounded in purely structural facts.

Now I can explain the distinctive thesis of Russellian monism, *consciousness as quiddity-involving*. On Russellian monism, our conscious properties, are quiddity-involving in two ways.

First of all, their instantiation is, at least partly, grounded in quidditistic facts about what the quiddities are, not merely “structural facts”. So, for instance, the micro-parts of your brain certain quiddities – those that in fact play the mass-role, the charge role, etc. It is partly by virtue of this that you have a pain, and see red, and so on. Moreover, according to Russellian monists, if a super-intelligence only knew the characters of the quiddities that play the fundamental nomic roles, as well as the structural facts about the world, then she could *a priori* deduce that you have these macro-level experiences. This makes Russellian monism a version of *a priori physicalism*.

The second respect in which conscious properties are quiddity-involving takes a bit more explanation. Russellian monists assume we have a basic grip on the idea of a “phenomenal property”, and they make a fundamental distinction between *quiddities that are phenomenal* and *quiddities that are non-phenomenal*. Phenomenal properties fall into two sorts: *experiences* and *qualities*. Examples of qualities

include red and middle-C. They can be experienced but they are not experiences. It's hard to explain the distinction between phenomenal and non-phenomenal quiddities. It's just supposed to be something you get. Now the second respect in which our macro-level conscious properties are quiddity-involving is this. They are not radically multiple realizable with respect to what the quiddities are. It is not the case that the quiddities in your brain help ground your current experiences, *but that any other quiddities could have done so just as well*. On the contrary, if your macro-level conscious properties are grounded in micro-quiddities, those micro-quiddities must be *phenomenal* quiddities. Non-phenomenal quiddities, in any combination, can't ground the kind of macro-experiences that we humans have.

Now, of course, in our world the quiddities (masses, charges, etc.) are widespread. They are not only instantiated by the subatomic parts of our brains; they are also instantiated by the subatomic parts of rocks. But unlike me and you, *composite objects like rocks presumably don't themselves see red or feel pain*. Even if experiences are everywhere at the *micro level*, they are not everywhere at the *macro-level*. So, although they haven't really addressed this issue in detail, Russellian monists must say that *structure* also matters for having these specific conscious experiences. For instance, in one version of Russellian monism, *seeing red* might be fully grounded in, or perhaps is even identical with, the property *having micro-parts with the "right" kinds of quiddities arranged brain-state-B-wise*. This is a property that you and I have but that rocks don't have, so this explains why we but not rocks have macro-level experiences of red. In short, on Russellian monism, seeing red is a bit like *having a yellow pixel-face* in pixel world 2. It is a matter of more than having parts with the right quiddities; it involves having those quiddities "arranged" in the right way.

It is the thesis of quiddity-involvement that sets apart Russellian monism from standard physicalism. Standard physicalists deny both of the two theses that make up quiddity-involvement.

Now I am ready to explain the second big payoff of Russellian monism: it provides an answer to the conceivability problem about consciousness in our world. The answer is analogous to the answer to the "conceivability problem" about yellow pixel faces in the case I started with.

The conceivability problem is that standard physicalism about consciousness doesn't jive with the easy conceivability of "Zombie cases" where the structural-physical facts are the same but where consciousness is totally absent. *It just seems obvious that consciousness is modally independent of the structural facts*. The easy conceivability of such cases suggests that they are possible, which refutes standard physicalism.

The standard response among standard physicalists is to hold that the necessary connection between the physical facts and the facts about conscious experience is deeply *opaque*. That is, when it comes to Zombies conceivability – even ideal conceivability – doesn't entail possibility.

But this kind of "opaque physicalism" has drawbacks. For one thing, we rely on conceivability all the time as proof of possibility. Why should our standard modal reasoning fail in the special case of consciousness (Stoljar, Levine & Papineau reviews)? For another thing, Chalmers argues that, if we have a basic grip on *ways thing could be* that goes beyond ways things could conceivably be, then it is unclear why the conceivability of things being a certain way should be any evidence at all that they could be that way (but see Rosen). As he puts it, "Why should there not

be just one metaphysically possible world, or thirty-seven?” Finally, Chalmers argues that necessary truths (framed in non-Twin-Earthable vocabulary) are generally *a priori transparent*. So it would be weird if the necessary link between the physical and the phenomenal were an exception, being by contrast radically opaque. We might call this the *singularity problem* for opaque physicalism.⁵

Russellian monism provides an answer to the conceivability argument that avoids such problems. The idea is simple. Since Russellian monism (unlike standard physicalism) holds that consciousness is *quiddity-involving*, in order to successfully refute it with a conceivability argument, we would have to conceive the structural *and* quidditistic physical facts being the same, with consciousness completely absent. In short, we would have to conceive of *quiddity-identical Zombies*. But, given *quidditistic-ignorance*, we cannot clearly conceive such a scenario, since we don't know the truths about what quiddities plays the fundamental nomic roles. Therefore, when used against Russellian monism, the conceivability argument fails

⁵ Schaffer (“The Ground Between the Gaps”, MS) defends a version of opaque physicalism that is meant to avoid the singularity problem. On his view, the opaque grounding of consciousness in the physical is *not* exceptional, because the grounding of *every other* element of the manifest image in the physical is *opaque in just the same way*. For instance, it is just as *radically opaque* that, if there are so and so atoms bonded to each other, then this grounds the existence of a molecule with so and so properties. (As I read him, Schaffer doesn't just think we cannot *a priori know* such a grounding principle; we also can't even any *less-than-clinching* immediate *a priori justification* for *believing* such a grounding principle: for if we did, then, since we lack such a justification for believing physical-to-consciousness grounding principles, we would be back with the singularity problem.) On this view, such grounding claims are, like *contingent laws of nature*, totally lacking in *a priori* support, that is, radically opaque. However, I think that Schaffer's position faces an enormous epistemic problem. In a nutshell, the problem is this: once you accept this view, then it becomes very hard to see how we might figure out, in a given case, whether we are dealing with *grounding connections* or *contingent laws of nature*. Here are some examples. (1) Suppose that (contrary to fact) we have strong evidence that the connection between the initial conditions and the subsequent states of the world is deterministic. Now consider two hypotheses. Hypothesis 1 – the standard hypothesis – says that there are contingent laws of nature that govern how the world unfolds given the initial conditions. Hypothesis 2 instead is the wild hypothesis that the initial state of the universe *grounds* the totality of future states – so that the initial state of the universe is fundamental and all subsequent states are entirely non-fundamental! Now, *if*, unlike Schaffer, we thought that true grounding connections are generally *not* radically opaque, but must enjoy at least some modicum of *a priori* support, then we would have reason to rule out Hypothesis 2 (for its grounding laws enjoy no *a priori* support). But if we think, with Schaffer, that they are generally radically opaque, then we no longer have this way of ruling out Hypothesis 2. (Indeed, since Schaffer think that only fundamental entities add to ontological complexity and grounded entities don't – see footnote 38 of this paper – he must say that there is a strong reason to *accept* Hypothesis 2 *over* Hypothesis 1, namely, that it is ontologically very simple, because it reduces our fundamental ontology to the initial conditions!) Now, in response to this problem, Schaffer has reminded me that he thinks that there are various *differences* between grounding connections and laws of nature: for instance, the latter but not the former come in probabilistic forms. (And he could say that we have a primitive grip on the difference between *grounding connections* and mere *law-like connections*.) But this is not to the point: for these differences don't help us to *decide between* Hypothesis 1 and Hypothesis 2, which is the problem. (2) We take it that the connection between individual H and O atoms coming together, and the existence of an H₂O molecule with various properties, is a *grounding connection*. But another hypothesis is that it is a *contingent law of nature* (*and, in another world, they could come together to compose an elephant*)! Again, the problem for Schaffer is the same: how can he rule this alternative hypothesis? Of course, this illustration is just the opposite of the first one: in the first one we have what most regard as a nomic connection, and I asked how Schaffer might rule out the perverse hypothesis that it is really a grounding connection. In this second example, we have what most regard as a grounding hypothesis, and I am asking how Schaffer might rule out the perverse hypothesis that it is really merely a nomic connection.

at the first step. For this reason, the Russellian monist has no call to reject the conceivability-possibility link. In fact, he can accept a liberal conceivability-possibility link.

True, this means that the Russellian monist must accept the possibility of *structural zombies*: beings who satisfy the same structural description as us but who lack consciousness. For, since we grasp structural properties, we *can* conceive of that. So they reject the view of standard physicalists that the experiential facts are necessitated by the *accessible* physical facts. But this is not a problem for her unique brand of physicalism. For, since she holds that consciousness is quiddity-involving, this would have to be a world where the quiddities are different, and, unlike in the actual world, are not “the right kind” to make for consciousness. They are not “*phenomenal*” quiddities.⁶ As Hawthorne (2006, 222) puts the idea, “quiddity switching turns the lights off and on”. Indeed, since quiddities are physical, this world would be physically different from our world, and hence would not be a counterexample to the kind of modal thesis that often goes with physicalism.

Here is another, related payoff of Russellian monism. It not only blocks the conceivability argument against physicalism; it also explains why we were tempted by it in the first place. Since we only grasp the *structural physical facts* and but we do not know the quidditistic facts, and since consciousness is quiddity-involving, we *think* we can conceive of the physical facts being the same but with the glow of consciousness absent. We find that there is a big “explanatory gap”. (Analogy: if you were only given a purely spatial, structural description of the pixel world in non-chromatic terms, you would find it easily conceivable that that world lack a “glowing” *yellow* pixel face, and you would find it mysterious how that world should produce a *yellow* pixel face.) But this is just based on our ignorance. We are missing out on the microphysical quiddities. The Russellian physicalist holds that, if only a super-intelligence knew the quidditistic physical facts in addition to the structural physical facts, then she would find it *inconceivable* that the high-level consciousness-facts should be any different. This means that she would be able to “see” *a priori* why, given *all* the physical facts, there *must* exist in brains macro-level states of consciousness, like seeing red or feeling pain. (Analogy: if you were told the spatial arrangement of the quiddity *yellow pixel* in world 2, you would “see” why in certain portions of that world *yellow pixel faces* appear.) Another way to put it: on Russellian monism, there is no explanatory gap for God. It is possible in principle to go *a priori* from the physical truths – the complex arrangements of a handful of phenomenal quiddities – to *all the truths about the conscious experiences of all sentient creatures (from humans to bats to aliens)*.

Now you can appreciate Chalmers’ Hegelian synthesis argument for Russellian monism. Since it is a form of physicalism, it appears to avoid the complexity problem associated with dualism. And at the same time it answers the conceivability argument in an attractive way.

However, at this point you might be thinking Russellian monism is too good to be true – that it is just incredible. The worry is based on the *combination problem*. Here is my favorite way of putting it. As I mentioned, there is only a *handful* of

⁶ Another possibility for the Russellian monist is that at least some Zombie worlds are purely “structural” worlds where there are no quiddities. On one version of this view, in such world there is a giant existentially quantified structural truth but it is not grounded in anything more basic – it is true but has no witnesses.

types of micro-phenomenal quiddities; they are identical with (or realize) the handful of fundamental physical properties. There are certainly *more* types of macro-level experiences across the universe: we have lots of types of experiences of qualities (colors, pains, tastes, etc), and other animals (think: bats) have still other types of experiences (not to mention aliens if such there be). This numerical difference in variety guarantees that, for at least *some* types of macro phenomenal property, *there is no phenomenal property of that type at the micro-level*. For instance, let us suppose that bats experience qualities that are alien relative to the qualities we experience. And let us suppose for the sake of illustration that this is one of the macro qualities that doesn't appear at the micro-level. Then *nothing like* this quality appears at the micro-level – the micro level and macro level qualities are totally different. Now, in general, there are not *a priori* connections between *totally different* qualities. So if the micro and macro qualities are totally different, then how might there be an *a priori* connection between them? How then could a super-intelligence, if she knew what the micro qualities are like, *a priori* deduce that bats experience this *totally different* macro quality? In general, whatever the handful of micro-phenomenal quiddities are, it is hard to believe that, if a super-intelligence knew what they were like and how they were combined in sentient creatures' brains, then she could *a priori* deduce from this limited basis the great variety of experiences that those creatures can undergo. On the contrary, there would *still* be a stubborn “explanatory gap”. If this is right, then in the end Russellian monism cannot after all close the explanatory gap and solve the conceivability problem. This is a big difference with the pixel case. For in this pixel case you have the *same* property at micro and macro level, namely *being yellow*. And we can easily see how yellow micro-pixels in the right combination must yield macro-level yellowness.

I think Russellian monism faces another type of problem. First imagine a person who has a huge diversity of conscious experiences – experiences of colors, pains, and so on. Now imagine a second person who is totally unconscious – say some one asleep or some one who is undergoing lots of sensory processing and sophisticated behavior but not at high enough levels for conscious experience (a kind of temporary zombie). According to Russellian monists, if a super-intelligence knew only the *structural facts* about them, she could *not* deduce *a priori* that there is this hugely significant macro-level difference between them – namely, the huge difference between enjoying technicolor consciousness and “darkness within”. Still, according to them, if only she knew secret the micro-phenomenal quiddities instantiated by the micro-parts of their heads, then this would make all the difference: all of the sudden she could deduce that there is this huge macro-level difference, and in fact could deduce exactly what experiences the first person has. But – and this is the problem – exactly *the same* micro-phenomenal quiddities are instantiated in their heads, only in different “combinations” or “structures”. For the micro-phenomenal quiddities are just identical with (or realizers of) mass and charge and so on, and mass and charge and so on are uniform in nature – everything is made from these same fundamental ingredients. This is puzzling. How could it be that, if the super-intelligence merely learns the identities of micro-phenomenal quiddities, which are exactly *the same* between the two subjects, then she can suddenly deduce that there is this hugely significant macro-level *difference*? Let us call this the *big difference problem*, because it is the problem of how the same micro phenomenal quiddities, when “structured” in slightly different ways in different systems, could intelligibly determine *big breaks* in nature at the macro level:

those that hold between all conscious creatures and physically similar but non-conscious creatures (arguably one of the most important break in reality).

I am impressed by combination problem and the big difference problem. Maybe you are impressed by them too. Do they mean that Russellian monism can be ruled out of court? Maybe not. The combination problem, at least, has already been much discussed and many Russellian monists think that it is *not* decisive. Another point is that the mind-body problem has proved to be nature's hardest nut to crack. So I think we should look seriously at all interesting ideas on how to solve it, even if at first blush they seem a bit implausible. In fact, in this paper I will just set aside the combination problem. I will also set aside the "big difference problem". I am going to argue that, even if we set these problems aside, Russellian monism fails on account of other problems.

In sum, as I understand it here, Russellian monism is the conjunction of the following theses:

- Physicalism
- Quidditism
- Quidditistic Ignorance
- Consciousness as Quiddity-Involving

A final point: Russellian monists reject the view of standard physicalists that the *accessible* physical facts - about brain states and functional organization and so on - necessitate the experiential facts. They reject this view on the basis of *conceivability reasoning*. They hold that "structural zombies" are easily conceivable: it is easily conceivable that these "structural" facts should be the same, but with consciousness absent. And they hold that this provide strong enough evidence that this is possible that we should accept that it is possible. (This is in contrast to standard physicalists who hold that we have special reasons to think that in this case conceivability is not strong evidence of possibility.) So they are under pressure to take seriously conceivability reasoning about experiential matters *in general*. This point will play an important role in the development of some of my problems. I will put this by saying that Russellian monists accept a "liberal conceivability-possibility link". This is deliberately vague. For instance, Chalmers formulates a general link between conceivability and possibility using the ideas of "two-dimensionalism". But other friends of Russellian monists could favor a different way of formulating the link between conceivability and possibility.⁷ In fact, they might not have any general principle at all, taking a more "particularist" attitude towards conceivability reasoning.

1.2 Reductive and Primitivist Russellian Monism

Next I want to introduce the distinction between *reductive* and *primitivist* Russellian monism that will form the basis of my dilemma. It corresponds to the distinction between reductive and primitivist versions of standard physicalism. I think it is very important, but Russellian monists haven't discussed it much.

