

## Russellian Monism or Nagelian Monism?\*

Daniel Stoljar, ANU

### Introduction

A natural reaction to the usual set of options in philosophy of mind—viz., dualism and materialism—is to disagree with both. Materialism is certainly an elegant view; the world as the materialist represents it to be is a coherent system founded on a small stock of fundamental physical properties and laws. But materialism seems not to do justice to the phenomena; in particular, it seems not to adequately find a place in the world for consciousness. Dualism by contrast does justice to the phenomena—or anyway does *more* justice than materialism to at least *some* relevant phenomena. But it seems extremely inelegant, for the world as dualism represents it to be is just the material world with consciousness grafted on. But if you take something elegant and graft something onto it, the thing you end up with may well not be elegant; on the contrary, it may well be *inelegant*. And so, it turns out, unfortunately enough, with dualism.

An important task in philosophy of mind, therefore, is to develop and assess positions that are neither materialism nor dualism in the standard sense. My concern in this paper is with two such potential positions: Russellian Monism, and a view that I will call, for a reason that will emerge, Nagelian Monism. As we will see, both Russellian Monism (RM) and Nagelian Monism (NM) have a considerable amount in common. Both deny, as against dualism, that experiential properties are metaphysically fundamental.<sup>1</sup> And both assert, as against standard forms of materialism, that the list of physical properties materialists usually operate with is inadequate—indeed both positions tend to be hyper critical about the usual attempts by both materialists and dualists to spell out what a physical property is. Finally, both advance an epistemic response to the conceivability argument and other arguments against materialism and for dualism—for both the apparent plausibility of these arguments is a symptom of our incomplete grasp of nature.

The difference between them is that RM relies, while NM does not, on a distinction that is in the same general class—to put it vaguely at first—as that between structural and

---

\* I am very much indebted to the following for their help: Torin Alter, David Chalmers, Yujin Nagasawa, Derk Pereboom, and Jon Simon.

<sup>1</sup> Some versions of Russellian Monism might allow that experiential properties of some sort are fundamental, but I will set them aside here. For some discussion of these forms, see Alter and Nagasawa 2012 and Kind, this volume.

non-structural properties, dispositional and non-dispositional properties, relational and non-relational properties, and extrinsic and intrinsic properties.<sup>2</sup> Nagelian Monists needn't reject distinctions in this class altogether; their suggestion is rather that they play no particular role in developing the most plausible position that is neither materialism nor dualism.

Even before we go into the issues in detail, there is at least the following consideration in favour of NM and against RM. Distinctions in the class I just mentioned have proved difficult to control; as Gareth Evans (1980, 102) memorably put it, one's views in this area seem to evince little more than a 'conceptual prejudice'. And if that is the case, why rely on them in your philosophy of mind? For many philosophers attracted to RM, however, this line of thought is mistaken. It may be true—they will say—that other things being equal one should prefer theories that do not rely on controversial distinctions to the ones that do, but that is not the situation here. For when properly understood NM is not an available position in the relevant sense; interpreted one way it is no different from RM, and interpreted another way it is no different from standard materialism.

Why suppose that NM is not an available position in the relevant sense? The argument most commonly advanced to establish this has come to be called 'the structure and dynamics argument', and according to the recent literature on RM, that argument is successful; see, in particular, Alter (2009, 2013), Alter and Nagasawa (2011), Chalmers (2010), McClelland (2013) and Pereboom (2011, 2014). By contrast, in earlier work (Stoljar 2006, 2009), I suggested it is not successful, and that NM is indeed the attractive view it seems to be. The present paper is an elaboration and defence of that point of view; in particular, I will divide the structure and dynamics argument into various versions, and argue that none of them is successful.

### **Russellian Monism and Nagelian Monism**

In order to put the two positions I want to discuss into sharper focus, I will begin with this simple statement of materialism:

---

<sup>2</sup> As Alter and Nagasawa (2012) put it, a distinction like this is "the central distinction" for RM.

D1. Materialism is true at a possible world  $w$  if and only if for every property  $G$  instantiated at  $w$ , there is some physical property (or some complex of physical properties)  $F$  instantiated at  $w$  such that  $F$  (metaphysically) necessitates  $G$ .<sup>3</sup>

Materialism so defined presents us with the impasse already alluded to. On the one hand, the conceivability argument—I will presume familiarity with it and similar arguments here<sup>4</sup>—tells us it is sound that experiential properties—that is, properties constitutively connected to consciousness—are neither physical nor are necessitated by the physical. Hence materialism is false if such properties are instantiated (which I assume they are). On the other hand, the usual alternative to materialism—the sort of dualism that regards experiential properties as metaphysically fundamental—faces widely discussed problems having to do with the overall elegance of the view and, related to this, the causal and explanatory role of such properties.

The Russellian Monist offers a way out of this impasse. A good way to see the shape of this view is to start from a distinction between two kinds of physical property, which we may call ‘standard’ physical properties and ‘non-standard’ physical properties. For the Russellian, this distinction is in turn explicated in terms of a distinction in the class mentioned above, i.e. between structural and non-structural, dispositional and non-dispositional, and so on. Suppose, to fix ideas, we focus on the first of these, the structural/non-structural distinction. Then the suggestion of the Russellian Monist is that standard physical properties are structural properties, and, correlatively, that non-standard physical properties are non-structural properties. The position further assumes, perhaps controversially, (a) that there is a good sense in which physics, and empirical inquiry generally, tells us only about structural properties and (b) that while this is so we nevertheless have good reason to suppose that non-structural properties are instantiated as well as structural properties.<sup>5</sup>

Now once we have the distinction between standard and non-standard physical properties before us, we can likewise draw a distinction between two kinds of materialism, standard materialism and non-standard materialism. Standard materialism is the view you get when you combine D1 with an assumption that the physical properties in question are

---

<sup>3</sup> For statements of materialism (aka physicalism) of this sort, see Stoljar 2010. The definition used in the text is a simple one; for example, it takes for granted exactly what metaphysical necessitation is. I will set aside such complications here. The definition is also simple (or better, incomplete) in that it does not say what a physical property is; I will take up that problem below.

