Accepted manuscript.
Accepted for publication in Noûs, 10/10/2019.

Russellian Monism and Mental Causation

Torin Alter, University of Alabama
Sam Coleman, University of Hertfordshire

Russellian monism has been hailed as a breakthrough theory of consciousness and its place in nature. The theory is often described as accomplishing something that has eluded traditional versions of both physicalism and dualism: doing justice to what makes consciousness distinctive while at the same time adequately integrating consciousness into the natural causal order.[1] But critics have challenged that description.

According to Russellian monism, consciousness is constituted at least partly by quiddities: intrinsic properties that categorically ground dispositional (or structural) properties described by fundamental physics.[2] It follows that consciousness and physical properties are closely connected. But how closely? *The contingency thesis* says that the connection is contingent. For example, negative charge might have been categorically grounded by a consciousness-constituting quiddity that is distinct from the one that actually grounds it. Must Russellian monists accept the contingency thesis? What are its implications for their view?

Some suggest that familiar versions of Russellian monism entail the contingency thesis.[3] Further, some take that result to entail that on Russellian monism consciousness

[1] Alter and Nagasawa 2012, Chalmers 2013, Goff 2015.

[2] That characterization is rough (Alter and Nagasawa 2012, Chalmers 2013). But it will suffice for present purposes. For simplicity, we will usually refer to properties that physics reveals as "dispositional" rather than "dispositional (or structural)." Also, we should note that throughout the dispositions we have in mind are those described by (completed) physics, not phenomenal dispositions.

[3] Howell 2015, Robinson 2018, Pautz n.d. C.f. Stoljar 2001.

lacks physical efficacy—and thus that the theory fails to adequately integrate consciousness into nature.[4] Call that *the integration-failure argument.* If the integration-failure argument is sound, then Russellian monism cannot live up to its considerable promise. With respect to mental causation, this theory would be no better off than epiphenomenalism.[5]

However, we will argue, the integration-failure argument is unsound. The contingency thesis could be interpreted in two different ways: (i) as a claim about metaphysical possibility, that is, about what is possible *tout court* or (ii) as a claim about nomological possibility, that is, about what is possible relative to the actual laws of nature.[6] If the thesis is interpreted in way (i), then Russellian monists can argue that it does not threaten their integrationist aspirations concerning consciousness. If the thesis is interpreted in way (ii), then they can (and should) reject it. Either way, the integration-failure argument fails.

We will begin by saying more about what Russellian monism and the contingency thesis are. Then we will formulate and criticize a generic version of the integration-failure argument. This will make it easier to see why no version could succeed. Then we will apply our critique to a particularly strong version developed by Robert J. Howell.[7] Finally, we

[4] Howell 2015, Robinson 2018. Pautz (n.d.) develops related arguments. Cf. Robinson 1993.

[5] For the idea that epiphenomenalism fails to adequately integrate consciousness into nature, see Papineau 2002, ch. 1.

[6] Kripke 1972, Chalmers 1996.

[7] Howell 2015. Our critique also applies, *mutatis mutandis*, to Robinson's (2018) version and to Pautz's (n.d.) related arguments.

will discuss the idea that Russellian monists can reject the contingency thesis even if interpreted in way (i). Howell anticipates that response but objects that it is not viable. However, we will argue that his objection depends on misunderstandings about Russellian monism's commitments. In the end, we will conclude, the contingency thesis does not undermine the claim that Russellian monism is a breakthrough theory of consciousness and its place in nature.

Russellian monism and the contingency thesis

Russellian monism comes in panpsychist and panprotopsychist varieties. Panpsychist Russellian monism, or *Russellian panpsychism*, identifies consciousness-constituting quiddities with phenomenal properties. That identification leads to the panpsychist view that at least some fundamental physical entities have conscious experiences.[8] We will refer to the sort of phenomenal properties Russellian panpsychists posit as *microphenomenal* properties. We will refer to the more familiar sort of phenomenal properties, such as pain and phenomenal redness, as *macrophenomenal* properties. We will sometimes use the unqualified term "phenomenal" where the distinction does not matter.