Let me start with the reductive vision of our world. Let me say right away that, following Sider and others, I will be understanding "reduction" broadly. Consider,

⁷ See, for instance, Yablo (2000, 121).

for instance, a functionalist theory of what it is to be a *hand* (Sider 2011). As I use “reduction”, this might count as a reduction. The general approach of physicalism has prominent adherents, including Armstrong, Lewis, Papineau, Sider, Tye, among many others. Reductive Russellian monism adds that conscious properties reduce to *quiddity-involving* complex properties. Chalmers says that Russellian monism achieves the virtues of simplicity and the other virtues of physicalism, but without physicalism’s biggest problems; we will see that it is a broadly reductive form of Russellian monism that would come closest to doing so.

Reductive physicalism can be explained in terms of the idea of an *identification*, which has played an important role in philosophy and science. An identification is a claim of the form *to be F is to be G* (Dorr MS). For instance, to be a vixen is to be a female fox, and to be a case of water is to be a case of H₂O. *I take it that these are a species of identity-statements, employing the general notion of identity; only here the identity sign is flanked by predicates and not names.* An identification yields a necessary bi-conditional claim: necessarily, all and only F things are G things. But an identification is not a mere necessary bi-conditional. Another name for this is a “*real definition*”. As these examples show, some identifications are *a priori* while others are *a posteriori*. Roughly, a *reduction* is an identification where the left-hand predicate is simple and the right-hand predicate is complex.

Such “identifications” or “real definitions” could be understood in different ways. One might instead appeal Sider’s notion of “metaphysical analysis” (2011). I myself am a realist about properties and I will suppose that identifications correspond to property-identities. If *to be F is to be G*, then the property of being F is identical with the property of being G. I will also assume that there are some (Russell-paradox-avoiding!) principles of property-formation which say that, given so and so properties, there are so and so complex properties (disjunctive, conjunctive, structural, functional, etc.), with such and such instantiation-conditions. Then we can say that a property *P reduces to* properties *Q, R . . .* iff P is identical with a complex property built from *Q, R, . . .* However, I would like to stress that nothing hangs on my decision to explain the reductive picture in terms of an ontology of complex properties. As I said, the central idea is that of an “identification”; and this can be understood without an ontology of complex properties – in fact it can be understood on a nominalist view that does away with properties altogether (Dorr MS).⁸

The reductive picture is very natural for the pixel world. We can imagine that the pixel world started out with certain initial conditions - a few yellow and black pixels. Afterwards, a few simple laws governed the evolution of these *base properties*. The reductionist holds all other properties that show up in this world are just *complex properties* built from the base properties. So, for instance, after awhile, things appear with the property of being a yellow pixel-face. The expression “being a yellow pixel-face” might refer, relative to one acceptable precisification, to a disjunction of all the pixel-arrangements that look a bit like the one shown in the picture above. This is a disjunction of structural properties. (Since there is vagueness about what pixel-arrangements are faces, there will be multiple acceptable preci-

⁸ What about “redness is Johnston’s actual favorite color”?..... By the way, in what follows, “identification” or “real definitions” *should not* be understood along the lines suggested by Gideon Rosen (2015). Elsewhere (Pautz MS) I argue that, on his account of real definitions, real definitions of experiential conditions in physical terms are compatible with weird forms of emergence. And I suggest that his account of real definition faces other problems.

sifications.) On another, perhaps more acceptable precisification, it might refer to the “functional” property having some arrangement of yellow pixels that has the second-order property of entailing a certain kind of shape.

Reductive Russellian monism provides a similar picture of our world. It comes in different versions, depending on different choices of a “base”. As I understand “the base”, it is just a *list* of individuals and properties. Different forms of reductive Russellian monism result from different choices of the “base”. Each must be evaluated on its own merits. I will assume that for the Russellian monist the list *at least* includes certain structural and topic-neutral properties and relations, as well as all the microphysical quiddities. An austere version would try to stick to these. A more liberal version would add more elements to the “base”; for instance, if you are a necessitist about properties, holding that necessarily all properties exist necessarily (somewhat like mathematical objects), then your base might be very plenitudinous, including all possible fundamental properties as well as the actually-instantiated fundamental physical and topic-neutral properties. (This would help with the issues of multiple realizability to be discussed below.) Reductive Russellian monism is then the thesis that all things are sums of the things in the relevant base and all properties instantiated in our world are complex properties built from the properties in the relevant base.⁹

For instance, simplifying grossly, the reductive Russellian monist might say that the property-designator “having THIS experience of red” refers, relative to a precisification, to some *quiddity-involving* complex property of the form: *having micro-parts with quiddities X, Y and Z arranged to form brain state B*.¹⁰ And “being a hand” refers, relative to a precisification, to another exceedingly complex property – only this time it is a quiddity-neutral one.¹¹

Many people would immediately reject reductionism on the grounds that many of the properties instantiated at our world are *multiple realizable*: they could have been instantiated in worlds where the correct physics is totally alien from our world. Schaffer (2013) has recently pressed this objection against Sider’s reductionism. But, since there are nomic properties and relations in the base, *reductionism in my sense it very broad: it allows for “functionalist”, “topic-neutral” reductions*. And this may help accommodate “multiple realizability”. There are other ways in which the reductionist might respond to multiple realizability.¹² It is also worth mentioning that reductionism doesn’t require that we can actually specify the relevant ex-

⁹ Because of quantum non-locality and etc. we don’t want to commit physicalism to *micro-physicalism*.

Will a “causal relation” be in the base? Most certainly not! Causal talk is context sensitive in ways that make it implausible that there is a single fundamental relation picked out by that talk. But maybe there will be a fundamental law-making relation (Armstrong-Tooley) or a fundamental law-making property *being a law* that attaches to regularities (a way of turning Maudlin’s basic-operator view into a basic-property view).

¹⁰ But is it a problem for this identification that every experience can go with every structure, given Russellian monism’s commitment to conceivability-possibility link? The quiddities of our world, when arranged in a brain, make for consciousness; but not when arranged into a rock. For the quiddities of other worlds, it is the reverse!

¹¹ Armstrong (1989, p. 101).

¹² Points that may help with multiple realizability: (1) Maybe alien properties are in the base. (2) Appealing to resemblance may help (Cian Dorr’s suggestion?). E. g. on one precisification “is a mountain” might refer to the property *having a property-profile that relevantly resembles the property-profile of THIS* (demonstrating a mountain). (3) skepticism or deflationism about the “outer sphere” and the possibility of “alien properties” (Armstrong).

ceedingly complex properties. As Ted Sider (2011) says, we often can at best provide “toy” reductions. Compare the pixel world. *It is obvious that the property of being a pixel face reduces even if it is multiply realizable and we certainly cannot specify the relevant reduction(s).*

Who advocates reductive Russellian monism? Grover Maxwell (1978), a pioneer of Russellian monism, can be classified as a reductive Russellian monist. He advocates a Russellian monist version of the mind-brain identity theory. *Also Herbert Feigl.* David Chalmers would at least be open to reductive Russellian monism. He says (2012, 381) that that ordinary terms like *friend* might have an exceedingly complex bi-conditional definition in very austere terms (some very basic topic-neutral causal-spatial terms and some basic phenomenal terms). So if he accepted Russellian monism, he might be open to there being similar bi-conditional definitions of phenomenal terms in terms of structure plus quiddities. And from *a priori* bi-conditionals it is a short step to out-and-out identifications.¹³

Why might Russellian monists opt for reductive Russellian monism? It is evidently a very *simple view of the world*. In this regard, it achieves the virtue of economy associated with physicalism. Further, I suggest that there is a kind of “*continuity*” *argument for reductionism*. Certainly at the very start of the universe, when things were relatively simple, reductionism reigned: all properties instantiated at the macro-level (e. g. being butane, etc.) were just complex properties built from the fundamental properties. It would be odd if at some point later on properties appeared for which reductionism fails. What point would that be?

Next I turn to *primitivist* Russellian monism. Whereas the key idea of reductive physicalism is that of an “identification” (“real definition”), the key idea of primitivist Russellian monism is that of “grounding” (Fine, Schaffer, Rosen, etc.). In brief, it combines the denial of reductionism for some experience properties with the view that their instantiation is grounded by the instantiation of the fundamental quiddities and structural properties. We will see later that, despite the appeal of reductive Russellian monism, there are special reasons for thinking that Russellian monists must accept primitivist Russellian monism.

To get a fix on the basic primitivist picture, let us start by considering the pixel world. It is plausible to think that, at least in the *early* stages of the pixel world, when the pixels were distributed in totally boring ways, reductionism reigned: all properties instantiated in this world were just either the base properties or complex properties built from them. So no special principles were required; just general principles about how to form complex predicates from simple predicates. But, after awhile, things become more interesting. For instance, after awhile, maybe some complex arrangements of pixels were *objectively good*. Maybe they were objectively beautiful and beautiful things are intrinsically good. And someone like Schaffer (2013) might argue on the grounds of multiple realizability that the property of being objectively good cannot be identified with a complex property built up from the “base properties”. If that is what you think, then what should you say about this

¹³ Chalmers (2014) holds that phenomenal terms are transparent. You might think this is inconsistent with the reductive Russellian monist idea that they have definitions in fundamental terms are in principle *a priori* knowable but that certainly not known by us now (Goff?). This is mistaken, though. By “transparency”, all Chalmers means is super-rigidity (an expression is super rigid iff it is epistemically rigid and metaphysically rigid *de jure*). And the super-rigidity of a term is compatible with its having a non-obvious *a priori* definition. For instance, the super-rigidity of “limit” (in math) or “knows” or “friend” is compatible with their having highly non-obvious *a priori* definitions.

property? Following G. E. Moore, you might say that it is a *simple or primitive* property that has no real definition in other terms. But then you might add this simple property is special: unlike other simple properties (masses, charges, etc.), it “needs” other properties in order to be instantiated. In particular, in the pixel world, its instantiation is always *grounded in* the instantiation of some such pixel-involving complex property. This view has no problem with multiple realizability, because in other possible worlds, its instantiation might be grounded in totally different complex properties (ones not involving pixel properties but alien properties). So on this view there is a slew of *special grounding laws* that are specific to the property of being good, linking it with specific distributions of pixels. These special grounding laws are brute: they can’t be derived from any more basic truths.¹⁴

Likewise, the *primitivist Russellian monist* holds (or should hold!) that our universe started out just as the reductive Russellian monist claims. In the early stages of the universe, after the big bang, everything was pretty boring. All properties instantiated at our world were just either properties in the base, or complex properties built from those properties. For instance, the property of being butane appeared, but this is just a complex property. So to explain its origin we need no special principles; we just need totally general quasi-logical principles about the formation of complex predicates from simple predicates. But, according to the primitivist Russellian monist, after awhile things became more interesting. When brains evolved, properties started to “pop up” that *aren’t* just identical with complex properties built up from the properties in the base. These are conscious properties like *being aware of red* and *being in pain*. These properties are “primitive” or “simple” (or, if they are complex properties, they are built up from simple mental ingredients, like a relation of *awareness* and *the color red*). There are no true, interesting “identifications” involving these properties. Yet there are special brute conditional *grounding laws* linking the distribution of micro-quiddities with the instantiation of these distinctive macro-properties. So the view is still physicalist, in name if not in spirit.

Some philosophers seem to accept primitivist physicalism in some form of other. Jonathan Schaffer (2013) might be an example. He accepts physicalism but rejects Ted Sider’s brand of reductionism about the manifest image on the basis of multiple realizability; and his objection would seem to carry over to any brand of reductionism. John Hawthorne (2006, p. 206) and Gideon Rosen (2010: §13) have flirted with such a view. The view is not immediately self-contradictory, since identifications and grounding claims are different sorts of claims. Primitivist Russellian monism combines the general “primitivist-physicalist” picture with the tenets of Russellian monism.

Since Russellian monism is a version of *a priori* physicalism, primitivist version of the view would require that the pattern of instantiation of a vast array of *simple, irreducible* phenomenal properties at the macro level (experiences of red, experience of pain, etc.) is in principle deducible *a priori* from pattern of instantiation of a handful of phenomenal quiddities at the micro level. For instance, God could just “see” the entailment. This may seem strange. But there may be other cases of *a priori* grounding connections without reducibility. For instance, *being scarlet* grounds *being red*, but many say that *being red* is not reducible to a disjunction with *being*

¹⁴ Fine thinks that they can be explained by “essentialist truths” but this doesn’t make a difference to what follows. For more on this see pp. 36-7.

scarlet as a disjunct (Rosen 2010). Or again, many think that facts about what ought to be the case are irreducible to facts about what is the case, but nevertheless hold that the facts about what ought to be the case are grounded in (and indeed *a priori* grounded in) facts about what is the case. Since there may be cases of *a priori* grounding without reduction, we cannot immediately rule out primitivist Russellian monism from the start. It is a view we should look at seriously.

Still, primitivist Russellian monism is an unattractively complicated view. I am calling it a physicalist view, but it seems to have many of the vices of dualism. Where dualism requires psychophysical laws, this view requires exactly parallel special, brute “grounding laws” (Rosen’s term) that do not follow from identifications or real definitions or from more standard grounding claims (disjuncts ground disjunctions, determinates ground determinables, and so on). Such grounding laws add to the complexity of the theory in much the same way as psychophysical laws add to the complexity of dualism. Reductive Russellian monism is initially much more attractive. In place of brute grounding laws, it posits identifications. Such identifications are *radically different from the kinds of “grounding laws” posited by the primitivist. They are special in that don’t add to the complexity of the theory – in fact, they reduce complexity.* (Compare: if you stipulate that a bachelor is an unmarried man, then having this identification in your theory doesn’t add to its complexity.)¹⁵ *This is why bi-conditional identifications are preferable to one-directional grounding claims. So Reductive Russellian monism would most clearly achieve the virtue of economy traditionally associated with physicalism.*

2. Problems with Reductive Russellian Monism

Now I turn to the main business of this paper: developing a dilemma for Russellian monism, arguing that neither reductive nor primitivist Russellian monism is adequate. In the present section, I look at the “reductive” horn. In the next, I turn to the “primitivist” horn.

I just said that *reductive* Russellian monism would most clearly achieve the virtue of economy traditionally associated with physicalism. But, as everyone knows, reductionism in general faces problems, even if it is a very attractive view and has very resourceful proponents (Sider, Lewis, Armstrong). I will argue that reductive Russellian monists face even *more intractable* problems. Far from combining the virtues of dualism and standard physicalism (achieving Chalmers’ “Hegelian synthesis”), it is just not an option for Russellian monists.

Why is reductive Russellian monism even worse off than reductionism more generally? Briefly, there are two reasons. (1) *Quiddity-involvement*: reductive Russellian monists identify conscious properties with complex properties that involve *very low-level, fine-grained properties*, namely the micro-phenomenal quiddities. This leads to a problem that I call the *problem of phenomenal reference* (§2.1). (2) *Conceivability*: the whole motivation for the Russellian monism relies on an inference from conceivability to possibility. After all, it is on the basis of this link that they reject standard physicalism. But it’s a double-edged sword: such reasoning can be used to show reductive Russellian monism is not an option for them (§§2.2-2.3).

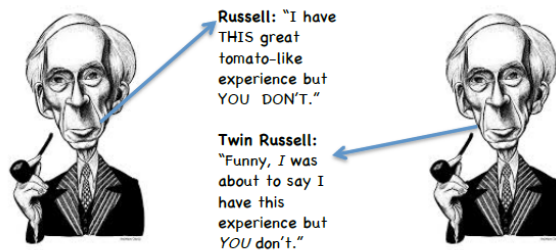
¹⁵ Many *also* say that identifications, unlike grounding claims, don’t “cry out for explanation” (Sider, Block, Stalnaker, Rayo, Dorr). But this is different from my claim that identities conduce to simplicity; and it is controversial (Chalmers).