<sup>4</sup> See, e.g., Stoljar 2014, Chalmers 2010 and Pereboom 2011

<sup>5</sup> In Stoljar 2014 I discuss four different versions of RM; the version in the text appears there as RM3.

standard, and so structural, properties. Non-standard materialism combines D1 with the assumption that physical properties are either standard or non-standard.

How do these distinctions show us a way out of the impasse? The suggestion of RM is that, while the conceivability argument is persuasive against standard materialism it is not persuasive against non-standard materialism. For according to non-standard materialism there is a class of relevant instantiated physical properties (i.e. the non-structural properties) of which we are ignorant; if so, the conceivability argument loses its force since it presupposes that we know at least in outline what the relevant physical properties are.<sup>6</sup> On the other hand, to give up standard materialism is not to adopt dualism, and so the fundamentality of experiential properties may be denied.

So RM offers a way out of the impasse presented by materialism—what then of NM? Well, NM offers a way out too, and in fact a directly analogous way. Like RM, NM starts from a distinction between standard physical properties and non-standard physical properties, though in this case the distinction is explicated differently. For the Nagelian, standard physical properties are properties of the sort described in contemporary physics (and contemporary empirical inquiry generally). Correlatively, non-standard physical properties are properties of the sort described in the physics, whatever it is, that one will or might formulate at the ideal limit of inquiry. The position assumes, perhaps controversially, that contemporary physics is incomplete in the story it tells about nature, and so there is a difference in kind between contemporary physics and the physics that we will (or might) ideally reach.<sup>7</sup>

Once we have this distinction before us, NM, exactly like RM, draws a distinction between two kinds of materialism, standard and non-standard materialism. How do these distinctions show us a way out of the impasse? The suggestion of NM—again: exactly like RM—is that while the conceivability argument is persuasive against standard materialism it is not persuasive against non-standard materialism. For according to non-standard materialism

---

<sup>6</sup> I will assume here familiarity with this epistemic style of response to the conceivability argument, which I have set out in detail elsewhere; see Stoljar 2006, 2014.

<sup>7</sup> In *The View from Nowhere*, Nagel writes: “The difference between the mental and the physical is far greater than the difference between the electrical and the mechanical. We need entirely new intellectual tools, and it is precisely by reflection on what appears impossible—like the generation of mind out of the recombination of matter—that we will be forced to create such tools. It may be that the eventual result of such exploration will be a new unity that is not reductionist. We and all other creatures with minds seem to be composed of the same materials as everything else in the universe. So any fundamental discoveries we make about how it is that we have mind and what they actually are, will reveal something fundamental about the constituents of the universe as a whole. (1986, 52–53)” It is because of its affinity with views like this that I will call the view described in the text Nagelian Monism.

there is a class of relevant instantiated physical properties (i.e. the properties described at the ideal limit) of which we are ignorant. On the other hand, to give up standard materialism is not to adopt dualism, and so the fundamentality of experiential properties may be denied.

So we seem to have two ways out of the impasse over materialism.<sup>8</sup> But as we noted before, if the structure and dynamics argument is sound the only available route here *in fact* is the Russellian one. But what exactly is that argument? I now turn to that question.

### **The Structure and Dynamics Argument**

The key text for the structure and dynamics argument is this well-known passage from David Chalmers:

First, physical descriptions of the world characterize the world in terms of structure and dynamics. Second, from truths about structure and dynamics, one can deduce only further truths about structure and dynamics. Third, truths about consciousness are not truths about structure and dynamics. (Chalmers 2002, 258.)

The argument presented in this passage is I think fairly summarized as having the following general form:

P1. Every physical truth is a truth of a certain kind, i.e., one that “characterizes the world in terms of structure and dynamics”.

P2. For every truth  $T$  of that kind, if  $T$  a priori entails<sup>9</sup> a truth  $T^*$ , then  $T^*$  is of that kind too.<sup>10</sup>

P3. No truth about consciousness is a truth of that kind.

---

<sup>8</sup> Reader, beware: in Stoljar 2014, I classified (what I am here calling NM) as a version of RM, viz., RM4. As I say there however, that it is not clear that RM4 *is* a version of RM strictly so called, because it does not appeal to the distinction between structural and non-structural properties or any similar distinction.

<sup>9</sup> I will assume here that for  $T$  to a priori entail  $T^*$  is for the material conditional ‘If  $T$  then  $T^*$ ’ to be both necessary and knowable a priori.

<sup>10</sup> There are a number of problems with P2 as stated. For example suppose  $T$  is a truth of a certain kind, and  $T^*$  is a disjunction of  $T$  and a truth of some other kind. In that case  $T$  will a priori entail  $T^*$  and yet  $T^*$  is arguably not of the relevant kind. However, since the problems with the argument I want to raise focus on P1 and P3 I will set this aside.

On the face of it, these claims present a major problem for NM. As we have seen, the Nagelian Monist supposes that the conceivability argument is persuasive against standard materialism but is not persuasive against non-standard materialism, where in turn this distinction is founded on a distinction between the properties described in contemporary physics and the properties described in ideal physics. But if each of P1-P3 is true, this distinction makes no difference to the impasse over materialism. For according to P1, both versions of materialism described by NM entail that truths about consciousness are necessitated by truths about structure and dynamics; and, according to P2 and P3, any theory of that kind lies in the target range of the conceivability argument. Hence, while we may for other reasons find non-standard materialism more plausible than standard materialism, doing so leaves the issues surrounding the conceivability argument unaffected.<sup>11</sup>

Nagelian Monists might respond by denying that both standard and non-standard materialism entail that truths about consciousness are necessitated by truths about structure and dynamics; in particular, they might say, non-standard materialism permits the idea that truths about consciousness are necessitated in part by truths about structure and dynamics and in part by truths not about structure and dynamics. However, while this avoids the problem, it abandons what is distinctive about NM as opposed to RM. For RM precisely holds that truths about consciousness are necessitated in part by truths that are not structural and dynamical. Hence we arrive at the conclusion of the structure and dynamics argument: either NM is no help as far as the conceivability argument goes (and so is no better than standard versions of materialism) or else it collapses into RM.