Russellian panprotopsychism identifies consciousness-constituting quiddities with what David J. Chalmers calls *protophenomenal* properties. Chalmers defines protophenomenal properties as follows:

> [L]et us say that *protophenomenal* properties are special properties that are not phenomenal (there is nothing it is like to have a single protophenomenal property)

[8] Chalmers 2013, pp. 246-47.

but that can collectively constitute phenomenal properties, perhaps when arranged in the right structure.[9]

Russellian panprotopsychism does not entail panpsychism. Some even argue that Russellian panprotopsychism is a form of physicalism.[10] Indeed, the view might seem hard to distinguish from traditional physicalism. After all, on both Russellian panprotopsychism and traditional physicalism, non-phenomenal properties associated with microphysical entities collectively constitute macrophenomenal properties.[11] To avoid that result, Chalmers writes,

> …one can unpack the appeal to specialness in the definition [of protophenomenal properties] by requiring that (i) protophenomenal properties are distinct from structural properties and…(ii) there is an a priori entailment from truths about protophenomenal properties (perhaps along with structural properties) to truths about the phenomenal properties that they constitute.[12]

[9] Chalmers 2013, p. 259.

[10] Pereboom 2011, Montero 2010, 2015, Kind 2015.

[11] At least that is true of many, perhaps most, traditional physicalist theories. There might be exceptions, such as a version of eliminative materialism on which phenomenal terms express concepts of primitive properties that are not instantiated (Rorty 1965, P. S. Churchland 1984, P. M. Churchland 1984/1988, pp. 43-49).

[12] Chalmers 2013, p. 260. Regarding structural properties, Chalmers writes, "Here a structural property is one that can be fully characterized using structural concepts alone, which I take to include logical,

(i) and (ii) would seem to rule out the possibility that the properties traditional physicalists take to collectively constitute macrophenomenal properties are protophenomenal.[13] With those stipulations in place, Russellian panprotopsychism entails neither panpsychism nor traditional physicalism.

The distinction between panpsychist and panprotopsychist versions of Russellian monism will play a small role toward the end of our discussion.[14] For the most part, however, what we will say about one version applies, *mutatis mutandis*, to the other.[15]

Russellian monism is often advanced as a promising alternative to traditional forms of physicalism and traditional forms of dualism.[16] The basic idea could be put as follows. Traditional physicalist views either disregard or distort the distinctive features of

---

mathematical, and nomic concepts, perhaps along with spatiotemporal concepts…" (Chalmers 2013, p. 256). For discussion, see Alter 2015, Stoljar 2015.

[13] This is because traditional physicalists either (a) deny that there is an a priori entailment from physical truths to (relevant) phenomenal truths or (b) accept that there is such an entailment but deny that any of the physical truths in the entailment base are about non-structural properties (Chalmers 1996, 2003). These commitments are often implicit, but some build them into the definition of (traditional) physicalism (Goff 2017, ch. 1).

[14] See the section below on "Necessitarian Russellian monism".

[15] For examples of Russellian panpsychism, see Strawson 2006, Chalmers 2013, Goff 2017. For examples of Russellian panprotopsychism, see Stoljar 2001, Pereboom 2011, McClelland 2013, Chalmers 2013, Montero 2015, Coleman 2016.

[16] For examples of traditional physicalism, see Smart 1959, Armstrong 1968, Lewis 1972. For examples of traditional dualism, see Jackson 1982, Hart 1988, Gertler 2007.

consciousness. Traditional dualist views spawn problems around mental causation and thus fail to adequately integrate consciousness into nature. Russellian monism, however, lacks all of those drawbacks. [17] Russellian monists reject the doctrine they believe leads traditional physicalists to disregard or distort the distinctive features of consciousness: the doctrine that macrophenomenal properties are nothing over and above complexes of dispositional properties that physics reveals. [18] On Russellian monism the microphenomenal or protophenomenal properties that constitute macrophenomenal properties are no less fundamental than physical, dispositional properties. Traditional dualist views say something similar. But unlike traditional dualism, Russellian monism is designed to integrate consciousness deeply into the natural order. On this view, (proto)phenomenal properties are quiddities: they contribute to physical causation by categorically grounding basic physical dispositional properties such as mass and charge. The macrophenomenal properties constituted by (proto)phenomenal quiddities inherit this causal efficacy at the macro-level.[19]

[17] We take no stand on whether all traditional theories have those drawbacks. We claim only that the belief that they do helps motivate Russellian monism.