2.1 The Problem of Phenomenal Reference

My *problem for phenomenal reference* for reductive Russellian monism is just an instance of a general problem for reductive theories, one that has been pressed by Ted Sider (2001) and David Papineau (2003). If you think that some manifest image property *M* is identical with some specific “scientific” property *S*, then you need to at least sketch an account how our term “*M*” latches onto the specific scientific property *S*, rather than the other candidates. If you can’t even begin to sketch such an account, then you should give up the identification $M=S$.

Reductive Russellian monists face an especially acute version of this problem. The reason is that they identify conscious properties with complex properties that involve *very low-level, fine-grained properties*, namely the micro-phenomenal quiddities, as I explained above. In particular, they hold that experiences are enormously complex *quiddity-involving* properties of our brains. In a nutshell, the problem of phenomenal reference for Russellian monists is this. If their view is right, then, whenever ordinary people talk about their experiences, they are (unknown to them) referring such complex quiddity-involving properties of their brains, properties they *don’t* share with certain “structural duplicates” of them. But people also have many quiddity-neutral properties that they *do* share with their structural duplicates, and in all respect they are equally good candidates to be what they are talking about. (Indeed, according to standard physicalists, such properties *are* what they are talking about.) I will argue that reductive Russellian monists cannot provide a plausible account of how our phenomenal terms determinately latch onto *quiddity-involving* properties, *rather than* quiddity-neutral properties.

This, as I said, is the gist of the problem. To develop the problem in more detail, I will use a fanciful example involving Russell himself. Here is the example:



Russell

Structural Russell

Here is what is going on in this example. For the sake of argument, let us just suppose that one element of Russellian monism is true: Russell’s micro-parts have *phenomenal* quiddities. And let us pretend that that Russell has a twin, only the twin’s micro-parts are *non-phenomenal* quiddities. And let us suppose that Russell and Twin Russell are both viewing a tomato. This quidditistic difference between

Russell and Twin Russell is totally inaccessible: even if we used electron microscope on twin Russell, we couldn't tell that there is any difference between him and Russell himself. However, let us suppose that Russell is somehow "tipped off" (by God if you like) that Twin Russell's micro-parts have non-phenomenal quiddities. Since Russell is (naturally!) a Russellian monist, he *thinks* (whether right or wrong) that this means that his twin is a *zombie*, even though (to be repeat) the twin is just like him in every accessible way. So he says "I have THIS tomato-like experience but you [Twin Russell] do not", referring demonstratively to the type of experience he is then giving. If reductive Russellian monism is true, then his speech is indeed *true*, because by "THIS experience" he is referring to a quiddity-involving property that *he possesses* but that *Twin Russell does not possess*.

In other words, even if Russellian monism is a metaphysical thesis, it takes on a semantic commitment:

Semantic commitment: It's determinately true that, when Russell uses "THIS experience", he is referring to some or other quiddity-involving property that he possesses but that twin Russell doesn't possess.¹⁶ (The same goes whenever anyone uses phenomenal terms.) So his speech "I have this experience but you don't" is *true*.

Now, when explaining why this is problematic, I will assume for the sake of discussion that the Russellian monist says that the property of having the tomato-like experience to which Russell is referring is identical with some or other extremely complex *intrinsic property of the brain*. It might have something like the following form: *having some of the "right" micro-phenomenal quiddities X, Y and Z arranged in some "right" way – as it might be, arranged so as to form brain state B*. Let us call this type of state *B+*. The plus sign indicates that this is a quiddity-involving property. Very roughly, it the property of *having a certain kind of brain state and being such that this brain state is realized by certain types of quiddities*. Grover Maxwell (1979) suggested a version of reductive Russellian monism along these lines.

The gist of problem of phenomenal reference will be this. Reductive Russellian monism is a reductive view of everything. So if reductive Russellian monism is true, then there must also be a *reductive account of how it is that Russell's use of "THIS experience" refers to quiddity-involving property B+*, a property possessed by Russell but not twin Russell. But, I will argue, there is no such account.

My argument for this is based on the fact there is actually a *huge multiplicity* of alternative physical-functional properties that are candidates to be the referent of Russell's phenomenal demonstrative "THIS tomato-like experience". They correspond to all the different major theories of consciousness. For instance, Russell has various *quiddity-neutral, wide* physical states of the form having some state or other that has the biological function of tracking the red reflectance-type (Tye, Dretske). He also has various neural states (Block) and various narrow functional states (Papineau, Prinz).

However, to develop the problem of phenomenal reference, it might be easiest if we grossly simplify the situation by restricting our attention to "narrow" physical-functional properties. In fact, let us suppose that there are only *two* narrow physi-

¹⁶ Of course she can say it is indeterminate which one.

cal-functional properties that are candidates to be the property that Russell is referring to. The first is the *quiddity-involving* property B+ that I already described. Roughly, this is the property of *having brain state B where this is realized by certain phenomenal quiddities at the micro level*.¹⁷ As for the second property, let us suppose that it is just like B+, *except it is quiddity-neutral*. That is, it is just a matter of *having brain state B, however it is realized at the inaccessible, lowest level of micro-quiddities*. Let us call this property simply B, with no plus sign. Roughly, B is just B+ *stripped of any quidditistic-requirements*. To appreciate the difference between B+ and B, consider Twin Russell. Recall that Twin Russell is just like Russell in every accessible way. Even if you use an electron microscope on his brainy parts, you couldn't tell a difference. So, for all intents and purposes, they are identical in their neural states. Because it packs in hidden quidditistic requirement, B+ is a property that Russell has *but Twin Russell lacks*. By contrast, B is neural a property that they both share. A standard type-type identity theorist like Block, McLaughlin or Papineau would identify Russell's experience with B, rather than with B+. It is a commonplace that the brain can be described at different "levels of abstraction", where properties at one level of abstraction are "multiply realizable" by properties at lower levels of abstraction. B corresponds to the level of description of neuronal patterns. By contrast, B+ corresponds to a richer, more fine-grained level of description that also includes hidden details of quidditistic realization.

Given this simplifying assumption, the question for Russellian monists is: what is the reductive account of phenomenal representation which implies that Russell's use of "THIS experience refers to B+, rather than referring to B (or being indeterminate in reference between B+ and B)?"

According to many meta-semantic theories, causation plays a big role in fixing the reference of many referring terms. But this will not help. For *both* B+ and B are "causally efficacious" in causing Russell to produce the demonstrative "THIS experience", if either is. They are after all both very similar, co-instantiated brain states. Indeed, they are nomically co-extensive, except in far out cases like the case of Twin Russell that we never encounter in real-life. They just correspond to different levels of abstraction. On pain of general macro-level epiphenomenalism, we cannot suppose that properties at different "levels" always compete for causal efficacy. True, B+ includes some quidditistic requirements that aren't included in B. But to suppose that only quidditistic-involving properties are causally efficacious would be to suppose that nearly all the properties we talk about are not causally efficacious. For nearly all the properties we talk about are not quiddity-involving. Think, for instance, of the properties of the special sciences. They are not quiddity-involving. Surely, these properties are instantiated in structural duplicate worlds which are just like our world in every accessible respect but which differ only in the inaccessible quiddities.^{18, 19}

¹⁷ Another idea is that B+ is some property characterized mainly in terms of structural stuff and just quantifying over microphenomenal properties. (Analogy: "having a bunch of regions whose colors are very different from the colors of their neighboring regions.") *But this doesn't help with the problem of phenomenal reference: even on this view the problem of phenomenal reference applies, just as I will develop it.*

¹⁸ You actually might think that B *rather than* B+ is "the" (a?) cause because of proportionality considerations: it packs in too much unnecessary quidditistic detail (Yablo 1992). But some think proportionality stuff about explanation, not causation (McLaughlin).

Now here finally is a statement of the problem:

The problem of phenomenal reference: Russell not only has the quiddity involving brain state B+, a property he fails to share with Twin Russell. He also has a multitude of distinct physical-functional properties, which he shares with Twin Russell. For instance, he has brain state B, which is just like B+ except that it is quiddity-neutral. In all respects, B+ and B are equally suitable candidates to be what Russell is referring to with “THIS experience”. They are equally “natural”. And they are both causally efficacious in the production of Russell’s employment of the demonstrative “THIS experience”. Given the facts, there is *no reductive theory of phenomenal reference* compatible with the Russell’s monist’s semantic commitment: that with “THIS experience” Russell refers to quiddity-involving B+, rather than to a quiddity-neutral property such as B. So reductive Russellian monism fails for meta-semantic reasons.

To drive the problem home, let me use some analogies. Imagine a crazy (non-actual!) philosopher who is a reductive Russellian monist about *rocks*. Her favorite theory that being a rock is a highly complex *quiddity-involving* property. In other words, according to her theory of rocks, to be a rock you need the “right” micro-quiddities. So there are structural duplicates of rocks that are just like rocks in every accessible detail (even using electron microscopes and so on) but that are *not* really rocks (they are “fool’s rocks”) just because their micro-parts don’t have the “right” inaccessible quiddities. This is a bad theory of the property of being a rock. One reason is that there is no plausible account of how our word “rock” could determinately latch onto such a quiddity-involving property, rather than a more structural property that rocks share with all structural duplicates of rocks that are just like rocks in every accessible respect (this, I take it, is the right view in this case). Here is another analogy. Timothy Williamson suggests that *being bald* is identical with some specific hair-condition, for instance *having exactly 1,171 hairs or less*. One reason why this cannot be correct (in the opinion of many critics of Williamson) is there is no good account of how “bald” might determinately latch on to such a fine-grained property as opposed to all the other candidates (Weatherson 2003). I think that reductive Russellian monism is implausible for the same kind of reasons that these other theories are implausible. Russell’s experience *can’t be* a quiddity-involving complex property B+, as reductive Russellian monists think. The reason is that, if this theory is correct, then it is what Russell is referring to when he uses “THIS experience”; but the reductive Russellian monist cannot even sketch a plau-

Denis Robinson says (argues?) that quiddities are epiphenomenal: “The upshot of the generalized version of the Epiphenomenalism Argument appears to be that the intrinsic essence of a fundamental property, that aspect which determines the intrinsic similarity of particulars independently of functional similarity, is epiphenomenal.” Robert Howell paper?

Jackson/Braddon Mitchell simply assert that quiddities are unknowable because acausal.

Hawthorne and recent Phil Studies paper argue that quiddity switching would be undetectable.

¹⁹ An interesting question is whether *Twin Russell’s* use of “THIS experience” refers to a *quiddity neutral* property. If it does, then by parity Russell’s does too – contrary to Russellian monism.

sible account of how Russell might be referring to this specific property to the exclusion of other candidates.

Let now address some solutions on behalf of reductive Russellian monists. I will suggest that they fall short.

(I) The Russellian monist might reply that she can give *reasons* for thinking that Russell, with “THIS experience”, is referring to a quiddity-involving complex property like B+. They are just the reasons for accepting the Russellian monist view that Russell’s experience is a quiddity-involving complex property like B+. Doesn’t this solve her problem of phenomenal reference?

Of course, this *doesn’t* solve the problem at all. The problem is not to supply *reasons* for thinking that Russell is referring to a quiddity-involving complex property, but to sketch a reductive theory of phenomenal reference which *explains* how this is possible. If no such theory can be sketched, then we have a good reason to reject reductive Russellian monism. Analogy: even if you can give reasons to thinking that the number two is an arcane abstract object, you still own an account of how it is possible that “two” refers to this arcane object. If you cannot sketch such an account, this gives us reason to reject this view of the number two (Benacerraf 1973, Hodes 1984).

(II) The reductive Russellian monist might point out that sometimes descriptive fit, as well as causal connections, help determine reference. And this is so even if Kripke is correct that pure descriptive theories fail. Thus, suppose I say “THIS is beautiful” while pointing to a statue. I am referring to the statue, and not the coincident hunk of clay, even though both are causally linked to my utterance, because the statue fits certain internally-represented descriptive information (for instance, that the thing cannot survive radical changes in form). So couldn’t descriptive fit help with the Russellian monist’s puzzle of how Russell refers to B+ rather than B, even though they are necessarily co-extensive and both causally efficacious in producing Russell’s utterance?

It could not. To begin with, recall that the problem is general: it doesn’t just concern Russell, but ordinary mortals too, including unsophisticated mortals. It just false that any normal must implicitly associate phenomenal terms with sophisticated descriptive information that is uniquely satisfied by quiddity-involving properties like B+ and not by quiddity-neutral properties like B.²⁰ Further, it is just intuitively implausible that descriptive fit plays *any* serious role in fixing the reference of phenomenal demonstratives like “THIS experience”. It seems to be a more primitive form of reference.

(III) This point – about how introspective reference seems to be an especially direct or primitive form of reference – leads to the last attempted solution I will consider. The reductive Russellian monists might simply respond a relation of *introspective attention*, or perhaps *acquaintance*, is the basis of phenomenal reference. (This response was suggested to me by David Chalmers.) Further, she might say that, when Russell uses “THIS experience”, *he is acquainted with the instantiation of the quiddity-involving property B+, and not the quiddity-neutral property B*. After all, she thinks B+, but not B, is an *experiential* property. That, she might say, is how he refers o B+, and not B. Call this the *acquaintance-based account*.

²⁰ One idea is that the relevant descriptive information, which is satisfied by B+ but not B, is something like *it is not necessitated by structural facts alone*. But a child or ordinary person doesn’t have anything sophisticated like this in mind.

But the acquaintance-based account is not yet a solution to the problem of phenomenal reference.²¹ Remember that the problem was that, if *reductive* Russellian monism is true, then there must be a *reductive* theory of phenomenal reference, which explains how Russell refers to B+ rather than B. The acquaintance-based account is not yet reductive. *Dualists* might be able to appeal to an *irreducible* relation of acquaintance to explain phenomenal reference, but reductive Russellian monists cannot. In order to be reductive, the acquaintance-based theorist must give a reductive theory of this “acquaintance-relation” along the following lines:

$$\lambda x \lambda y (x \text{ is acquainted with } y) = \lambda x \lambda y (x \dots y)$$

where the dots are only filled with *physicalistically-acceptable* terms (including quiddity terms). Further, this reductive account of the acquaintance relation must be compatible with the claim that Russell bears this relation to *the instantiation quiddity-involving property B+*, and not the *quiddity-neutral property B*. That is to say, this physical relation must be exquisitely tuned to B+ and not B. Only then will the reductive Russellian monist have a *reductive* account of how Russell refers to B+ and not B.

But what on earth could this reductive account be? One natural idea for the reductive Russellian monist is to try reductively explain the acquaintance relation (at least in part) in terms of causation. Maybe to be acquainted with a state is to have a language of thought term for it that stands in some special causal relation to it. But, as I emphasize above, *B+ and B* are both causally efficacious in the production of Russell’s phenomenal demonstrative “THIS experience”. So a reductive account of acquaintance in terms of causation would not give the reductive Russellian monist what she needs: an account of how Russell refers to *B+ but not B*.

To conclude: we cannot even begin to sketch a reductive theory of phenomenal reference, which is compatible with the Russellian monists semantic commitment, namely, that Russell refers to the quiddity-involving B+, rather than the quiddity-neutral B. The problem of phenomenal reference provides a strong reason to think that reductive Russellian monism is unworkable.

Let me end with a couple of clarifications. Of course, all reductive physicalists – not just reductive Russellian monists – face similar general problems about reference-determination – not just for phenomenal terms like “this experience” but also for non-phenomenal terms like “water” and “rabbit”. I agree. But I think that sort of “reference problem” concerning phenomenal terms faced by reductive Russellian monists is *much worse*, for a couple of reasons. First, there are well-known accounts, or at least account-sketches, about how a non-phenomenal term like “Godel” or “water” might refer without extreme indeterminacy. By contrast, reductive Russellian monists cannot offer a half-way plausible reductive account of how Russell’s use of the phenomenal term “THIS experience” might determinately refer to a *quiddity-involving* property such as B+. Second, in response to problem of reference, *standard* reductive physicalists could just accept a great deal of indeterminacy, as Sider and others have suggested. Indeed, Papineau (2003), who accepts extreme phenomenal indeterminacy, would claim that Russell’s phenomenal demonstrative is *massively indeterminate in reference between quiddity-involving properties*

²¹ Besides the problem I will raise, many would say that the acquaintance-based idea goes against “transparency”. But see Chalmers against transparency (2013).

like B+ and quiddity-neural properties like B. (Compare Field (1973) on how “mass” in Newton’s mouth had *divided reference*.) But, even if such indeterminacy were acceptable, it is not a possible way out for reductive Russellian monists. Their view requires that it is determinate that phenomenal terms refer to *quiddity-involving* properties, not that they refer indeterminately to quiddity-involving properties and purely “structural” properties.²² I have argued that there is no reductive theory of phenomenal reference which explains how this might be so.