Is this argument persuasive? In my view the crucial issue in deciding whether it is concerns what exactly “characterizes the world in terms of structure and dynamics” means. (In what follows I will concentrate on what ‘structure’ means, since ‘dynamics’ refers to how structure changes over time.) What then is it for a truth to characterize the world in terms of structure; what is it (as I will say) for a truth to be a structural truth? Since there are different answers to this question, there are different versions of the structure and dynamics argument.

---

<sup>11</sup> The argument in the text is in fact a bit stronger than is dialectically necessary. One way in which it might be weakened is that quantifier ‘no truth’ in P3 can be replaced with ‘not every truth’; I will mention this at several points below. A different way is that a proponent of the argument may wish to weaken the premise even further and say, not that no structural and dynamical truth is a truth about consciousness, nor even that not every such truth is, but rather that the conceivability argument (however that argument turns out) concerns the relationship between truths about consciousness and structural truths, and hence the difference between standard materialism and non-standard materialism (as understood by NM) makes no difference to the argument. (I discuss this second way to weaken the argument in Stoljar 2006, 2009).

As I have indicated, my strategy will be to distinguish three different versions of the argument, and argue that each is unpersuasive.

### **The First Version of the Argument**

According to the first and simplest suggestion, a truth is a structural truth if it concerns a relation; that is, if it concerns the instantiation of an  $n$ -place property where  $n > 1$ . Plugging this idea into the template above we arrive at these premises:

- P1.a.            Every physical truth is a truth of a certain kind, i.e., one that concerns relations.
- P2.a            For every truth  $T$  of that kind, if  $T$  a priori entails a truth  $T^*$ , then  $T^*$  is of that kind too.
- P3.a            No truth about consciousness is a truth of that kind.

As I have argued elsewhere (Stoljar 2006, 2009), however, in this form the argument is not plausible. First, as against P1.a, there is no reason to suppose that physics cannot tell us about properties that are not relations. Second, as against P2.a, it is false that from truths about relations only truths about relations follow—e.g. from the truths about the relations between points in space one can derive a truth about the shape of the region these points constitute. Third, as against P3.a, some truths about consciousness are themselves relational.

Now a natural response to these points is to abandon the first version in favour of another; the bulk of the discussion to follow will be concerned with suggestions of this type. But before turning to them, it is worth mentioning a different reply suggested at one point by Alter (2009, see also Alter and Nagasawa 2012).<sup>12</sup> According to this reply, “truths about consciousness are not truths solely about structure and dynamics” (Alter 2009, 764). On the assumption that ‘structure and dynamics’ is here interpreted along the lines of the first version of the argument we are considering, Alter’s suggestion is that truths about consciousness are not truths solely about relations.

Actually, there are two ways to take this suggestion. First, in saying that truths about consciousness are not solely relational, Alter might be making the point that truths about consciousness are not *all* relational, i.e. some are relational and some are not. Now this point

---

<sup>12</sup> Alter and Nagasawa 2012 develop this point by appealing to a notion introduced by Pereboom 2011, viz, that of a purely extrinsic property. I express some doubts about this notion below.

is indeed of some interest, for it brings out that a version of the structure and dynamics argument may go through even if P3.a—according to which *no* truth about consciousness is a relational truth—is replaced with the weaker premise that *not every* truth about consciousness is a relational truth. Combining this weaker premise with P1.a and P2.a yields the result that some truths about consciousness are not physical, and that might well be enough to generate a problem for NM. However, even if this is so, pointing it out does not help this first version of the argument. For while it may be that not every truth about consciousness is a relational truth, it is likewise the case that not every physical truth is a relational truth; indeed this is precisely one of the problems with this version of the argument.

Second, in saying that truths about consciousness are not solely relational, Alter might be making the point that such truths, even when they are relational, are not only about relations. For example, take the truth that this pain is more intense than that pain. That is a relational truth, but it is not only about a relation, i.e. because it is also about the intensity of pain. Once again, however, while this point is correct it is hard to see how it helps with the first version of the structure and dynamics argument. For if P3.a is interpreted that way, then P1.a must likewise be interpreted as saying that every physical truth is solely relational. But that is not so. Take any relational physical truth, for example, that the earth is heavier than the moon. That is a truth about a relation, but it is also a truth about the physical world. (It is not about pain, for example, or about biology.) So physical truths aren't solely relational either. So there is no way to defend this first version of the argument by appealing to the idea that truths that are not solely relational.

### **The Second Version of the Argument**

While the first version fails, Derk Pereboom (2011, 2014) has suggested a second version, and his suggestion has become widely adopted (see, e.g. Alter 2014, Alter and Nagasawa 2012, McClelland 2013).

Pereboom's key idea is to draw a distinction between two sorts of intrinsic property, a comparatively intrinsic property and an absolutely intrinsic property. To illustrate the idea of a comparative intrinsic property, consider the shape of my brain, that is, the organ currently located in my skull. That my brain has this shape is an intrinsic property of it, at least on most understandings of what an intrinsic property is. Moreover, that my brain has this shape is plausibly a function of the parts of my brain having certain properties and standing in certain relations. Now suppose something that may or may not be true, viz., that my brain has the shape that has, not because its parts instantiate various *intrinsic* properties, but because



they instantiate various *extrinsic* properties. That does not conflict with shape of my brain being an intrinsic property of it, but it would make it (in Pereboom's sense) a comparatively intrinsic property.