[18] Some physicalist views might be classified as traditional even though they do not entail that doctrine (Stoljar 2006, 2015). We take no stand on this issue.

[19] Howell 2015, pp. 32-33. Here we do not attempt to explain how this inheritance works. Providing an adequate explanation is a difficult problem, which may or may not have an adequate solution (Chalmers 2016). But it is not a problem we need solve here. Proponents of the integration-failure argument, such as Howell, assume *arguendo* that there is such an explanation and argue that even so Russellian monism fails to secure a role for consciousness in physical causation. At least, most such proponents make this assumption. Robinson (2018) might be an exception. In our estimation, he does not well distinguish the integration-failure

Let us turn to the contingency thesis. It concerns scenarios (or worlds) in which quiddities are *swapped* or *absent*.[20] In a swapped-quiddity scenario, which (proto)phenomenal quiddities ground which dispositional properties varies in some way from the grounding relations that, according to Russellian monism, obtain in the actual world. In an absent-quiddity scenario, dispositional properties that fundamental physics describes lack (proto)phenomenal quiddistic grounding altogether. The contingency thesis says that swapped- and absent-quiddity scenarios are possible.

Howell describes three scenarios that can be used to illustrate swapped-quiddity scenarios:

> Consider a world w1 in which R, phenomenal redness, grounds the property of negative charge given the causal laws governing R in w1. Now consider world w2 where G, phenomenal greenness, is covered by those same laws so that G grounds the causal powers associated with negative charge and R instead grounds the powers associated with negative spin. Finally, consider a third world, w3, in which

argument from objections to Russellian monism that concern micro-to-macro inheritance of causal efficacy. In any event, the latter objections should be set aside if the integration-failure argument is to be properly assessed.

[20] Morris 2016.

> the laws are such that either R or G can ground the powers of negative charge—R and G are governed by exactly the same laws in exactly the same ways.[21]

Suppose w1, where phenomenal quiddity R categorically grounds negative charge, is actual. In that case w2, where negative charge is categorically grounded in not R but distinct phenomenal quiddity G, is a swapped-quiddity scenario. And if in the actual world R and no other phenomenal quiddity grounds negative charge, then w3, where either R or G can ground negative charge, is also a swapped-quiddity scenario.[22] One might also consider a world w4 in which negative charge lacks categorical grounding altogether. If w1 is actual, then w4 would be an absent-quiddity scenario.[23]

---

[21] Howell 2015, p. 28. We assume that "phenomenal redness" and "phenomenal greenness" are proxies for microphenomenal properties, which presumably are phenomenally quite different from phenomenal redness and phenomenal greenness.

[22] In this paper, we consider w3 as a potential counterfactual scenario. But w3 raises a question about how the world might actually turn out: might it actually turn out that negative charge is actually grounded disjunctively, as in w3? That question might give rise to an epistemic variant of the integration-failure argument that we consider in this paper. We believe that variant would face difficulties similar to those we raise for the argument we consider. But that is a topic for another occasion.

[23] Howell (2015) does not discuss absent-quiddity scenarios. To some philosophers, they might seem especially dubious. Indeed, even those who accept the coherence of swapped-quiddity scenarios might nonetheless reject absent-quiddity scenarios as incoherent—perhaps because the very idea of ungrounded dispositional properties is incoherent, or perhaps because it is incoherent that ungrounded dispositional properties should be causally efficacious. Those so inclined may simply ignore all references we make to

One might imagine scenarios where the swaps or the absences are systematic, such as an absent-quiddity scenario in which every basic dispositional property is categorically ungrounded. One might also imagine relatively isolated swaps or absences.[24] For our purposes, however, scenarios such as w1-w4 (or protophenomenal versions thereof, featuring the swapping or absence of protophenomenal properties) will suffice.[25]

The integration-failure argument

Let us turn to the integration-failure argument, according to which the contingency thesis undermines the claim that on Russellian monism consciousness is adequately integrated into nature. In the next section, we will discuss Howell's version of that argument. First, in

absent-quiddity scenarios. We add those scenarios to the mix for completeness. None of our arguments depend essentially on claims about them.