A final clarification. You might wonder why I have described the problem of phenomenal reference as a problem for *reductive* Russellian monism. Isn’t it an insuperable problem for all Russellian monists, including primitivist Russellian monists? I think not. Primitivist Russellian monists have more resources to solve the problem. I will explain why at the start of the next section §3, where I turn to primitivist Russellian monism.

2.2 The Problem of Phenomenal Representation for Reductive Russellian Monism

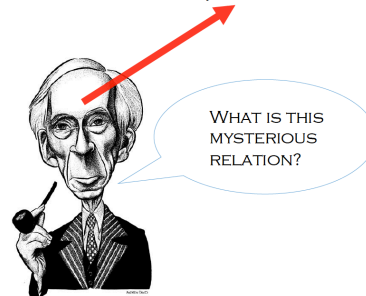
To illustrate my next problem, we can continue to focus on the example of Russell having a tomato-like experience. The problem of phenomenal reference for reductive Russellian monism was about how, in saying “THIS experience”, Russell might be referring more or less determinately to a quiddity-involving property that *he himself instantiates*. By contrast, next problem for reductive Russellian monism is about how, in having this experience, Russell is presented with, and can refer to, properties that are instantiated, if they are instantiated at all, *outside himself*.

The problem I have in mind is based on a very minimal form of representationalism, a generally accepted view (see for instance Dretske 1995, Tye 1995, Chalmers 2006b, Pautz 2010). I will just assume that some experiences are inherently representational. So Russellian monists need to be representationalists for some experiences.

To illustrate, let us now pretend that Russell’s tomato-like experience is hallucinatory. According to a minimal form of representationalism, necessarily, in having this tomato-like experience, Russell stands in what we might call the “the phenomenal representation relation” to the general properties being round and being reddish (see picture). So they seem to be instantiated before him. Yet in this case they are not instantiated before him – not even by a “sense datum” (contrary to his own view of the matter). This theory is needed to explain the obvious fact that Russell is having a non-veridical experience. It also needed to explain how, by having the hallucination, Russell can think about and refer to these properties, even though they are not instantiated before him. According to minimal representationalism, having visual experience is in general necessarily connected with “phenomenally representing” properties out there.

Elsewhere (Pautz 2010) I have argued that phenomenal representation creates a big problem for standard reductive physicalism. I will now suggest that it also

ROUNDNESS, REDDISHNESS



²² Of course, the reason is that they hold that in some structural duplicate worlds consciousness is determinately absent.

creates a decisive problem reductive Russellian monism – one in addition to the problem of phenomenal reference discussed in §2.1. The problem is simple. As Field in his seminal “Mental Representation” (1978) and in a recent (2001) post-script to that essay, reductive physicalists must not only give reductive accounts of monadic mental properties but also of dyadic mental relations. So, given minimal representationalism, the reductive Russellian monist must be able to sketch at least a “toy” identification of the following form:

$$[\text{PR}] \lambda x \lambda y (x \text{ phenomenally represents property } y) = \lambda x \lambda y (x \dots y)$$

where the dots are filled by physicalistically acceptable terms. But, on reductive Russellian monism, what on earth might this reduction look like? This is the problem of phenomenal representation.

One idea on how to complete [PR] comes from Tye (1995) and Dretske (1995). In explaining the problem of phenomenal reference for reductive Russellian monism in the previous section §2.1, I assumed for the sake of discussion a form of *internalism*: experiences are intrinsic properties of brains. Tye and Dretske suggest that to solve the problem of phenomenal reference we should rather accept a form of externalism. First, they locate the “sensible properties” in the external world along the shapes and spatial features. So, for instance, phenomenal colors are just reflectance properties. (The Russellian monist could accept a modified version of this view, identifying sensible colors with *quiddity-involving* reflectance properties – this would help with the “objective explanatory gap.”) Then they identify the phenomenal representational relation with a complex causal-functional relation: $\lambda x \lambda y (x$ is in a inner state that is apt to cause so-and-so cognitive-behavioral responses and that *would* be caused by the instantiation of y were conditions biologically normal). Call this the *tracking relation*. This is an *externalist-functionalist theory* of phenomenal representation. And it yields an externalist-functionalist theory of the character experience, quite different from the internalist view we were working with in §2.1. For if experience is essentially tied to representation, and if representation depends on extrinsic factors, then so does experience.

But reductive Russellian monists cannot accept this standard externalist-functionalist approach to phenomenal representation. The whole motivation for the Russellian monism relies on an inference from conceivability to possibility. After all, it is on the basis of this link that they reject standard physicalism. Because of their commitment to this very type of conceivability argument, externalist-functional theories of phenomenal representation are off the table for them. For we can conceive of Russell bearing that relation to a property without *phenomenally representing* that property.

In short, in the special case of phenomenal representation, Russellian monists cannot accept any of our standard causal-information models for reductively explaining representation.

Instead, Russellian monists must hold that, if there is a reductive account of the phenomenal representation relation at all, it must be quiddity-involving, not causal-structural like Dretske-Tye. It would have to be radically different from standard accounts. Such a reductive account would enable Russellian monists to block the conceivability argument. In particular, in response to this argument, Russellian monists can wield their standard “ignorance” response. Since we are ignorant of the relevant quiddities, we don’t even have the capacity to imagine a case in which

they are the same but in which Russell doesn't stand in the phenomenal representation relation to redness.

Now here is the crux of the problem of phenomenal representation as it arises for reductive Russellian monists. The problem is that it is not even remotely possible to see what the relevant quiddity-involving reduction might look like. The phenomenal representation relation is a *dyadic relation*. The quiddities are *typically* monadic (mass, charge, and so on). How might one construct a dyadic relation *Q* out of these generally monadic properties, a relation with which the phenomenal representation relation might be identified?²³ More generally, my point is that it is not possible to even gesture at a toy quiddity-involving reduction of the phenomenal representation in the form of [PR]. *And, as Sider (2011, 117), if you cannot even gesture at a toy reduction of a relation, that is strong evidence that it has no reduction.*²⁴

Let me make a couple of clarifications. First, the problem of phenomenal representation for reductive Russellian monism is totally independent of the combination problem, a general problem for all forms of Russellian monism that I described in §1.1. (Here I am responding to Chalmers who suggested in discussion that perhaps the problem of phenomenal reference is just a version of the combination problem.) The combination problem is an *epistemic problem*. Recall that Russellian monism is committed to *a priori* physicalism. So, for instance, it is committed to the epistemic claim that, if a super-intelligence knew all of the hidden micro-quiddities in Russell's head, and their combination, then she could *a priori* deduce that he is phenomenally representing reddishness out there. The problem is just that this is an epistemic claim implausible for various reasons. By contrast, the problem of phenomenal representation is a totally separate *metaphysical problem* specifically for reductive Russellian monism. To see this, let us suppose that the combination problem is based on a mistake. In fact, if a super-intelligence knew all of the hidden micro-quiddities in Russell's head, and their combination, then she *could* deduce *a priori* that he is phenomenally representing reddishness out there. This epistemic claim wouldn't yet solve the metaphysical problem of making it plausible if there is a two-place physical-functional relation with which the Russellian monist might identify the two-place phenomenal representation relation [PR]. That problem is still there with as much force. Again, the problem is that, if Russellian monism is true, there is just no plausible candidate. Russellian monists cannot identify it with a causal-tracking relation – that view is open to a conceivability argument. So even if the combination problem has a solution, the Russellian monists must say, contrary to reductive Russellian monism, that the relation of phenomenal representation is a *simple, irreducible* relation, even if it is *a priori* grounded in

²³ It is true that there are some fundamental physical relations – and so, on Russellian monism, fundamental relation quiddities. (In fact, on a Ladyman-style “structuralism”, it's all relations at the fundamental level.) But my main point still stands. It is not possible to even gesture at a toy quiddity-involving reduction of the phenomenal representation in the form of [PR]. *And, as Sider (2011, 117), if you cannot even gesture at a toy reduction of a relation, that is strong evidence that it has no reduction.*

What about: *x* has some *Q*-involving *B* state of so and so type which tracks property *y*? (Deflationary?) Open to all my empirical problems. Also indeterminacy in slow switching; and Harry-Inverted-Sally.

²⁴ Ney (2015) asserts that “representationalism” is inconsistent with Russellian monism in general. I am not sure about that. My point is that it sits ill with reductive Russellian monism only. I will suggest in the next section that it is consistent with primitivist Russellian monism.

combinations of quiddities in the brain. (Compare the Moorean view that goodness is simple, even if it is grounded in complex combinations of natural properties.) In other words, even if the combination problem is based on a mistake, Russellian monists must accept “primitivism” – they are pushed to second horn of my dilemma that I will examine in the next section §3.

Here is a second clarification. The problem of phenomenal representation goes beyond the general problem for reductive physicalists of giving a reductive theory of conscious experience (thanks here to Chalmers). The problem is more specific: it depends on the representationalist idea that experience involves a dyadic “phenomenal representation relation”. If we reject this idea, the particular problem of phenomenal representation goes away. For then Russellian monists only have to worry about monadic experience properties. And they might just identify these with complex, quiddity-involving properties of the brain like B+ (though this would face the problem of phenomenal reference discussed in the previous subsection, as well as the problem of multiple realizability in the next subsection). By contrast, if they accept the representationalist idea that experience involves a dyadic “phenomenal representation relation”, as I think they must, then they also must give a reductive theory of this *relation* in quiddity-involving terms, one that takes the form of [PR] above. But, to repeat, we cannot even begin to see how this might be done.

2.3 Multiple realizability (Or: The Simple-Subjects Problem)

The problem of phenomenal representation for reductive Russellian monism derived in part from the fact that Russellian monism is based on conceivability reasoning. In fact, this is at the foundation of their case for the view over standard physicalism. But it is a double-edged sword. It rules out for reductive Russellian monists reductive proposals that standard physicalists can accept – for instance functionalist-externalist theories of phenomenal representation. I will now argue that it leads to another problem. Reductionism in general faces a problem about multiple realizability. I think with Sider and others that this problem is not decisive for standard reductionists (see Sider 2013 and note 10 of the present paper). However, because of their endorsement of conceivability arguments about experience, reductive Russellian monists face an especially intractable version of the problem.

The problem is this. We can conceive of structural zombies and structural inverts and so on. That is why Russellian monists think that these scenarios are possible, contrary to standard physicalism. But we can also conceive of a quite different, even more radical sort of scenario: mereologically simple things having all the same experiences as us. For instance, it’s certainly conceivable that mereologically simple things (“souls) could have the same tomato-like experience that Russell has, or the same headache that Russell has. So, by parity of reasoning, Russellian monists ought to accept that this is possible. That is, they should accept a radical form of multiple realizability. In the actual world, having a tomato-like experience is realized in complex systems like me and you. In other worlds, it is realized in mereologically simple things.

This extreme form of multiple realizability is hard to square with reductive Russellian monism. For instance, previously, we considered the idea that having a tomato-like experience (*E*) is necessarily identical with having some micro-parts with some of the “right” quiddities arranged into some “right” kind of brain state

B. We called this B+. But the possibility of a simple thing having the same tomato-like experience rules out this view, since mereologically simple thing cannot have B+.

Standard reductionists don't face such an extreme problem about multiple realizability. Unlike Russellian monists, they reject conceivability reasoning when it comes to experience (§1.1). That is how they respond to the Zombie problem. So they can say that, while "simple experience" is conceivable, it is not possible. That is, they can deny that experience is "realizable in" a simple thing. In general, they don't have to be super liberal about multiple realizability, because they can always say that non-actual cases multiple realizability are impossible even if they are conceivable. By contrast, Russellian monists, since their view is founded on conceivability considerations, are under pressure to accept the possibility of a simple thing having experience – a possibility which seems to undermine reductive versions of the view.²⁵

Notice that this adds to my previous argument that reductive Russellian monists cannot give a quiddity-involving reductive account of the phenomenal representation relation. What reductive account of this relation is compatible with the possibility that a *simple thing* should have a tomato-like experience and bear this relation to redness and roundness?

As always, there are possible replies. Here are two. *Disjunctivism*: Reductive Russellian monists might say that simple creatures have a tomato-like experience by virtue of having some alien, non-structural property S that simple things can have. Then they might say that the property of having a tomato-like experience is actually a *disjunctive* property: *having B+ or S*. So when Russell's speaks of his experience, he is actually referring to a highly disjunctive property. This trick of course accommodates the possibility of a simple thing having a tomato-like experience. *Denial*: Alternatively, reductive Russellian monists might say that a simple thing having experiences is impossible, even if it is *prima facie* conceivable. After all, they *already* hold that, if only we grasped the micro phenomenal quiddities instantiated in Russell's brain, we would see *a priori* that having parts with those properties, in the right combination, entails having a headache. So, they might say, we should be open to view that the entailment also goes in the opposite direction. Maybe, if only we grasped the micro phenomenal quiddities, we would also see *a priori* that having experience requires having parts with the "right" quiddities.

However, neither reply is an entirely comfortable one for the reductive Russellian monist. The idea that having a tomato-like experience is a highly disjunctive property (with disjuncts of a radically different nature) just seems wrong. Also it would make the problem of phenomenal reference (§2.1) even more intractable. How could Russell be referring to this specific disjunctive property, which has an *alien* property as a disjunct? What is the mechanism of phenomenal reference which explains this? As for the option of denying the possibility of a simple thing having experiences, it is clearly at odds with the Russellian monist's liberal use of a conceivability-possibility link in motivating her rejection of standard physicalism.²⁶

²⁵ Thanks to Philip Goff and Brian Cutter for emphasizing this problem for reductive Russellian monism in some helpful discussions.

²⁶ **Here is a somewhat similar problem for pansychist versions of reductive Russellian monism. On this view, both complex things (like me and you) and simple things (fundamental particles or whatever) have the determinable property *having experience*. What *reductive* account of this**

2.4 Conclusion

Here now is the first horn of my dilemma:

First horn: Reductive Russellian monism faces especially intractable versions of the worst problems with reductive physicalism (§§2.1-2.3). (In fact, in §4, I argue that *all* form of Russellian monism also share a profound problem with dualism about “psychophysical luck”. If this is right, then reductive Russellian monism *combines* with worst problems with physicalism with one of the worst problems with dualism.) It can hardly be said to achieve Chalmers’s Hegelian synthesis of dualism and physicalism. Even if you think progress can be made in solving a particular problem, the cumulative case against the view is overwhelming.

3. Problems for Primitivist Russellian Monism

We just saw that reductive Russellian monism fails. This brings me to the second horn of my dilemma: primitivist Russellian monism. I myself think that some primitivist view of consciousness is right (understood broadly to include “primitivist physicalism” and dualism), because of the kinds of problems for reductionism discussed in the previous section. But I think it is far from clear primitivist Russellian monism is the best version of primitivism. In particular, I will argue that primitivist Russellian monism faces exactly the same problems as traditional dualism, another version of “primitivism”. So, contrary to Chalmers’s Hegelian synthesis argument, there is no version of Russellian monism that “combines the virtues of dualism and physicalism while avoiding their vices”.

My plan in this section is as follows. In §3.1, I explain the version of primitivist Russellian monism I will focus on, and I will explain how it avoids the problems with reductive Russellian monism discussed in the previous section. In §3.2, I will argue that it is very similar to traditional dualism and faces the same problems.²⁷

3.1 A Version of Primitivist Russellian Monism

The problems discussed in the previous section make it all but impossible to be a reductive Russellian monist. But they can all be avoided by a version of primitivist Russellian monism.