On the assumption that the notions of a comparatively intrinsic property and an absolutely intrinsic property may be inter-defined, we might capture the relevant notions like this:

- D2.  $F$  is a comparatively intrinsic property of  $x$  if and only if (a)  $F$  is an intrinsic property of  $x$ ; and (b) there are extrinsic properties  $G_1, G_2 \dots G_n$  such that (i) the (proper) parts of  $x$  exhibit a pattern of instantiation of these properties and (ii) necessarily, if the parts of  $x$  exhibit that pattern of instantiation, then  $x$  has  $F$ .
- D3.  $F$  is an absolutely intrinsic property of  $x$  if and only if (a)  $F$  is an intrinsic property of  $x$ ; and (b)  $F$  is not a comparatively intrinsic property of  $x$ .

This suggestion about how to understand the comparative/absolute distinction is close but not quite the same as that found in Pereboom 2014 which is in turn close but not quite the same as that found in Pereboom 2011. In his earlier discussion Pereboom did not focus squarely on necessity, and in both discussions he appeals to the idea of purely extrinsic property, which is an extrinsic property that has no intrinsic aspects or components (2014, 50). I find the latter notion obscure. For example, take the property of being one among many, which is Pereboom's main example of a purely extrinsic property. Doesn't this property have the property of being numerical, and isn't this an intrinsic aspect (i.e. an intrinsic property) of that property? Indeed, take any extrinsic property at all; does it not have the intrinsic property of *being extrinsic*? If so, it is not a purely extrinsic property, and indeed no property is a purely extrinsic property in Pereboom's sense. For these reasons, it is unclear what a purely extrinsic property is. Fortunately, however, it will not be important for our purposes here to settle this matter. In drawing our attention to the comparative/absolute distinction, Pereboom has picked out something correct and important; precisely characterizing what he has picked out may be set aside.

But how does this distinction help with the structure and dynamics argument? Well we now have a further way to understand the notion of a structural truth. On this view, a structural truth is one that concerns either extrinsic properties or comparatively intrinsic properties; equivalently, it is a truth that does not concern absolutely intrinsic properties. Plugging this into the template, we arrive at this second version of the argument:

- P1.b Every physical truth is a truth of a certain kind, i.e., one that concerns either extrinsic properties or comparatively intrinsic properties.
- P2.b For every truth *T* of that kind, if *T* a priori entails a truth *T\**, then *T\** is of that kind too.
- P3.b No truth about consciousness is a truth of that kind.

Once again, these premises entail that the difference between standard and non-standard materialism, as understood by NM, does not matter as far as the conceivability argument is concerned.

This second version of the argument is better than the first. First, that physics can tell us about properties that are not relations does not threaten P1.b. It is consistent with P1.b that physics tells about intrinsic properties and if these are non-relational properties, then physics tells about them. Second, pointing out that a thing can have a non-relational property because its parts stand in various relations does not threaten P2.b. It is consistent with P2.b that this happens.

Nevertheless, a serious problem for this argument emerges when we focus on why P3.b is true<sup>13</sup>. On the face of it many truths about consciousness concern both relations and extrinsic properties; indeed this point was one of our criticisms of the first version of the argument above. Suppose for example I am touching a piece of velvet. This involves a relation between the velvet on the one hand and me on the other; and yet that I am touching velvet is a truth about consciousness. Or consider the idea held by many philosophers of mind, viz., that to be in a conscious state is to bear a relation to something, e.g. a relation of awareness or acquaintance either to an individual or a property or proposition. Any position of this kind entails that there are truths about consciousness that are truths about relations.

Now a proponent of the second version of the argument can revise P3.b to avoid some of these problems. One might for example draw a clear distinction between intrinsic properties and relational properties, and suggest that while truths about consciousness are truths about relations they are also truths about intrinsic properties; e.g. being acquainted with a property might be a relational feature of a person as well as an intrinsic feature.<sup>14</sup> One might

---

<sup>13</sup> There is a second problem for Pereboom's argument as well, viz. why and in what sense physics does not tell us about absolutely intrinsic properties. I will set this problem aside here, however.

<sup>14</sup> The literature on intrinsic properties standardly draws a distinction here; see, e.g., Weatherston and Marshall 2013.

also say that what is important here is not that *all* truths about consciousness concern absolutely intrinsic properties, but merely that *some* do (cf. Pereboom 2014); indeed we noted the possibility this move when discussing Alter’s response to the first version of the argument above.

But even while these manoeuvres are available, it remains unclear that P3.b is true. For suppose that I feel as if I am touching piece of velvet; and suppose that this property—feeling as if I am touching velvet—is an intrinsic property. Is it an *absolutely* intrinsic property? Surely that is an open question; it *might* be an absolutely intrinsic property, but then again it *might* not be. Indeed, this is true for many paradigmatic intrinsic properties. Suppose my brain has some particular shape *S*; and suppose that this property is an intrinsic property of my brain. Is it an absolutely intrinsic property? Once again, that is surely an open question. But the problem with P3.b is that, according to it, it is not an open question whether feeling as if one is touching velvet is absolutely intrinsic. On the contrary, the premise says directly that it is an absolutely intrinsic property; and that that is why it should be rejected.

Pereboom has responded to an earlier and briefer (see Stoljar 2014) presentation of this point by offering two arguments that properties such as feeling as if one is touching velvet are absolutely intrinsic (see Pereboom 2014). The first argument concerns the notion of primitive property, where “a primitive property is (i) one whose entire qualitative nature or essence is revealed in our sensory or introspective representation of it, and thus is not identical to a property with a qualitative nature distinct from what is revealed by the representation, and (ii) one that is metaphysically simple and thus not constituted by a plurality of other properties” (2014, 59). On the basis of this, Pereboom argues as follows:<sup>15</sup>

- (1) Experiential properties are primitive
- (2) Experiential properties are intrinsic
- (3) If Experiential properties are intrinsic and primitive, then they are absolutely intrinsic
- (4) Ergo, experiential properties are absolutely intrinsic.