[24] Note that to generate swapped- or absent-quiddity scenarios, we need not assume that on Russellian monism there is (in either the actual world or the scenario in question) a one-to-one mapping from distinct basic dispositional properties (mass, spin, charge, etc.) to distinct (proto)phenomenal quiddities. Indeed, we need not assume that there is actually more than a single (proto)phenomenal quiddity. Suppose there is actually only one such quiddity Q that grounds all basic dispositional properties, or some class thereof. To generate a swapped-quiddity scenario, imagine a scenario in which Q is replaced, in at least some token instances, with one or more alien (proto)phenomenal quiddities: (proto)phenomenal quiddities that might but do not actually exist.

[25] One might also consider *absent dispositions*: a world of pure (proto)phenomenal quiddities, devoid of basic dispositional properties. However, that scenario is not directly relevant to the integration-failure argument that we will consider. Swapped-quiddity, absent-quiddity, and absent-disposition scenarios correspond roughly to more familiar inverted spectrum, zombie, and ghost scenarios (Chalmers 1996, Goff 2010).

this section, we will discuss a generic version. This will make it easier to see why no version, including Howell's, could succeed.

The generic integration-failure argument has four main steps:

1. Russellian monism entails the contingency thesis.
2. The contingency thesis entails that consciousness has no physical effects.
3. Any view on which consciousness has no physical effects fails to adequately integrate consciousness into nature.
4. Therefore, Russellian monism fails to adequately integrate consciousness into nature.

That argument appears to be valid, at least at first glance. And initially the premises might seem plausible. Regarding premise 1, consider that the distinction between quiddities and dispositional properties is central to Russellian monism. The theory is formulated in terms of that (or a similar) distinction.[26] Having taken such pains to distinguish the two sorts of property, how could Russellian monists deny that such properties could come apart in the ways the contingency thesis implies?

Regarding premise 2, the possibility of swapped- or absent-quiddities entails that, at the most fundamental microphysical level, the world might have been dispositionally the same regardless of the (proto)phenomenal quiddistic facts, that is, regardless of whether consciousness-constituting quiddities were swapped or absent. If the world might have been dispositionally the same despite any (proto)phenomenal quiddistic difference, why

[26] Stoljar 2015. Alter and Nagasawa (2012, p. 72) call it "the central distinction."

think that (proto)phenomenal quiddities make any difference to physical causation? Wouldn't the possibility of swapped- or absent-quiddities indicate that, with respect to physical causation, such quiddities are, as Howell writes, "just along for the causal ride"?[27] And if such quiddities are physically inefficacious, then so is consciousness—assuming that they constitute consciousness, as Russellian monism says.

Regarding premise 3, consider epiphenomenalism. Epiphenomenalism is widely rejected because it is seen as failing to adequately integrate consciousness into nature. And epiphenomenalism is thought to have that drawback precisely because it implies that consciousness has no physical effects.[28] Similar reasoning would apply to any view that has that implication: if a view implies that consciousness has no physical effects, then that view fails to adequately integrate consciousness into nature.

Nevertheless, the generic integration-failure argument is unconvincing. The contingency thesis asserts that swapped- and absent-quiddity scenarios are possible. But in what sense of possibility? Does the thesis say that such scenarios are *metaphysically* possible? Or does the thesis say that they are *nomologically* possible, that is, possible relative to the actual laws of nature? In other words, is the thesis that the set of metaphysically possible worlds includes swapped- and absent-quiddity scenarios? Or is the thesis that the set of metaphysically possible worlds in which all actual natural laws obtain includes swapped- and absent-quiddity scenarios? Either way, we will argue, the generic integration-failure argument fails.