Let’s start with the *problem of phenomenal representation* (§2.1). In this paper, I am assuming a minimal form of representationalism, which is very widely accepted. On this view, when Russell has his tomato-like hallucination, he stands in “the phenomenal representation relation” to sensible redness and roundness. We have seen that there is reason to think that Russellian monists cannot provide any re-

property could be compatible with this? (A similar problem arises for reductive panqualitism regarding having a quality.)

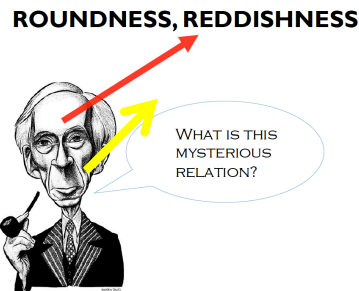
²⁷ In Pautz (2010) I briefly argue that primitivist physicalism in general (which I there call “primitivist physicalism”) is no better than dualism.

ductive theory of this relation. They cannot provide a bi-conditional real definition of the relation. For instance, they can't identify it with a quiddity-neutral "tracking relation" (Tye, Dretske), because that is open to a conceivability argument.

This is not a problem for primitivist Russellian monists. Here is how Russellian monists can account for the representational character of experience. They can say that, if a super intelligence knew that Russell is in quiddity-involving brain state B+, and if she knew the nature of the quiddities involved in B+, she could *a priori* deduce that he is having a tomato-like experience. Since having this experience is inseparable from phenomenally representing sensible redness and roundness, she could also *a priori* deduce that Russell phenomenally represents these sensible properties. That is to say, it would be inconceivable to her that Russell have B+ but fail to phenomenally represent these properties. So, according to this view, Russell's having the quiddity-involving brain state B+ *grounds* his bearing the phenomenal representation relation to sensible redness and roundness. At the same time, primitivist Russellian monists can just concede that there is no general *reduction* of this relation, for the reasons given above. They can agree that they cannot identify it with a tracking relation (Dretske, Tye), for instance. They can just say it is a "simple" relation. That is, there is no general completion of the schema [PR]. This is perfectly consistent. It is consistent to say that the phenomenal representation relation is a simple and irreducible relation, but that a subject's bearing this relation to some sensible properties is always grounded in his quiddity-involving brain state.²⁸ On this view, it is just a kind of brute "grounding law" that, necessarily, if some one has B+, then he stands in the simple phenomenal representation relation to sensible redness and roundness (the term "grounding law" comes from Rosen 2010). Here is an analogy. A follower of G. E. Moore might say that goodness is a simple property (it has no bi-conditional reduction), but that its instantiation is *a priori* grounded in complex natural facts.

In short, the primitivist Russellian monist can accept the following picture of the situation:

Picture 1: primitivist Russellian Monism. The upward arrow indicates the mysterious form of grounding that holds on this view between Russell's brain state and his phenomenally representing redness and roundness.



Picture 1: primitivist Russellian monism

Notice that this is an "internalist" theory of visual experience and phenomenal representation. The character of visual experience, and what properties one phenome-

²⁸ They can also say it is indeterminate what precise neural-quiddity conditions ground phenomenal representation (somewhat as Mark Johnston and others say it is indeterminate what sums of particles constitutes the statue).

nally represents, is fixed by the brain. But this is a good thing. Elsewhere (2013) I have argued for internalism on empirical grounds, and Chalmers (2006) has argued for internalism on a priori grounds.

A remaining question is the traditional issue of the “sensible properties”. Chalmers (2006) and Pautz (2006, 2016) defend the view a neo-Galilean view on which they are simple properties that nothing has. McGinn (1996) defends a kind of realist, response-dependent primitivist view. Primitivist Russellian monists could go either way on this issue.

By the way, it is worth mentioning that *primitivist* Russellian monists, unlike reductive Russellian monists, can accept “revelation”, another Russellian doctrine. That is, they can accept that, when Russell phenomenally represents redness, he is in a position to know redness “perfectly and completely” so that “no further knowledge of [its essence] is even theoretically possible”, as Russell himself put it in *The Problems of Philosophy*. And they can accept that he is in a position to know the complete essence of the relation of phenomenal representation that he bears to this quality. True, the primitivist Russellian monist says that he standing in this relation to his quality is *grounded in* his having B+. And Russell is *not* in a position to know this. But for the primitivist there is no threat here to revelation. For the primitivist Russellian monist can say that this truth about the *grounds* of phenomenally representing redness does not touch on the *essence* of phenomenally representing redness itself. (This would require that he give up Fine’s principle that grounding derives from essence.)

Next, let us turn to the *problem of phenomenal reference* (§2.1) for reductive Russellian monists. Remember this problem concerned how, when Russell’s uses “this experience”, he might be determinately referring to the complex quiddity-involving neural property B+, as the *reductive* Russellian monist view requires, given that he also has the complex quiddity-neutral neural property B which is an equally good candidate referent. On primitivist Russellian monism, the problem goes away. On this view, Russell is *not* determinately referring to the complex quiddity-involving neural property B+. Instead, he is referring to his tomato-like experience-type – call it E. And, on the primitivist picture, his experience is not identical with B+. Rather, E is an irreducible state (involving standing in a simple relation of phenomenal representation to a simple quality, redness). True, E is grounded in B+; but E is distinct from B+, on the primitivist picture. It is, of course, also distinct from B. (In fact, on this view, whereas B+ grounds E, B doesn’t ground anything interesting at all, and certainly doesn’t ground an experiential condition.)

True, the primitivist Russellian monist also faces a question about phenomenal reference, namely: how is it that Russell manages to determinately refer to E, as his primitivist view requires, rather than for instance to B+ (the *ground* of E) or to B? But, unlike the reductive Russellian monist, the primitivist Russellian monist can easily answer this kind of question, since he thinks that E “stands out”. Let me briefly mention three types of answers he can choose between. (1) *Reference magnetism*. E, by contrast with B+ and B, stands out as being 100% natural, and hence a “reference magnet” for Russell’s demonstrative “this experience”. (2) *Epistemic specialness*. By contrast to B and B+, E also stands out in being epistemically special: if you have this property, then you have a special super-justification for believing you have this property. Further, given a rationality-based theory of belief invoking a principle of charity/humanity (Davidson, Lewis), it follows that the correct inter-

pretation of Russell is that his demonstrative beliefs is determinately about E, rather than B or B+.²⁹ (3) *Acquaintance*. Chalmers (2013) claims that it is just in the nature of experience properties, and only experience properties, that having such a property grounds “being acquainted with” one’s having that property. On primitivist Russellian monism, neither B+ nor B is itself an experience, but B+ grounds the tomato-like experience E. Thus, given Chalmers’s claim, having E, but not B+ or B, necessitates that Russell is acquainted with itself. This in turn might explain why Russell refers, with “THIS experience”, to E, not B+ or B. Unlike reductive Russellian monists, primitivist Russellian monists don’t need a reductive account of this relation. They can say it is a *simple* inner-directed relation (just as they say that the phenomenal representation relation is a simple outer-directed relation). They only need a “grounding account”. And this they can easily supply. Given Chalmers’s claim, Russell’s standing in the acquaintance relation to his having E is grounded in his having E. And, on primitivist Russellian monism, his having E is in turn grounded in his having B+.³⁰

Finally, let us look at the problem of *multiple realizability*. While this is a problem for reductive Russellian monists, it is not a problem for primitivist Russellian monists. It is a problem for reductive Russellian monists because identifications or real definitions entail necessary bi-conditionals. By contrast, grounding, the key notion of primitivist Russellian monism, is a conditional notion. Thus, primitivist Russellian monists can say that having a tomato-like experience is grounded in B+ in our world but totally different quiddities in other worlds (just not *non-phenomenal* quiddities because, remember, they think that macro experiences can only be grounded in *phenomenal* micro-quiddities).

In fact, they can even accept the possibility of mereologically simple things having experiences. The idea is that, while having a tomato-like experience is grounded in more basic quiddity-involving conditions in our world and nearby worlds, in far away worlds it is not grounded in anything at all. This requires that the self-same property can be grounded in one world, but in other worlds occur without being grounded. But maybe this is a coherent idea. So the possibility of “simple subjects” is not a problem for *primitivist* Russellian monism.

A final point. For purposes of illustration, I have focused on Russell’s tomato-like visual experience, which is a paradigm of a representational experience. But, of

²⁹ Option (1) invokes a toy theory of content often attributed to Lewis but that he never accepted. Option (2) invokes a theory of mental content that is closer to Lewis’s actual view. See Pautz 2013b and Weatherson 2013.

³⁰ There is of course still the “combination problem”. And there is another problem. While reductive physicalism may lead to radical indeterminacy, primitivist physicalism may lead to hyper-determinacy. On this view, since experiences are super-eligible or stand out referents, and since you only have one determinate total experience at a time, there is no (or very little) indeterminacy in the reference of certain of our phenomenal demonstratives (“THIS experience”), and perhaps in the reference of the determinable expression “has an experience”. These expressions *determinately* refer to certain properties P, Q, R, . . . Now suppose there is no “wordly indeterminacy”: necessarily, for any property X and any thing y, either it is determinate that y instantiates X or it is determination that y fails to instantiate X. These two claims imply that there can be no indeterminate issues concerning the instantiation of certain experiences. For instance, there was a first moment when a system determinately had an experience, before which it was determinate that that none had any experience, etc etc. (And, if you are a complex thing, then, among the multitude of complex things in your vicinity, there is a single one that determinately has THIS experience, and so on.) To avoid this result, the primitivist would need to accept “worldly indeterminacy”: in some cases, we determinately refer to certain experience-types but it can be (worldly) indeterminate whether they are instantiated by something.

course, the primitivist Russellian monist will generalize the view, *mutatis mutatis*, to other experiences, including non-representational experiences if such there be (the feeling undirected depression perhaps).

In what follows, when I speak of primitivist Russellian monism, I will have in mind this version of it.

3.2 Primitivist Russellian monism shares dualism's problems

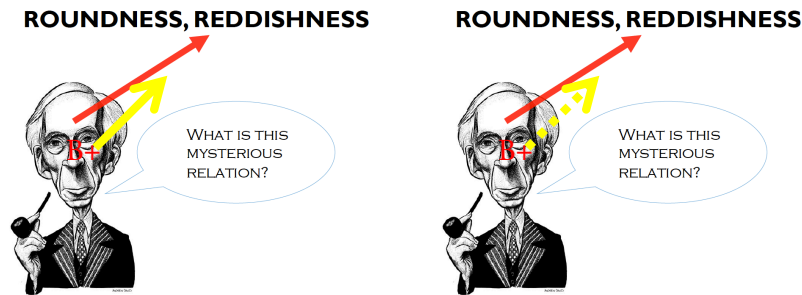
We saw in the previous section that primitivist Russellian monism may solve the problems specific to reductive Russellian monism. So there is a case for primitivist Russellian monism. But I will now argue that it faces a separate set of problems. In fact, it faces the very same problems as traditional dualism. So there is no strong argument for accepting it over a parallel form of dualism.

To see this, let me first describe such a parallel version of dualism. The version of dualism I have in mind can be described quickly: it is exactly like the version of Russellian monism discussed above, but for one difference: wherever the primitivist Russellian monist posits a brute, necessary grounding law, the dualist posits a brute, contingent psychophysical law. In fact, to make the views even more parallel, we could suppose that the relevant form of dualism, like primitivist Russellian monism, recognizes the “quidditistic facts” in addition to the structural facts that holds that phenomenal quiddities are widespread at the micro level. *In that case, their only difference concerns the nature and strength of the modal connection between these facts and the irreducible experiential facts at the macro level.*

For instance, we have seen that, on primitivist Russellian monism, there is a brute “grounding law” that, necessarily, if someone (Russell himself in our example) has B+, then this grounds his standing in the simple phenomenal representation relation to the simple color, redness. So, on this view, it is metaphysically impossible that Russell should have a Zombie twin: a total physical duplicate of Russell that has the quiddity-involving neural state B+ but doesn't bear the simple phenomenal representation relation to redness. The version of dualism I have in mind is exactly the same, but it replaces this with a parallel brute *psychophysical law*: a brute, *contingent law of nature* that, if some one has the quiddity-involving neural state B+, then he stands in the simple phenomenal representation relation to redness.³¹ So, on this parallel dualist view, it *is* metaphysically possible that a physical duplicate of Russell with B+ should be a Zombie that fails to bear the simple phenomenal representation relation to simple redness – because the relevant law is contingent, it might have failed. This is modal difference is the only difference between the views. When I speak of “dualism” in what follows, I shall have in mind this form of dualism.

³¹ I assume that the “brute law” will have this form for simplicity. More realistically, the basic psychophysical laws will be more general than this (cf. Chalmers's online reply to Latham; and Chalmers 1996, p. 160, 214). For instance, there may be a single psychophysical law for the experience of color, along the following lines: it is nomically necessary that, if an individual has brain state N+, then they phenomenally represent sensible color $f(N+)$, where f is some systematic mapping from brain states onto sensible colors. (R. M. Adams raises profound puzzles about whether this kind of generality is achievable, but I will set them aside.) Likewise, the primitivist Russellian monist might say that her “grounding law” will have a similar general form: it is metaphysically necessary that, if an individual has brain state N+, then this grounds the fact that they phenomenally represent sensible color $f(N+)$. I will work with the simple picture in the text for convenience; none of my points (e. g. my point that primitivist Russellian monism is just as complex as dualism) will depend on it.

Here is a picture that illustrates just how similar are primitivist Russellian monism and dualism:



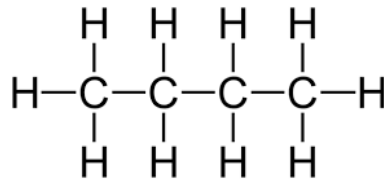
Primitivist Russellian view on the left; the dualist view, on the right. Ontologically, the views are alike. Both hold that Russell's tomato like experience (*E*) is irreducible. Both hold that in having this experience Russell bears the simple "phenomenal representation relation" (the red arrow) to a simple quality, redness. The two views just differ *modally*. In particular, on primitivist Russellian monism, Russell's irreducible experience depends on his quiddity-involving brain state *B+* by way of a *metaphysically necessary* "grounding law" (represented by the solid yellow arrow). By contrast, on dualism, it depends on his quiddity-involving brain state *B+* by way of a merely nomically necessary psychophysical law (represented by the dotted arrow).

Now, Chalmers's Hegelian synthesis argument is that Russellian monism combines the virtues of physicalism and dualism while avoiding their vices. But this cannot be said of primitivist Russellian monism. In fact, given its close similarity to dualism, it shares all the vices of dualism. To show this, I will describe each of the main problems with dualism, and show that primitivist Russellian monism shares that very same problem.

I. The problem of non-uniformity. One virtue of physicalism is supposed to be that it provides a uniform view of nature. By contrast, dualism sees the universe as radically non-uniform. On this view, something happens in human brains that is radically different from what happens in the rest of nature. J. C. Smart said he found this "frankly unbelievable".

Now I will argue that primitivist Russellian monism is also radically non-uniform. To see this, I will contrast the "emergence" of Russell's tomato-like experience (*E*), under primitivist Russellian monism, with the formation of a type of chemical, *butane*, in the non-mental world.

True, on primitivist Russellian monism, in both cases, their occurrence is supposed to be *a priori* deducible from the fundamental physical facts. Nevertheless, on this view, the occurrence Russell's experience is totally *sui generis* and different from anything else in nature:



it is form of what I will call "weird emergence". Just look at picture 2. The emergence of Russell's experience is weird in several respects. After explaining how it is weird, I will show that the occurrence of butane is not weird in those respects.

First, *E* is *irreducible*. So when Russell's quiddity-involving brain state *B+* necessitates his having *E*, then this is due to a special, brute "grounding law". In fact, since *E* is multiply realizable with respect to physical states, its emergence requires a whole slew of special grounding laws, *that only kick in when individuals get cer-*

tain brain states. Compare the emergence of G. E. Moore's primitive property of being good. In this regard it is just as non-uniform as dualism.