Here, (3) is intended to be a logical truth that follows from the definitions of the relevant notions; and (1) and (2) are intended to be truths which we can come to know or at least

---

<sup>15</sup> While this argument is set out in his 2014, Pereboom has since indicated (p.c) that it does not represent his considered position, in part for the reasons I go on to indicate. Since it is somewhat natural argument however, I will consider it in the text in any case.

believe on the basis of introspection.<sup>16</sup> Since (1-3) entail (4), we seem to have a good argument that experiential properties are absolutely intrinsic.

But this argument is unpersuasive. First, regardless of its soundness, it is not an argument that the Russellian may make against the Nagelian, and so is of limited value in the present discussion whose focus is on the contrast between these views. The reason is that (1) is inconsistent with RM, at least in the version we have been considering. For RM in that version holds as against the dualist that experiential properties are not fundamental and so are not primitive in Pereboom's sense.

Second, it is far from clear that (1) is true or is something that we are liable to believe or know on the basis of introspection, as Pereboom suggests. When I feel as if I am touching velvet, I can certainly come to know or believe certain things on the basis of introspection. For example, I can come to know or believe that I feel as if I am touching velvet. Perhaps too I can come to know (in a variety of senses) the way one feels when one feels velvet. But to say that I can come to know on the basis of introspection that feeling velvet is a primitive property seems to me extremely implausible. To adapt a phrase David Lewis used in a related context, "making discoveries in metaphysics is not so easy"—that is, you can't tell just by introspection that some property is or is not determined by some other property.<sup>17</sup> At least, the Nagelian Monist is perfectly entitled to deny this, and so this argument is unpersuasive.

Pereboom's second argument in favour of the view that experiential properties are absolutely intrinsic is an interesting variation on the original conceivability argument against materialism.<sup>18</sup> Suppose T is any truth at all about extrinsic properties and comparatively

---

<sup>16</sup> In the text I use the phrase 'on the basis of introspection' with the intention of abstracting away from various controversies in the philosophy of introspection that are lying just below the surface here. For example, in Pereboom's 2011 discussion of introspection we find a distinction between non-belief-like introspective representations on the one hand and introspective beliefs (i.e. beliefs occasioned in various ways by introspective representations) on the other. From this point of view, to say that S knows or believes p on the basis of introspection might be read as entailing that S has a non-belief-like introspective representation according to which p. I don't intend the phrase that way here, for two reasons. First, while he accepts the existence of non-belief like introspective representations, Pereboom himself does not suppose that we have introspective representations according to which (e.g.) (2) is true; his suggestion is rather that belief in (2) is occasioned by introspective representations of other sorts. Second, it is controversial whether such non-belief like introspective representations exist in the first place, since to suppose so is to treat introspection as deeply analogous to perception in ways it may not be.

<sup>17</sup> What Lewis in fact said is "making discoveries in neurophysiology is not so easy" (1995, 329). In effect, this is the argument from revelation against materialism as discussed in Lewis 1995 and Stoljar 2009.

<sup>18</sup> Here is Pereboom's (2014, 59) way of stating this argument: "...if we let PRP be any epistemically possible description of the world that features only purely relational properties, the intuition that 'PRP and ~O' is ideally, primarily and positively conceivable will be very strong. If from this we can conclude that the phenomenal truths are not necessitated by or derivable from the purely relational truths, we can also conclude that the phenomenal truths are not truths exclusively about purely relational and merely comparatively intrinsic

intrinsic properties, and T\* is any truth about consciousness—e.g., the truth that I feel as if I am touching velvet. The premise of Pereboom's second argument is that it is conceivable that T is true and T\* is false, in just the sense of 'conceivable' that is at issue in the original conceivability argument. The conclusion is that T\* is not a truth about extrinsic or comparatively intrinsic properties, and so is a truth about (a truth which attributes) an absolutely intrinsic property.

But this argument is unpersuasive too. As we have seen, it is an open question whether the experiential property of feeling velvet is an absolutely intrinsic property or not; hence it is an open question whether T\*, the truth that I am feeling velvet, attributes an absolutely intrinsic property or not. And this means that the conjunction Pereboom has in mind is not conceivable in the sense he intends. For consider: if T\* concerns an absolutely intrinsic property (which it might do), then Pereboom's conjunction will be conceivable, but if it does not (which it might not) Pereboom's conjunction is not conceivable. Either way, one cannot rely on this argument to establish that this truth about consciousness concerns an absolutely intrinsic property.

In sum, while the second version of the structure and dynamics argument offered by Pereboom is an improvement on the first, it is only successful if we have good reason to suppose that its third premise (P3.b) is true. But Pereboom has provided no reason to suppose that P3.b is true. Hence this version of the argument should be rejected.

### **The Third Version of the Argument**

The two versions of the structure and dynamics argument we have considered so far have in common the view that a structural truth is a truth that *concerns* or is *about* properties of a certain sort. The first version says that a structural truth is one that concerns properties that are not one-place properties; the second version says that a structural truth is one that concerns properties that are either extrinsic or comparatively intrinsic.

Now if one looks at the passage from Chalmers 2002 quoted above it would be very natural to assume that this was in fact how the argument should be understood; that is, it would be very natural to assume that structural truths are truths with a certain metaphysical subject matter. But more recently, in *The Character of Consciousness*, Chalmers has corrected this interpretation. A structural truth, he says, is not a truth that *concerns* a property, relational or not, absolutely intrinsic or not; rather it is a truth that can be *formulated* in a

---

properties. And this in turn would entail that the phenomenal truths are at least partly about absolutely intrinsic properties and/or absolutely intrinsic aspects of properties"

certain vocabulary: “formally, a structural-dynamic description is one that is equivalent to a Ramsey sentence whose O-terms are limited to spatiotemporal expressions, nomic expressions, and mathematical and logical expressions. (2010, 210n18).”<sup>19</sup>

Plugging this into the template we arrive at this third version of the structure and dynamics argument:

- P1.c Every physical truth is a truth of a certain kind, i.e., one a priori<sup>20</sup> equivalent to a Ramsey sentence whose o-terms are limited to (a) spatiotemporal expressions; (b) nomic expressions; (c) mathematical and logical expressions.
- P2.c For every truth  $T$  of that kind, if  $T$  a priori entails a truth  $T^*$ , then  $T^*$  is of that kind too.
- P3.c No truth about consciousness is a truth of that kind.

Once, again, if these premises are true then the difference between standard and non-standard materialism, as understood by NM, does not matter as far as the conceivability argument goes.

Now this version of the argument is clearly different from the previous ones, and as a consequence avoids their problems. In particular, the fact that a truth can be formulated in a certain vocabulary does not tell us anything about the metaphysical character of the properties or relations the truth is about. For suppose there is an expression in the class Chalmers describes which has as its semantic value an intrinsic property (or an absolutely intrinsic property). That would not by itself stop a sentence containing that expression from being structural in the sense he has in mind. Hence a truth can be structural in Chalmers’ sense and not be structural in (e.g.) Pereboom’s sense, and vice versa.

But how plausible is this version of the argument? I want to address this question by focusing on what spatiotemporal expressions are. It is clear that the argument presupposes that spatiotemporal expressions are a certain class of expression, and so to assess the argument we need to know what determines membership of that class. Now in *Constructing the World*, Chalmers offers an answer to this question which, following him, I will call

---

<sup>19</sup> The reference to ‘O-terms’ and ‘Ramsey-sentence’, Chalmers is alluding to the Ramsey-Carnap-Lewis account of theoretical terms. For relevant presentation of these ideas see Lewis 2009, and Chalmers 2012. I also connect them to Russellian Monism in Stoljar 2014.

<sup>20</sup> In the passage quoted, Chalmers does not specify the sense of ‘equivalent’. I will assume however that what is intended in ‘a priori equivalent’, where for any sentence  $S$  and  $S^*$ ,  $S$  is a priori equivalent to  $S^*$  just in case the bi-conditional ‘ $S$  is true iff  $S^*$  is true’ is knowable a priori.

‘phenomenal functionalism’. On this view, spatial expressions “function to pick out that manifold of properties that serves as the normal causal basis of a corresponding manifold of properties in our spatial experience” (2012, 335). On the assumption that this applies to *spatiotemporal* expressions and not simply to *spatial* expressions, spatiotemporal functionalism says that a spatiotemporal expression is one that functions to pick out the manifold of properties that serves as the normal causal basis of a corresponding manifold of properties in our spatiotemporal experience.

However, when we combine this suggestion with the third version of the structure and dynamics argument just outlined we confront a major problem, viz. that if phenomenal functionalism is true, the structure and dynamics argument is unpersuasive. There are two different ways to bring out this fundamental point.<sup>21</sup>

The first way starts by bringing to the surface a presupposition that is held by at least some proponents of the structure and dynamics argument, viz., that whatever physical or structural truths are exactly, no such truth a priori entails any truth about consciousness. If we do assume this, and if we further assume that phenomenal functionalism is true, it immediately follows that P1.c is false. For P1.c says that physical truths are equivalent to various truths formulated in a certain vocabulary including spatiotemporal vocabulary; and phenomenal functionalism says that truths formulated in spatiotemporal vocabulary are a priori equivalent to (or a priori entail) truths about consciousness. Putting this together, P1.c says that physical truths are a priori equivalent to, or a priori entail, truths about consciousness. But if that is so, and if no physical truth a priori entails any truth about consciousness, P1.c is false.

The second way focuses instead on P3.c, the claim that no truth of the relevant kind is a truth about consciousness. If phenomenal functionalism is true, this claim is false too. For, given phenomenal functionalism, some truths about consciousness will indeed be structural truths on the present understanding, or at any rate will be a priori equivalent to such truths. To illustrate, take the truth ‘there are spatiotemporal experiences that have causes’, which is a truth about consciousness. This truth entails ‘there are things which cause spatiotemporal experiences’, and this is a structural truth *if* we assume phenomenal functionalism. Similarly,

---

<sup>21</sup> Indeed the point here is of considerable interest when placed in the larger context of Chalmers’ work. For the Chalmers of *The Character of Consciousness* says that the structure and dynamics argument is sound and should be interpreted in a particular way, in particular as involving the notion of a spatiotemporal expression. And the Chalmers of *Constructing the World* tells us that phenomenal functionalism is the best account of what a spatiotemporal expression is. The problem is that the first Chalmers seems to be on a collision course with the second!

take the truth ‘there are things which cause spatiotemporal experiences’ which is a structural truth. This truth entails ‘there are spatiotemporal experiences that have causes’, and this is a truth about consciousness.

How might one respond to the point that if phenomenal functionalism is true, the structure and dynamics argument is unpersuasive? There are four main objections to this; I will go through them one by one:<sup>22</sup>

*Response 1:* “Phenomenal functionalism tells us *at most* that some truths about consciousness are spatiotemporal truths, viz., those about spatiotemporal experience. However, aren’t there also *other* truths about consciousness, i.e., those not about spatiotemporal experience? If so, the argument will still go through for the by now familiar reason, viz., that the third premise may be weakened so that it says only that not every truth about consciousness is structural.”

*Reply:* This is true in principle but in practice it would require separating out experiences that are spatiotemporal from experiences that are not, and this is very difficult if not impossible to do. Any truth about consciousness will concern experiences of some sort, and any sort of experience plausibly includes spatiotemporal experience as a part. Take for example perceptual experiences; experiences of this sort seem to represent objects as having various spatial and temporal properties and so are themselves partly spatiotemporal.