Second, Russell's having E involves standing in a simple phenomenal representation relation to sensible redness and roundness. But this is grounded in Russell's having the brain state B+, which doesn't itself involve standing in a relation to sensible redness and roundness. Thus the grounded fact and the grounding fact are very different, indeed radically *non-congruent*. This doesn't happen elsewhere in nature. For instance, the fact that you stand in the relationship of friendship to Johnny is itself grounded in your standing in certain (psychological and social) *relations to Johnny*. The grounded fact and the grounding fact are congruent.³²

Third, E belongs to a "stand out" determinable: having a conscious experience. This determinable is radically different from all other determinables that appear at the macro level. The determinates in this determinable have an "internal unity" form an extremely "natural class". They also have a unique normative significance. For instance, having an experience necessarily provides immediate justification for believing one has that experience (Chalmers, Smithies, etc).³³ Some – for instance Chalmers – hold that this is because having experiences is special in that it entails standing in a simple acquaintance relation to the fact that one has the experience. On primitivist Russellian monism, this "stand out" character of experiences is what makes them "reference magnets", thus solving the problem of phenomenal reference. So primitivist Russellian monism is non-uniform in yet another way: it requires that certain of our brain states ground properties belonging to a determinable that is radically different from all other determinables that appear at the macro level in nature. *In short, they retain the dualist idea that consciousness "stands out in nature" – even while trying to have a physicalist view.*³⁴

Now contrast this kind of "weird" emergence with the totally non-weird occurrence of *being butane*. This property is reducible. To be butane *just is* to have as parts carbon and hydrogen in a certain arrangement. To explain its occurrence, no special principles are required. We only need general, quasi-logical principles going from simple predicates to complex predicates. Also, the occurrence of this property doesn't involve any kind of weird "non-congruent grounding". Finally, this property doesn't belong to any normatively special, super natural "stand out" determinable. It belongs to many determinables, where none of them stands out as being especially natural. And, of course, none stands out as being normatively or epistemically special.

Of course, I picked a simple property, the property of being butane. But I think that parallel remarks apply to all other properties at the macro level, like *being a*

³² Numbers, instantiation.

³³ This in fact is part of one solution to the problem of phenomenal reference.

³⁴ In fact, on primitivist Russellian monism, not even the brain states that *ground* our experiences belong to such a stand out determinable. Take the brain states that ground conscious experiences. True, they may have something in common (recurrent processing, high Φ level, or whatever) and so fall into a determinable. But any brain state falls into *multiple* determinables or families; and it is not true that one of these sets of brain states stands out in the radical way in which experiences stand out on primitivist Russellian monism. No such family of brain states is super-natural in the way in which the family of experiences is super-natural or possesses high internal-unity; and none is united in having special epistemic significance (on primitivist Russellian monism, it is our experiences, not the brain states that ground them, that have this significance).

rock.³⁵ Here is a kind of continuity argument for the claim that these properties, too, are reducible. Consider the evolution of the universe after the big bang but prior to the appearance of conscious creatures. Obviously, at the *very* beginning, the only macro properties that appeared were complex properties reducible to the fundamental physical properties, like *being butane*. So reductionism reigned at the very beginning. Where in the evolution of the universe prior to the appearance of conscious creatures did properties appear for which reductionism is *not* true? It is implausible that there was such a moment. So the most reasonable and simplest view is that even properties like *being a rock*, *being a mountain*, and *being a tree* are reducible. They are just like *being butane*, or *being a pixel face* in the pixel world discussed in §1. To be a rock, or a mountain, or a tree *just is* for the fundamental physical and topic neutral properties to be arranged in certain ways. So to explain their occurrence we just need general principles for forming complex predicates – no additional, special “grounding laws”. Also, the occurrence of these properties, too, doesn’t involve any peculiar sort of “non-congruence”. And they don’t belong to any super-natural, normatively significant determinables.³⁶

In fact, to show how weirdly non-uniform primitivist Russellian monism is, we don’t have to consider mountains and other things outside the brain. It is massively non-uniform *within the brain*. On this view, B+ is connected through a special grounding law with *standing in the simple phenomenal representation relation to sensible redness and roundness*. But it is known that a great deal of neural processing can occur in the total absence of phenomenal consciousness. Let B* be a brain state that occurs under these conditions. We can suppose that B* is broadly similar B+. B* involves the same micro-level quiddities as B+ (after all, they are uniform in nature as we discussed in connection with the “big difference” problem

³⁵ David Lewis raises a non-uniformity problem for emergentism about ethical properties. It is reported by Frank Jackson (1998, 127) as follows:

We can distinguish a more and a less extreme {version of ethical emergentism}. The extreme view says that for every (contingent) descriptive way there is, there is a quite distinct, necessarily [connected] ethical way there is. This extreme version is hard to take seriously. It seems an absurdly anti-Occamist multiplication of properties: for every descriptive property, we have a corresponding non-descriptive one! **But if the idea is that the duplication only happens occasionally, where is the principled basis for saying when it happens and when it does not?**

Lewis’s point is that, since “reductionism” is certainly true in most of nature, a local primitivist view requires a kind of arbitrary-looking non-uniformity in nature. That is the point I am pressing.

³⁶ David Chalmers has suggested to me that, under primitivist Russellian monism, the “emergence” of Russell’s experience from his brain state is no more odd than the “emergence” of the property of being red from the property of being scarlet.

I disagree. The relationship between scarlet and redness is an instance of the more general relationship between determinates and determinables – which in turn might be an instance of the even more general relationship between disjuncts and disjunctions (if determinables are disjunctions of their determinates – see Rosen 2010). By contrast, if primitivist Russellian monism is true, then, to explain the “emergence” of Russell’s experience from his brain state, we need to posit a “special” sort grounding law, which is unique to experience and that only kicks in when individual get brain states of a certain complexity. This ground law is a “dangler” which cannot be derived from any other familiar sort of grounding connection. Moreover, as I explain in the text, it will relate radically disparate, non-congruent kinds of states (namely, Russell’s brain monadic state B+, and his standing in the primitive representation relation to redness, and perhaps also his standing in an irreducible acquaintance relation to this).

in §1.1). B* and B+ involve neurons firing, and the types of neurons are the same everywhere within the cortex (Prinz 2012, 129). B* and B+ just involve different patterns of firing, and perhaps different brain regions. Still, on primitivist Russellian monism, there is a giant difference between them: B+ grounds Russell's standing in the simple phenomenal representation relation to redness and roundness, whereas B* doesn't ground anything like this or indeed anything of interest at all.

For all these reasons, just like dualism, primitivist Russellian monism requires that, in connection with certain brain states, something truly weird happens: properties "emerge" at the macro *whose nature and origin* is totally different from the nature and origin of all other properties that appear in nature at the macro level, in mountains, rocks and trees (and indeed totally different from what happens in connection with other brain areas). It requires special inter-level principles ("grounding laws") that are specific to consciousness. They only kick in when certain brain states occur, and they are like no other principles that operate in nature. In short, primitivist Russellian monism is just as non-uniform as dualism. So considerations of uniformity cannot provide a strong argument for preferring primitivist Russellian monism to dualism.³⁷

II. The Complexity Problem. Dualism requires a swarm of *brute, special* psychophysical laws going from the physical states of the brain to the endless variation of experiences we can have. They are *brute*, in the sense that they cannot be derived from more basic, general principles. For instance, they cannot be derived from the laws of physics. They are "nomological danglers". They are "special principles" in that they are specific to conscious experience and they only kick in when certain brain states occur.

Now, Chalmers says that one of the main motivations of Russellian monism is to avoid the complexity of dualism. It is meant to be economical view of the world. *Reductive* Russellian monism has this virtue. But primitivist Russellian monism is clearly just as complex as dualism. For each and every brute psychophysical law that dualism requires, primitivist Russellian monism requires a parallel "grounding law". For instance, the dualist claims that it is a nomic law that, if some one has B+, then he stands in the simple phenomenal representation relation to the simple quality, redness. By contrast, the primitivist Russellian monist says that it is a grounding law that, necessarily, whenever some one has B+, then this grounds the further fact that he stands in the simple phenomenal representation relation to the simple quality, redness. Such "grounding laws" are "danglers" just as much as the dualist's psychophysical laws. They cannot be derived from any other general principles of modal metaphysics: general claims about logic, essence, grounding. They are *special* principles, which are just true of conscious experiences. It is only conscious experiences that are grounded in brain states in this way. Furthermore, they add to the complexity of the view, just as dualism's psychophysical laws add to the complexity of dualism. *So primitivist Russellian monism and dualism are equally complex.*³⁸

³⁷ True, elsewhere in nature there are properties that have *some* of these features. For instance, masses and charges are perfectly natural and extremely potent similarity/difference makers. But, elsewhere in nature, they only appear at the *micro-level*. What primitivist RM requires is that this sort of thing happens at the macro-level in connection with brains – and at the macro level this is *not* something that happens elsewhere in nature.

³⁸ (i) Some philosophers (Bennett MS, Schaffer 2014?) might suggest that to measure the complexity of a theory you only need to take into account its fundamental ontology. But I think that there are

In response, the primitivist Russellian monist might point out that her brute “grounding laws” are supposed to be in principle knowable *a priori*, even if we don’t currently know them *a priori*. This is because Russellian monism is a form of *a priori* physicalism. (see §1.1 and §3.1). By contrast, the dualist’s psychophysical laws are only knowable empirically. Doesn’t this make a difference to the complexity of their views?

In fact, this *doesn’t* make a difference when it comes to assessing the relative complexity of primitivist Russellian monism and dualism. To see this, consider an analogy. Imagine two philosophers, the Empiricist and the Rationalist. They have exactly the same *fundamental physical theory* of the universe. They totally agree about what the fundamental physical laws are. They only differ as regards the status of these laws. The Empiricist thinks that they are contingent truths that are only knowable empirically. By contrast, the Rationalist holds that they are necessary truths (he accepts a version of the “Shoemaker-Swoyer” view). He also speculates that they are in principle knowable *a priori*. He speculates that, if only we knew the natures of the fundamental physical properties, we would “see” that they must figure in just these laws.³⁹ Now, do we have a reason to prefer the Rationalist’s physical theory to the Empiricist’s physical theory, on the grounds that it is *simpler*? Obviously not! The Rationalist and the Empiricist posit the same set of brute, fundamental laws. Their difference over the (modal and epistemic) status of these laws does *not* make the Rationalist’s theory simpler. Likewise, the dualist and the primitivist Russellian have exactly parallel psychophysical theories of the mind-brain connection, positing exactly parallel inter-level psychophysical principles. Their difference over the *epistemic status* of those principles (the fact that the Russellian monist speculates that they are knowable *a priori*) doesn’t make the primitivist Russellian’s system of psychophysical principles *simpler* than the dualist’s exactly parallel system.⁴⁰

Let me briefly consider another response to my contention that primitivist Russellian monism is just as complicated as dualism. Kit Fine (e. g. 2012: §11) holds that, whenever the fact that C is grounded in other facts, then there is a gen-

clear counterexamples. For instance, take two philosophers. One, Ted, is an austere *reductive* physicalist. For Ted, experiences are just brain states and so on. The other, Mark, is a funny sort of physicalist who combines physicalism with old fashioned sense datum theory. He says that, when you see a tomato, then your having a certain brain state grounds the coming-into-existence of a red and round sense datum, which is like nothing in your brain; and it also grounds your being acquainted with this sense datum. When he says that “there are sense data grounded in brain states” and “there are electrons”, he is using “there are” with the same meaning in each case. *Clearly Mark’s theory is more complex than Ted’s*. It has a more bloated ontology. And it requires all these funny grounding laws, linking brain states with the coming-into-existence of sense data, and our acquaintance with them. (Did Karen disagree when I brought up this case with her at her talk?) In the same way, the “grounding laws” of primitivist Russellian monism undeniably contribute to the complexity of that theory. On primitivist Russellian, they must be included in our complete theory of the world.

(ii) *Emails with Kit Fine*: I was wondering whether you think that there is some more general & basic essentialist truth about what it is for a mountain to exist which has E1, E2, as consequences. *Kit*: Not sure, but if there is a general principle then I think it would have to be formulated with the use of vague terms. Thus the particles would have to constitute something massive but being massive is a vague matter. Once we had the appropriate vague terms at our disposal then perhaps the problem would not be so difficult.

³⁹ Spinoza accepted a view along these lines – a kind of causal rationalism. See pp. 29-30 of Bennett, *A Study of Spinoza’s ethics*.

⁴⁰ Another idea: grounding laws are not especially bad on Sider’s “big list” theory of metaphysical necessity. //Daniel S told me he has a Fine-like view but with the laws flowing from quiddities?

eralization that logically entails the particular grounding connection and that holds in virtue of the essence of some item involved in the fact that C. If primitivist Russellian monists accept this Finean view, then they will say that “grounding laws” are not brute as I have been assuming, but rather derive from more basic essentialist truths. So, for instance, they might say that *it is just “in the essence” of standing in the primitive relation of phenomenally representation to redness that, if some one is in brain state B, then this grounds the fact that they stand in this primitive relation to redness.* (Compare: even if yellowness is a primitive quality having no definition in other terms, as many think, it may be in its essence that it implies extension.)⁴¹ This more basic truth explains the grounding law. So whereas dualism requires a system of brute “psychophysical laws”, primitivist Russellian monism requires a system of brute “essentialist laws”. But – and here is the crucial point – they might just assert that such “essentialist laws” don’t add to the complexity of their theory in the way that psychophysical laws add to the complexity of dualism. The idea (I guess) is that such essentialist truths just flow from the nature of the items involved. They just “come for free”.⁴²

I have two points in reply. First, if primitivist Russellian monists take this Finean route, then a cost is that they must after all give up on the thesis of “revelation” discussed in §3.1. For on this Finean view, there are “hidden” facts about the essences of experience properties and qualities that certainly aren’t revealed by mere experience and reflection. Second, the view *doesn’t* have the advertised payoff anyway: it doesn’t help with the complexity problem. For now, where dualism requires

⁴¹ You might think that this sort of essentialist claim is just incoherent. How can it be in the essence of a *primitive relation* to be grounded in the physical in so and so ways? But Fine puts no restrictions on what can fill in “X” and “p” in “It is in the essence of X that p”. (This leads to a worry expressed by Ted Sider:

I have a bit of a concern . . . with Fine on essence. The initial examples he uses (e.g. Socrates and singleton-Socrates) suggest that claims about essence can be underwritten by a special sort of explanatory story. In the case of Singleton Socrates, the underlying explanatory story involves the generation of all sets by an operation of set-building (as in Fine’s paper “Towards a theory of part”); the essential properties of each set arise because of how that set is built. But the eventual locution Fine uses to talk about essence is so much more powerful: **a general operator, where “p” can be any sentence and X is any list of things.** My complaint is that we are given no hint of how to understand these claims in terms of the tamer kinds of essential facts.

So, there is nothing about the “it is in the essence of X that p” notion that forbids the view that it is in the essence of a *primitive relation* to be grounded in the physical in so and so ways. Here is another coherent view: (i) the property of being something you ought to do is *primitive* in the sense that there is no interesting identification of the form “to be something you ought to do is to be F” which would result in a biconditional, necessary and sufficient claim “necessarily something is good iff it is F”. (ii) In fact, this property is very *different from* other properties in various ways – it is 100% natural, has special reason-giving powers, etc. (iii) Still, it is part of the essence that certain other properties are *sufficient* for this special property, for instance, the property of *saving a kid’s life at no cost*.

I think that examples like this show that one-directional essence claims alone aren’t *sufficient* for a plausible form of physicalism (contrary to Dasgupta “The Possibility of Physicalism”). Indeed, the following view would count as physicalist by this criterion: there are non-physical sense data, but it is in the essence of being acquainted with (say) a red and round sense datum that brain state B is sufficient for it, etc!!!