*Response 2:* “Phenomenal functionalism is only one sort of functionalism. In *Constructing the World*, Chalmers also describes a position called ‘non-phenomenal functionalism’, which is the same as phenomenal functionalism except that “the roles will not involve connections to our spatial experiences, but rather will be structural roles within the physical realm” (2012, 332). If we adopt a non-phenomenal functionalism, the problem is avoided.”

*Reply:* The phrase “structural roles within the physical realm” is not very clear, especially in a context in which it is contested just what ‘structural’ and ‘physical’ are supposed to mean. But as I understand him, what Chalmers intends by non-phenomenal functionalism is a position exactly like phenomenal functionalism except that spatiotemporal expressions are *themselves* defined in terms of mathematical/logical and causal/nomic expressions. So on this interpretation, P1.c says that every physical truth is one that is a priori equivalent to a truth formulable using only nomic/mathematical and causal/nomological vocabulary.

---

<sup>22</sup> I am indebted to discussion with David Chalmers in formulating these responses.



But the problem now is that this is a very implausible account of what physical truths are. To illustrate this, consider the temporal phase of the actual world from its beginning up to the point at which life (and therefore consciousness in the ordinary sense) evolved—call it *Pre-life*—and suppose I am omniscient with respect to the structural truths (in the sense at issue) that obtain there. In effect, this means that I know a huge body of truths about Pre-life of the following type:

- (i) There is an object, a, and a property F, such that a instantiates F;
- (ii) There is a pair of objects a and b, and a pair of properties F and G such that a instantiates F, and b instantiates G;
- (iii) There is a pair of objects a and b, and a pair of properties F and G such that a instantiates F, and b instantiates G, and F and G are lawfully correlated.

Now if physical truths are structural truths in this sense, structural omniscience with respect to Pre-life is also physical omniscience with respect to it; that is I would know not merely all the structural truths that obtain in this part of the world but all the physical truths too. But that is clearly not so. For consider some truth V—as it might be, the truth that the volcano at this place erupted at this time. On the face of it, no matter how many truths I know that are similar to those in (i-iii), I will not know V. The reason is that truths of the (i-iii) type tell me only that there are certain objects that have certain properties that stand in certain lawful relations; they do not tell me what *kind* of objects, properties and lawful relations these are. So, for example, even if I know that there is some object that has some property, I will not know that the object is a volcano and that the property is the property of erupting.<sup>23</sup> But if that is true, P1.c should be rejected because it offers an implausible account of what a physical truth is. More generally, just as phenomenal functionalism is inconsistent with the persuasiveness of the structure and dynamics argument, so too is non-phenomenal functionalism; if either view is true, the argument is should be rejected.

*Response 3.* “Functionalism in either its phenomenal or non-phenomenal form is only one type of account of what spatiotemporal expressions are. Another account discussed in *Constructing the World* is the primitivist account that treats these expressions not as a priori equivalent to definite descriptions of a certain form but as instead as names for a certain class

---

<sup>23</sup> The point in the text is closely related to a problem mentioned by Chalmers in his discussion of ‘nomic structuralism’, viz., that it is “open to multiple realization, and that it will leave certain truths about the nature of the parameter unsettled” (2013, 412).

of properties and relations. If we adopt primitivism, then the conflict between the phenomenal functionalism and the structure and dynamics argument is avoided.”

*Reply:* The first point to make about primitivism is that, unlike functionalism of either sort, it does not by itself answer the question we started with, viz., what sort of expression is a spatiotemporal expression? Hence, if one appeals to primitivism, one would need to elaborate it to answer that question; as I said above spatiotemporal expressions are a type of expression and the persuasiveness of the argument depends in part on how we understand that type. I see two main possibilities here.

On the first—which I take to be Chalmers’ own view— spatiotemporal expressions are those that denote a certain class of property and relation, namely, the ones that obtain in what Chalmers calls Eden, the mythical place in which, as he puts it (2010) “we had unmediated contact with the world. We were directly acquainted with objects in the world and with their properties. Objects were simply presented to us without causal mediation, and properties were revealed to us in their true intrinsic glory.” But the problem with this is that properties instantiated in Eden, including spatiotemporal properties, are not in fact instantiated, or so Chalmers plausibly argues. And if this is the case, introducing primitive expressions of this sort makes no difference to the nature of the physical truths mentioned in P1.c, i.e., for no physical truth will attribute any such property. Hence appealing to primitivism of this variety is no advance on the suggestions already considered.

On the second, a spatiotemporal expression is any expression that one might introduce in the course of explaining the scientific counterparts of Edenic properties. Since the scientific counterparts of Edenic properties are different from the Edenic properties themselves, and are presumably instantiated, this suggestion avoids the problem just noted. But the issue now is that, understood this way, the class of the primitive spatiotemporal expressions is extremely open ended, i.e. since it depends on what exactly those scientific counterparts are and what their explanation turns out to be. Given that it is open ended, it is in turn unclear that P3.c is true. More generally, appealing to conceptual primitivism does not alter the underlying situation.

*Response 4.* “Suppose phenomenal functionalism is indeed in conflict with the structure and dynamics argument. Even so, this is merely a pyrrhic victory for the Nagelian Monist, since that position cannot appeal to phenomenal functionalism in any case. What phenomenal functionalism does in effect is treat truths about spatiotemporal experience, and experience more generally, as conceptually primitive truths of the world. That does not immediately entail that experiential truths are metaphysically fundamental but it comes very

close to doing so, and the path from a truth's being conceptually primitive to it being metaphysically fundamental is not something the Nagelian is liable to block. But Nagelian Monism is inconsistent with the truths about consciousness being fundamental.”

*Reply.* It is true that Nagelian Monism sits uneasily with phenomenal functionalism. Indeed, this is not only for the reason given but for the further reason that the program that phenomenal functionalism is central to—the program of *Constructing the World*—involves a kind of rationalism according to which any truth at all follows a priori from some truth of a kind currently known to us, or at least currently known to some idealized versions of us. This is something that NM rejects from the start. But none of this detracts from the points we have just made against the third version of the structure and dynamics argument, for at least three reasons. First, while, as we have seen, the argument is unpersuasive if phenomenal functionalism is true, it is likewise unpersuasive if various alternatives to phenomenal functionalism are true, e.g. non-phenomenal functionalism or primitivism. So our criticism of the argument does not depend on the truth of phenomenal functionalism. Second, it is not as if these alternatives are unavailable to NM, though admittedly this issue requires much more discussion than I can give it here—for example it is an open possibility that the Nagelian Monist could hold the second kind of primitivism mentioned above, the kind that treats spatiotemporal expressions as those introduced in the course of explaining the scientific counterparts of Edenic properties. Finally, recall that our interest in the structure and dynamics argument in the first place is to ascertain whether we have any reason to favor RM over NM; but if Nagelian Monism is in conflict with phenomenal functionalism, that is also true of Russellian Monism—for it too is a position that denies that experiential properties are fundamental. So, again, this third version of the argument is not successful.

## **Conclusion**

I have considered three versions of the structure and dynamics argument, and suggested in each case that it is unpersuasive. The problem has not been—as is sometimes suggested (cf. Alter and Nagasawa 2012)—that the various notions of a structural truth we have examined are unclear or vague; by most standards they are clear enough. The problem is rather that there are lots of different notions here and no matter which one is in play the structure and dynamics argument is unpersuasive against NM.

Suppose now we agree that the structure and dynamics argument is unpersuasive; where does that leave the project we identified at the outset, the project of finding a view that is neither materialism nor dualism? In my view it leaves us much better off than we would

otherwise be. For the structure and dynamics argument is one of those arguments in philosophy that is supposed to narrow the field; if successful it would suggest, in particular, that RM is the only position here that fits the bill. If what I have been saying is right, however, this is not so. RM is *one* sort of view that fits the bill but is it not the *only* sort, for NM also remains a possibility. And this makes it more, rather than less, likely that a position that is neither materialism nor dualism is the truth about nature.

## References

- Alter, T. 2009 “Does the Ignorance Hypothesis Undermine The Conceivability and Knowledge Arguments?” *Philosophy and Phenomenological Research* 79:756-765
- Alter, T. 2013. Review of Pereboom’s *Consciousness and the Prospects of Physicalism*. *Mind* (2013) doi: 10.1093/mind/fzt008 First published online: March 22, 2013
- Alter, T., and Y. Nagasawa. 2012. “What Is Russellian Monism?” *Journal of Consciousness Studies* 19 (9–10): 67–95.
- Alter, T., and Y. Nagasawa, eds. Forthcoming. *Russellian Monism*. New York; Oxford University Press.
- Chalmers, D. 2002 “Consciousness and Its Place in Nature”, in Chalmers, D (ed.) *Philosophy of Mind: Contemporary and Classical Readings*, New York: Oxford University Press, pp.247-272.
- Chalmers, D. 2010. *The Character of Consciousness*. New York; Oxford University Press.
- Chalmers, D. 2012. *Constructing the World*. Oxford: Oxford University Press.
- Evans, G. 1980 “Things without the Mind—A Commentary on Chapter 2 of Strawson’s Individuals” in Zak Van Straaten (ed) *Philosophical Subjects* (Oxford: Clarendon Press) pp. 76-116.
- Johnston, M. 2002 ‘How to Speak of the Colors’ *Philosophical Studies* 68: 221-63.
- Kind, A. Forthcoming ‘Pessimism about Russellian Monism’, This Volume
- Lewis, D. 1995. “Should a Materialist Believe in Qualia?” *Australasian Journal of Philosophy* 73: 140–144.
- Lewis, D. 2009. “Ramseyan Humility.” In *Conceptual Analysis and Philosophical Naturalism*, ed. David Braddon-Mitchell and Robert Nola, 203–222. Cambridge, MA: MIT Press.
- Locke, J. 1975. *An Essay Concerning Human Understanding*, ed. P. H. Nidditch. Oxford: Oxford University Press.

- McClelland, T 2013 The Neo-Russellian Ignorance Hypothesis: A Hybrid Account of Phenomenal Consciousness, *Journal of Consciousness Studies*, Volume 20, Numbers 3-4, 2013 , pp. 125-151(27)
- Pereboom, D. 2011. *Consciousness and the Prospects of Physicalism*. New York: Oxford University Press.
- Pereboom, D. 2014 ‘Russellian Monism and Absolutely Intrinsic Properties’. In Kriegel, U. (ed). *Current Controversies in Philosophy of Mind*, London:Routledge pp.40-69.
- Russell, B. 1927. *The Analysis of Matter*. London: Kegan Paul.
- Stoljar, D. 2001. “Two Conceptions of the Physical.” *Philosophy and Phenomenological Research* 62 (2): 253–281.
- Stoljar, D. 2006. *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness* New York; Oxford University Press.
- Stoljar, D. 2009. “The Argument from Revelation.” In *Conceptual Analysis and Philosophical Naturalism*, ed. David Braddon-Mitchell and Robert Nola, 113–138. Cambridge, MA: MIT Press.
- Stoljar, D. 2010. *Physicalism*. London: Routledge.
- Stoljar, D. 2014 ‘Four Kinds of Russellian Monism’, in Kriegel, U. (ed.) *Current Controversies in Philosophy of Mind*. London: Routledge pp.17-35
- Weatherson, Brian and Marshall, Dan, 2013, "Intrinsic vs. Extrinsic Properties", *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2013/entries/intrinsic-extrinsic/>>.