⁴² Some people says that (constitutive) essentialist truths “don’t cry out for explanation”. The claim we’re looking at now is different: it is that they don’t “add to the complexity of a theory”.

a system of brute “psychophysical laws”, primitivist Russellian monism requires a system of exactly parallel brute “essentialist laws”. While the psychophysical laws “dangle” from the rest of the body of laws of nature (the physical laws), these essentialist laws “dangle” from the rest of the body of essentialist truths. Consider for instance standard essentialist truths relating determinates to determinables, relating complex properties to simpler properties (e. g. it is in the essence of a conjunction that it is grounded in its conjuncts), and so on. The primitivist Russellian monist who takes the Finean route needs a slew of *additional* essentialist laws that are special to consciousness and that are exactly parallel in form the psychophysical laws of dualism. Our only reason to believe them is to account for the emergence of macro level experiences. They must also be included in any complete inventory of the truths of our world. (By contrast, as we saw, the reductionist doesn’t need any such special principles.)

Therefore, my original claim stands: primitivist Russellian monism is just as complicated as dualism. We cannot use “simplicity considerations” to support accepting this view over traditional dualism.

III. The problem of mental causation. Does primitivist Russellian monism have an advantage over dualism when it comes to solving “the problem of mental causation”?

No – *given the similarity between the views, they have the same options when it comes to mental causation. To illustrate, look back at Picture 2.* Suppose that Russell, while having his tomato-like experience, utters the sentence “there is a red and round thing there”. Both the primitivist Russellian monist and the dualist have two options: they can accept overdetermination, or they can accept epiphenomenalism. On the epiphenomenalist option, Russell’s brain state B+, *but not his experience E*, causes his utterance. On the overdetermination option, *both* are causes of his experience.⁴³

Perhaps it will be replied that under primitivist Russellian monism the overdetermination option is more palatable than it is under dualism. On dualism, it would have to be a form of overdetermination that is like no other kind of overdetermination in the rest of nature (Jackson and Braddon-Mitchell, 2007 p. 17). By contrast, on primitivist Russellian monism, it would be more like the kind of benign overdetermination that we encounter in the rest of nature, such as the “overdetermination” that happens when a pigeon trained to peck at red things can be said to peck at a color-chip “because it is scarlet” *and* “because it is red”.

⁴³ Perhaps it will be said that dualists but not primitivist Russellian monists are *committed* to the epiphenomenalist option and cannot accept the overdetermination option, given a counterfactual “dependency” view of causation (Loewer 2001, 51-52.). But I think that the issue is not so clear-cut. For one thing, Chalmers (1996, 192) notes that, under a *non-reductive* theory of causal and nomic facts (Armstrong-Tooley), dualists are certainly not committed to epiphenomenalism and can certainly accept overdetermination (see also Bealer XX). For another, there is some reason to think that, given a counterfactual “dependency” view of causation, even primitivist Russellian monists may be committed to epiphenomenalism. Here is why. In the Russell example, if Russell had failed to have his tomato-like experience E, would he also have failed to say “hey there is a red and round thing there”? Well, on Russellian monism, this is not obvious. On this view, one way for this to happen would have been for there to have been *local quiddity-switch*, where the phenomenal quiddities in Russell’s brain were replaced by non-phenomenal quiddities (Hawthorne 2006, p. 206), but with *all the structural facts staying the same*. This is because, on this view, mere quiddity switching is enough to turn Russell into a *functionally-equivalent Zombie* - “make the lights go out”. In this case, despite not having the experience, Russell would have *behaved in exactly the same way*.

But this response ignores all of the previous discussion – and in particular the discussion of “non-uniformity”. On primitivist Russellian monism as on dualism, the relevant overdetermination would be supported by a kind of dependence of experience on the brain that is like *nothing else in nature*. It is *not* like the dependence of being red and being scarlet. That is an instance of the more general dependence of a determinable on a determinate (or, if a determinable just is a disjunction of its determinates, the even more general dependence of disjunctions on their disjuncts). By contrast, on primitivist Russellian monism, the dependence of Russell’s experience on his brain state B+ derives from a special “grounding law”, which is specific to consciousness, which only kicks in when certain brain states occur, and which links radically non-congruent conditions (the brain state and standing in a primitive representational relation to redness).

So, on primitivist Russellian monism as on dualism, the overdetermination *would* be like no other form of “benign” overdetermination in nature.

IV. The problem of luck. It is sometimes said that dualism faces a problem about “luck”. I believe that in any version Russellian monism faces an exactly parallel problem. This is a large and fascinating issue, so I will address it in a separate section (§4). If I am right, then again there is a parity between the views.

3.3 Conclusion

Now here is the second horn of the dilemma.

Second horn. Chalmers’ Hegelian synthesis argument is that Russellian monism combines the virtues of dualism and physicalism while avoiding their problems. This cannot be said of primitivist Russellian monism. It is extremely similar to dualism and shares exactly the same problems. So there is no strong argument for accepting it over dualism. **In fact, if anything, there is reason to prefer dualism. Primitivist Russellian monism faces the “combination problem” and the “big difference problem” (§1.1).** Dualism avoids these problems. If we are forced to choose between dualism and primitivist Russellian monism, it’d be a toss up.

4. A General Problem for Russellian Monism: Psychophysical Luck

As I said, the central argument of this paper takes the form of a dilemma: both reductive and primitivist Russellian monism are problematic and unmotivated. But I will conclude by developing a problem for *any* form of Russellian monism. The problem is one that Russellian monism shares with dualism. Dualism isn’t just complicated. It faces an additional deep problem I’ve elsewhere called the problem of *psychophysical luck*. Although he hasn’t addressed it in detail, David Chalmers (*p. c.*) has described it as “the most serious worry” for dualism – even worse than the complexity problem. I will explain the problem as it arises for dualism, and then explain why Russellian monism shares the very same problem.

In the case of dualism, the problem starts from the fact that dualists hold that experiences and physical states are “distinct existences” which are modally independent from each other. They believe in “psychophysical laws” but they hold that

these laws are totally contingent. So they are committed to the possibility of *inharmonious worlds*. Here are some examples of what I have in mind:

- In *W1*, the “structural” neural, functional facts are the same as in the actual world. We undergo exactly the same neural processing and exhibit exactly the same physical behavior. But where in the actual world we have rich visual experiences, in *W1* we only have visual experiences of pervasive darkness. It is an “altered qualia case”. In *W1*, we are not visual zombies: we *do* have visual experiences. Only they are experiences of blackness. Otherwise, our experiences are like our experiences in the actual world.
- In *W2*, the “structural” neural, functional facts are the same as in the actual world. We undergo exactly the same neural processing and exhibit exactly the same physical behavior. But where in the actual world we have sweet taste experiences, in *W2* we have rotten-flesh taste experiences. It is another “altered qualia case”. In this world, whenever our twins eat oranges, they get rotten-flesh taste experiences. But, since they neural-functional duplicates, they find themselves acting exactly as we do when we eat oranges – *as if* they like those experiences. Still, they seem worse off than we are.
- In *W3*, *humans* are all zombies but *rocks* constantly have macro-level experiences of severe pain pervading their little rock-bodies.

Now the problem is simply this. On dualism, it is a contingent fact that the psychophysical laws are *generally* “harmonious” rather than “inharmonious”. Does this fact have an explanation or not?

Dualists could say that this fact about the psychophysical laws is just a brute fact without any explanation. But then they are obnoxiously multiplying the types of brute facts we have to believe in. True, they *already* believe in a bunch of brute facts: they accept all these psychophysical laws, over and above the laws of physics. But now they are making their view even worse: they are adding that it is a just a brute fact that they are always have the property of being “harmonious”. This makes their view even more complex. Also, this fact “cries out for explanation”.

On the other hand, dualists could try to cook up some explanation for why the psychophysical laws are always harmonious. For instance, they could say that harmonious psychophysical laws are “simpler” or “more good” than unharmonious laws, and then say that it is some kind of objective law that the universe tends to be “simple” and “good”. But this looks pretty objectionable too. (For one thing, in many inharmonious worlds, the psychophysical laws needn’t be any less simple, or any less “good”, than the actual psychophysical laws.) Alternatively, they could say that the fact that the psychophysical laws are harmonious, when added to the other wonderful facts about the universe (puppies, iPhones, etc.), provides enough evidence of a harmony-loving God that we should *believe* in such a God. But it’s far from obvious that such a conclusion would be warranted (especially if the rational “prior probability” of such a God is extremely low to begin with). (Why not believe in a cheese-loving God whose main goal was to create cheese, given the plenitude

of cheese on earth? Or a suffering-loving God, given all the suffering?) And, anyway, this theistic view is complex and objectionable in other ways.

Standard physicalists avoid the problem of psychophysical luck. On their view, inharmonious worlds like W1-W2 are metaphysically impossible. The experiential facts *can't* vary in these bizarre way with respect to the neural-functional facts. Why? Well, at least, on a reductive version of standard physicalism, the reason is simple: the experiential facts are just *identical with* the neural-functional facts.

Now for my main point: Russellian monism shares dualism's problem of psychophysical luck. The reason is that, like dualists, Russellian monists accept a liberal conceivability-possibility link. So, for instance, they agree that the conceivability argument establishes the real possibility of accessible Zombies. But mismatch worlds are no less conceivable than structural Zombies.⁴⁴ Altered qualia are no less conceivable than absent qualia. So by parity of reasoning they are committed to the possibility of structural mismatch worlds, just as dualists are.

Now the resulting problem for Russellian monists is the same as the problem for dualists: if the links between the neuro-functional and the phenomenal could have been inharmonious in ever so many ways, why is it actually generally so harmonious?

But before we get to this, we have to understand how Russellian monists might accommodate the possibility of inharmonious worlds. Now the way they accommodate the possibility of structural Zombie worlds is by saying that to have states of conscious experience you don't only need the neural states you in fact have. In addition, the micro-quiddities of the parts of your brain must be phenomenal quiddities. What goes wrong in the structural Zombie worlds is that everything is the same in your brain in every accessible way but the hidden micro-quiddities are not phenomenal. They are, so to speak, boring, grey quiddities. So, on their view, phenomenal quiddities make a difference to the *presence or absence* of conscious experience. Therefore, it is reasonable for Russellian monists to claim that they also make a difference to the *character* of experience. More precisely, let S be the total quiddity-neural, structural description of our world, including the neural processing in our brains. Let the actual phenomenal quiddities that realize the mass role, the charge role and the spin role, and so on, be $\langle A, B, C, \dots \rangle$. The combination of S and $\langle A, B, C \rangle$ grounds our actual experiences, which are nicely harmonious with the structural neural and behavioral facts in S. The idea I am suggesting is that in mismatch worlds the phenomenal quiddities are different. They are C, D, E or G, H, I, or whatever. In these worlds, S, together with these other combinations of quiddities, grounds the different experiences our twins have, which are totally inharmonious with the structural neural and behavioral facts in S. Hence these worlds are experientially different at the macro-level, because they are quidditistically different at the micro-level, in "hidden" ways.

Now we can say the form the luck problem takes for this brand of Russellian monism. We saw that for dualism the problem takes this form: why are the psychophysical laws harmonious? That is because for the dualist it is the psychophysi-

⁴⁴ Nagel disagrees? You might claim that it's a priori that the taste of rotten flesh necessarily results in the physical movement of withdrawal, so that this crazy scenario is impossible. This however is a deeply implausible claim. For one thing what if you believe eating more will make the taste go away? At most, what is a priori is that if you have the experience, you ought to desire that it go away – this falls short of an actual link to physical behavior. There is no a priori link to physical movements. (Chalmers someplace says these cases are conceivable.)

cal laws that explains phenomenal variation across worlds. By contrast, for the Russellian monist the problem takes this form: given the vast (infinite) different combinations of micro-phenomenal quiddities, and corresponding macro-experiences, *why do we have exactly the right combination of quiddities to yield harmonious experiences?* Why did we win the “experiential lottery”? This form of the problem is no less serious than the form of the problem for dualists discussed above.

In response, the Russellian monist could say that we just must accept a number of brute facts about what the quiddities happen to be, in addition to the standard brute facts we must accept about what the laws are and what the initial conditions are. According to this response, they not only happen to be *phenomenal* quiddities (so that we are lucky to have experiences and not be Zombies); they also happen to be exactly the “*right*” combination of phenomenal quiddities needed in order to have harmonious experiences. These are just two sorts of brute facts we must accept with “natural piety”. On this view, the fact that we in the actual world have exactly the “*right*” combination of phenomenal quiddities needed in order to have harmonious experiences is a bit like a monkey sitting a typewriter and typing the combination of letters “tomorrow, tomorrow, and tomorrow, creeps in this petty pace from day to day”. Responding in this way is like the dualist claiming that it is just a brute fluke that all the psychophysical laws are harmonious. **And it is objectionable in the same way. It multiplies the number of brute facts we must accept. It complicates our world picture. This goes against the claim Chalmers makes that Russellian monism, like standard physicalism, avoids dualism’s “complexity problem”. It doesn’t: it just requires a different form of complexity and luck than standard dualism.**⁴⁵ Another problem with this response is that intuitively the fact (if it is a fact) that we happen to have the “*right*” combination of phenomenal quiddities needed in order to have harmonious experiences *cries out for explanation*. At this point, the Russellian monist could offer a theistic explanation. But even if the Russellian could justify such a grand inference, he would certainly still have to give up his claim to having a view that combines the simplicity of physicalism with the virtues of dualism!

So far I have assumed that on Russellian monism inharmonious worlds are ones in which our experiences are weirdly different *because of differences in the combinations of micro-phenomenal properties*. However there is a second way in which the Russellian monist might accommodate the possibility of inharmonious worlds. Let me explain his idea, and then address the question of whether helps with my problem of luck for Russellian monists.

On a second version of Russellian monism, having micro-phenomenal quiddities is an *enabling condition* for conscious experience, but precisely what the micro-phenomenal quiddities are makes no difference to the *specific characters* of our experience. Instead, once this enabling condition is in place, it is the *structural facts* S that totally determine the specific character of our experiences. This means that our macro-level experiences are radically *multiply realizable* with respect to the

⁴⁵ Everyone already believes that we are lucky in that the laws and initial conditions didn’t lead to 6 brains circling each other and feeling constant severe pain ad infinitum, and that instead they led to our great lives, Starbucks, and iphones. But I am suggesting that Russellian monists need an *additional* form of luck, adding to the complexity of their view. The reason is that they have another “moving part” in their theory of the world, namely what the quiddities are, on which the existence of character of consciousness depends. (Also, the multiverse hypothesis plus an observation selection effect can’t really explain this form of luck.)

micro phenomenal quiddities. That is, the structural facts S, together with *any* combination of micro phenomenal quiddities (<A, B, C> or <C, D, E> or <G, H, I>) would have realized exactly the same experiences we have in the actual world.

How does this accommodate the possibility of zombie worlds? They are worlds in which the structural facts S are the same but the general enabling condition is not met – worlds in which our micro quiddities are not at all *phenomenal* quiddities, but boring, grey quiddities. How does this view accommodate the possibility of mismatch worlds? The idea is that these are worlds in which (i) S obtains, (ii) we don't have *phenomenal* quiddities needed to ground the occurrence of experiences (instead we have lame non-phenomenal quiddities), and yet (iii) we nevertheless do have experiences, because in this world *Dualism* is true, so that in this world our experiences don't have to be grounded in anything more basic. In particular, they are worlds in which dualism is true, *and* our experiences are out of whack with the neural-behavioral facts. On this view, experiences are non-fundamental (in the sense of being grounded in more basic facts) in our world but fundamental in mismatch worlds. We might call this *variable fundamentality*. And, just to give it a name, we might call this whole package the *multiple realizability* version of Russellian monism, since the initial idea is that our experiences are radically multiple realizable with respect to what micro phenomenal quiddities are.⁴⁶

Notice that this view requires the *primitivist* form of Russellian monism. For if experiential properties were necessarily identical with certain complex properties, as reductionism holds, then they would be grounded in more basic facts *in all worlds*. That is why this view requires a primitivist picture on which at least some experiential properties are not complex properties. So the idea would have to be that in our world these irreducible experiences are grounded in the micro phenomenal quiddities plus the structural facts S, and in mismatch worlds the same irreducible experiences aren't grounded at all (and distributed in weird ways).

For our purposes the key question is whether this "multiple realizability" version of Russellian monism answers the kind luck problem I have developed. I don't see that it does. After all, this view *agrees* with Dualism that an endless variety of mismatch worlds are metaphysically possible, in addition to the harmonious world we occupy. It just accommodates those possibilities in a different way. So on this

⁴⁶ David Chalmers suggested something like "multiple realizability" view as a response to my problem about luck. In fact, Chalmers suggested a slightly different response to my luck problem. On the multiple realizability response discussed in the text, differences in phenomenal quiddities, while holding "structure" fixed, cannot make for *any* phenomenal differences in macro experiences. By contrast, on the response that Chalmers suggested, they can make for phenomenal differences – but only "ones that preserve structure". So, for instance, quiddity switching could result in spectrum inversion, but could not make everything looking grey.

My main objection to this view is similar to one I will develop in the text below to the more radical multiple realizability view, namely, that it is *a priori* odd. If quidditistic differences can result in structure-preserving phenomenal differences, then why not non-structure-preserving phenomenal differences? What explains the restriction? How could there be just a brute restriction? I have an additional objection: the view is just unclear as well as arbitrary. What does "structure preserving mean"? It is supposed to allow for spectrum inversion, but, since color space is asymmetrical in various ways, spectrum inversion is not structure preserving. Also, couldn't sweet-bitter inversion preserve phenomenal differences and similarities, and hence be structure-preserving? In that case, the view does nothing to explain why we are lucky in that we have a combination of quiddities that grounds our actual taste experiences, rather than a different (unlucky) combination of quiddities that grounds sweet-bitter inverted ones. Presumably, Chalmers wants valence to be part of "structure". But what is the general conception of structure at play here?

view we can raise the question: why isn't the *actual* world a mismatch world? I don't see why the "multiple realizable" form of Russellian monism has any advantage when it comes to answering this question. The view must say that this is just a brute, contingent fact, over and above the structural facts that we must all believe in. This adds to the complexity of the view. Another point is that this fact looks lucky and "cries out for explanation". Of course, the proponent of this "multiple realizability" version of Russellian monism might say that it is a kind of objective law that things tend to be simple, and add that mismatch worlds where Dualism is true are necessarily more complex than harmonious worlds like our own. But, given the fact noted above that this view requires a primitivist view of experience in our own world, this claim of comparative simplicity false (see previous section). More to the point, this "law" has no clear content, and so it is hard to believe that there could be such a fundamental law. Even if a clear law along these lines could be formulated, it would add to the complexity of the Russellian monist picture. And of the main points of Russellian monism was to *avoid* the complexity of Dualism – the postulation of brute laws over and above the laws of physics.

Moreover – and this is my main objection - the multiple realizability version of Russellian monism is decidedly odd. If the phenomenal quiddities make all the difference between the *presence* or *absence* of conscious experience, one would expect that they can make a difference to the *specific character* of those conscious experiences, contrary to this view. So we would expect that, if mismatch worlds are possible, then in some of them the reason why our experiences are inharmonious is that the combinations of phenomenal quiddities are different.

So my original "luck problem" stands: *just like dualism*, Russellian monism requires that we accept a slew of brute, contingent facts about our world that look lucky. These facts must just be accepted with "natural piety". It's just that the relevant sorts of facts differ between the views: for the dualist the relevant psychophysical laws turned out to be nice and harmonious, whereas for the Russellian monist the relevant fact is that we have exactly the right combination of micro-phenomenal quiddities to make for harmonious experiences. By contrast, we saw that physicalism avoids the luck problem entirely. So this is a major problem that Russellian monism shares with dualism – and that standard physicalism avoids.

5. Conclusion

Chalmers' Hegelian synthesis argument is that Russellian monism combines the virtues of dualism and physicalism while avoiding their problems. This cannot be said of reductive Russellian monism. This version *combines* the vices of these views. In fact, I think we can say it is false. It also cannot be said of primitivist Russellian monism that it combines the virtues of dualism and physicalism while avoiding their problems: it faces the same problems as dualism. So there is no form of Russellian monism for which the Hegelian synthesis argument is successful.⁴⁷

⁴⁷ The point of this paper is to develop a dilemma for Russellian Monism. Briefly, the dilemma is that this view comes in a "reductive" form and a "primitivist" form. The Russellian monist cannot accept the "reductive" form, while the "primitivist" form is just as *complicated*, and otherwise problematic, as dualism. Philip Goff (MS) develops a theory of experience – indeed, a metaphysics for the whole world – which is vaguely in the same spirit as Russellian monism. He has suggested to me (in discussion) that it avoids the kinds of problems I've developed for Russellian monism. I disagree. I plan to

say more on this elsewhere, but I'd like to briefly explain here why I disagree. Let me start by explaining Goff's view.

In short, Goff accepts Jonathan Schaffer's *cosmos-first* picture of reality, but adds a unique twist. On Schaffer's cosmos-first picture, the whole world is fundamental; everything else is grounded in it. So, at the big bang, the cosmos was F1, where "is F1" is an utterly fundamental predicate expressing the character of the cosmos as it was then. This grounded the existence and character of certain particles and fields which then existed. Then the cosmos started evolving: next it was F2, then F3, then F4, then F4, etc. Its having these evolving fundamental, primitive properties grounded the changing existence and character of the particles and fields and whatnot. When sentient creatures came on the scene, this too was grounded, like everything else, in the fundamental properties of the whole cosmos. Goff accepts this cosmos-first view, but (for reasons I won't go into here) adds some pansychist twists; in particular, he adds that these fundamental world-properties, F1, F2, F3, . . . are *experiences* that the cosmos enjoys. He calls his view *cosmopansychism*.

My main problem with this sort of view is that it is very *complicated* in a way I will now explain. In fact, I think this is a problem for *any* cosmos-first view, including Schaffer's.

The cosmos-first picture requires a *huge, endless swarm of "big-to-small grounding laws"*:

--Necessarily, if the cosmos is F1, then this grounds that there is so and so particles with so and so properties.

--Necessarily, if the cosmos is F2, then this grounds that there such and such particles with such and such different properties.

.

.

.

-----Etc., etc. for every possible world state F_n .

(The proponent of the cosmos-first view might say that these big-to-small grounding laws are explained by a swarm of more basic *essentialist laws*: for instance, it is in the essence of F1ness that if the cosmos is F1, then this grounds that there is so and so particles with so and so properties. But this wouldn't affect the points I want to make, for the sorts of reasons I discussed in this paper at around pp. 36-38.)

Furthermore, I think that the cosmos-first theorist must take all these "big-to-small" grounding laws as brute and must also admit that *each and every one adds to the complexity of our theory of the world*. They must be admitted as further basic principles, beyond we already accept (logical truths, connections between determinates and determinables, etc.). Just as dualists need brute principles that add to theory complexity of our theory of the world ("psychophysical laws"), so too do proponents of the cosmos-first view. This problem I'm pushing here is *not* based on any general Lewisian hostility to "brute necessities" (see e. g. Lewis on the magical conception of complex properties in "Against Structural Universals"); it is rather based on the point, which everyone should accept, that we should try to have a simple-as-possible theory of the world.

To help with this "complexity problem", the proponent of the cosmos-first view might reply that his large-to-small grounding laws (or essentialist laws – see the parenthetical paragraph above) can be derived from a *simple set of more basic truths*.

However, as far as I can see, there is only one way to do this – and it is not at all plausible. I call it the *disjunctivist gambit*. To illustrate, consider the big-to-small grounding law "necessarily, if the cosmos is F15, then there is an electron at place p." On the disjunctivist gambit, this follows from two things. (1) First, an "identification" or "real definition" of the form *what is for there to be an electron at place p just is for the cosmos to be F1, or F2, or F3, . . .*, where F1, F2, F3 are *all the possible world states that ground the existence of an electron at p*. (2) The general principle that, if p, then this grounds the fact that p or q. This "disjunctivist" approach would indeed achieve greater simplicity, because it would show that all big-to-small grounding connections can be derived from (1) identifications (which, as emphasized in this paper, intuitively *don't* add to the complexity of our theory of the world) and (2) the *single principle that disjuncts ground disjunctions*. But such radically disjunctivist real definitions implausible. (Further, this view seems to require *denying* "there are electrons", where "there are" is used in the same strict sense as in "there is a cosmos" – which, *I think*, would rule it out for both Schaffer and Goff.)

Therefore, I think that the proponent of the cosmos-first view is just stuck with *an endless list of big-to-small grounding laws that add to the complexity of his theory of the world*. (By the way, calling

References

- Barnes, E. (2012). Emergence and Fundamentality. *Mind* 121 (484):873-901.
- Benacerraf (1973). Mathematical Truth. *Journal of Philosophy* 70 (19):661-679.
- Bennett, K. MS *Making Things Up*.
- Bennett, K. 2003. Why the exclusion problem seems intractable and how, just maybe, to tract it. *Noûs* 37:471-97.
- Byrne and Hilbert. 2007. Color Primitivism. *Erkenntnis* 66 (1/2): 73 - 105
- Chalmers. 2015. Pansychism and Panprotopsychism. In Alter and Nagasawa (eds) *Consciousness in the Physical World*. Oxford: Oxford University Press.
- Chalmers, D.J. forthcoming. The combination problem for panpsychism. In (G. Bruntrup and L. Jaskolla, eds.) *Panpsychism*. Oxford University Press.
- Chalmers, D. 2014. Strong Necessities and the Mind-Body Problem: A Reply. *Philosophical Studies* (symposium on The Character of Consciousness) 167:785-800.
- Chalmers, D. 2013. The Contents of Consciousness: Reply to Hellie, Peacocke, and Siegel. *Analysis* (symposium on The Character of Consciousness) 73:345-68.
- Chalmers, D.J. 2012. *Constructing the World*. Oxford University Press.
- Chalmers, D. 2006. Strong and Weak Emergence. In (P. Clayton and P. Davies, eds) *The Re-Emergence of Emergence*. Oxford University Press.
- Chalmers, D. 2006b. Perception and the Fall from Eden.
- Chalmers, D. 1996. *The Conscious Mind*.
- Dorr, C. (MS) To be F is to be G.
- Dorr, C and J. Hawthorne. (2013). Naturalness. In K. Bennett and D. Zimmerman *Oxford Studies in Metaphysics*, v. 8.
- Dretske, F. 1995. *Naturalizing the Mind*. MIT Press.
- Field, H. (2001). Postscript to 'Mental representation. In *Truth and the Absence of Fact*. Oxford University Press.
- Field, H. (1978). Mental representation. *Erkenntnis*, 13, 9–61.
- Field, Hartry. 1973. "Theory Change and the Indeterminacy of Reference", *Journal of Philosophy* 70: 462-481.
- Fine, K. 2012. Guide to Ground. In Fabrice Correia & Benjamin Schnieder (eds.), *Metaphysical Grounding*. Cambridge University Press 37- 80.
- Hall, N. 2010. Reduction. <http://plato.stanford.edu/entries/lewis-metaphysics/reduction.html>
- Hawthorne, J. 2006. *Metaphysical Essays*. Oxford
- Hodes, H. 1984. Logicism and the Ontological Commitments of Arithmetic. *Journal of Philosophy* 81 (3):123-149
- Horgan, T. 2010. "Materialism, Emergentism, and the Hard Problem of Consciousness". In Bealer and Koons, *The Waning of Materialism*. Oxford: Oxford University Press.
- Jackson, F. and D. Braddon-Mitchell. 2007. *Philosophy of Mind and Cognition*
- Jackson, F. 1998 *From Metaphysics to Ethics*. Oxford
- Lewis, D. 1994. Reduction of Mind.
- Lockwood, M. (1989). *Mind, Brain, and the Quantum*. Oxford University Press.
- Loewer, Barry. 2001. From physics to physicalism. In Carl Gillett and Barry Loewer, eds.,

the properties *being F1, being F2, being F3*, and so on, "distributional properties", does nothing to change this fact.)

To bring out the complexity of the cosmos-first view, we can compare it the kind of standard reductive physicalist view that is defended (in different versions) by Dorr, Sider, Armstrong, Lewis and many others, which I discussed in this paper. On this view, to explain the origin of the non-fundamental in the fundamental, all we need are identifications or "metaphysical analyses" like *to be in pain is to F* and *to be a city is to G* - where again, intuitively, *these identifications don't add to the complexity of our theory of the world*. This view is clearly much simpler than the cosmos-first view with its huge swarm of brute big-to-small grounding laws. And that is a reason to prefer it to the cosmos-first view.

Because Goff's view is complicated, it is not clear he can use simplicity considerations to support it over other complicated views, like traditional dualism or what I called "primitivist Russellian monism". (In discussion, he agreed with this point, but said that it does have an advantage over dualism, namely, that it avoids the "causal exclusion argument".)

- Physicalism and its Discontents. Cambridge: Cambridge University Press, pp. 37-56.
- McGinn, C. 1996. Another Look at Color. *Journal of Philosophy*.
- Ney, A. 2015. A Physicalist Critique of Russelian monism. In Alter and Nagasawa (eds) *Consciousness in the Physical World*. Oxford: Oxford University Press.
- Papineau (2003). Could There Be a Science of Consciousness? *Philosophical Issues* 13 (1):205-20.
- Pautz (MS) "Rosen on Real Definitions"
- Pautz, A. 2016. How Does Color Experience Represent the World? In Brown and MacPherson (eds.) *Routledge Handbook of the Philosophy of Color*.
- Pautz. 2013. The Real Trouble for Phenomenal Externalists: The Science of Pain, Taste and Smell. In Brown (ed.) *Consciousness Inside and Out*, pp. 237-298. Springer.
- Pautz, A. 2013b "Does Phenomenology Ground Mental Content?" In Kriegel (ed.) *Phenomenal Intentionality: New Essays*, pp. 194-234. Oxford: Oxford University Press.
- Pautz 2010. A Simple View of Consciousness. In Bealer and Koons, *The Waning of Materialism*. Oxford: Oxford University Press.
- Pautz (2006) "Can Physicalists Explain the Structure of Colors?" *Australasian Journal of Philosophy*.
- Prinz, J. 2012. *The Conscious Mind*. Oxford: Oxford University Press.
- Rosen, G. (2015) Real Definition. *Analytic Philosophy*
- Rosen, G. 2010. Metaphysical Dependence.
- Russell, B. 1912. *The Problems of Philosophy*.
- Schaffer (2014). What Not to Multiply Without Necessity. *Australasian Journal of Philosophy* 93 (4):644-664.
- Schaffer, J. 2013. Metaphysical Semantics Meets Multiple Realizability. *Analysis* 73 (4):736-751
- Schiffer, S. *Remnants of Meaning*. MIT Press.
- Sider (2013). Symposium on Writing the Book of the World. *Analysis* 73 (4):751-770.
- Sider, T. 2011. *Writing the Book of the World*. Oxford: Oxford University Press.
- Sider 2001. Criteria of Personal Identity and the Limits of Conceptual Analysis. *Philosophical Perspectives* 15:189-209.
- Stoljar, D. 2001. Two Conceptions of the Physical. *Philosophy and Phenomenological Research* 62 (2):253-81
- Tye, M. 1995. *Ten Problems of Consciousness*. MIT Press.
- Weatherston 2013. The Role of Naturalness in Lewis's Theory of Meaning. *Journal for the History of Analytical Philosophy* 1 (10).
- Weatherston, B. 2003. Vagueness, Parasites and Names. *Australasian Journal of Philosophy* 81: 276-9.
- Wilson, J. (1999). How Superduper Does a Physicalist Supervenience Need to Be? *Philosophical Quarterly* 50 (194):33-52.
- Yablo, S. 2000. Textbook Kripkeanism and the Open Texture of Concepts, *Pacific Philosophical Quarterly*, 81: 98-122.
- Yablo 1992. Mental Causation. *Philosophical Review* 101 (2):245-280.