[27] Howell 2015, p. 34.

[28] Papineau 2002, ch. 1.

Horn 1: Suppose that the contingency thesis says that swapped- and absent-quiddity scenarios are metaphysically possible. Call that *the metaphysical contingency thesis.* On this interpretation, premise 1 (Russellian monism -> the metaphysical contingency thesis) might seem plausible, for the reasons given above. However, on this interpretation, premise 2 (the metaphysical contingency thesis -> consciousness is physically inefficacious) is doubtful. For why should one take the mere metaphysical possibility of swapped- or absent-quiddity scenarios to have any implications for physical causation in the actual world?[29]

Arguably, causal efficacy does not in general require metaphysically necessary connections among relata. For example, the assumption that event A actually causes event B does not entail that A's occurring metaphysically necessitates B's occurring. Perhaps A causes B in virtue of those events falling under contingent causal laws. Thus, the metaphysical possibility of A's occurring without B's occurring does not threaten the claim in the actual world A is causally efficacious with respect to B.

A parallel point holds for the role of categorical grounding in causation. Suppose property G categorically grounds dispositional property D. Suppose also that G is thereby causally efficacious with respect to D's effects: if D's being instantiated has some effect then, by categorically grounding D, G's being instantiated contributes to the causation.[30] It

[29] Hawthorne 2002, p. 44.

[30] Arguably, the second supposition follows from the first. As Stoljar (2001, p. 267) writes, "As Mark Johnston puts it…if dispositional properties are efficacious, they are so 'at one remove and by courtesy' (1992; p. 235). …But it seems clear that, if dispositional properties *are* causally efficacious, so too are their categorical grounds." As we understand Stoljar and Johnston, they do not mean to imply that a disposition

does not follow from this that G grounds D in all possible worlds in which G is instantiated. Nor does it follow that G grounds D in all possible worlds where D has the effects it actually has—or even that, in each such world, D is categorically grounded at all. In the actual world, G might contribute to D's effects in virtue of metaphysically contingent grounding laws. Thus, the claim that, in the actual world, G contributes to D's effects is not threatened by metaphysical possibilities such as G's being instantiated without grounding D, or D's having the effects it actually has without being categorically grounded by G (or without being grounded at all).

In short, causation is arguably metaphysically contingent. That contingency applies not just to causally related events but also to the causal relationship between dispositional properties and the properties that categorically ground them. Russellian monists could argue that there is no reason why there should be an exception to that general rule in cases where the dispositional properties are those described by fundamental physics and the properties that categorically ground them are (proto)phenomenal quiddities. But if the contingency thesis concerns metaphysical possibility, then premise 2 (the metaphysical contingency thesis -> consciousness is physically inefficacious) relies on there being such an exception. In other words, premise 2 would then rely on an assumption that runs contrary to the metaphysical contingency of causation such as this: if by grounding a dispositional property D a (proto)phenomenal quiddity Q actually helps D cause an effect E then, in all possible worlds where D causes E, Q helps D cause E by grounding D. Russellian monists

---

cannot be causally efficacious unless it has a categorical ground. Instead, on our understanding, their claim is that a grounded disposition is causally efficacious only if its ground is, i.e., for any disposition D that has a categorical ground G, if D is causally efficacious then so is G.

can reject such assumptions. Thus, they can reject premise 2, if the contingency thesis is interpreted as the metaphysical contingency thesis.

Horn 2: Suppose now that the contingency thesis says that swapped- and absent-quiddity scenarios are nomologically possible. On that *nomological contingency thesis*, swapped- and absent-quiddity scenarios are not only metaphysically possible but also consistent with the actual laws of nature. On this interpretation, premise 2 (the nomological contingency thesis -> consciousness is physically inefficacious) is defensible. Here is why. Physical causation is governed by physical laws. If the nomological contingency thesis is true then, one might argue, the actual physical laws are insensitive to facts about (proto)phenomenal quiddities: with respect to those laws, neither which (proto)phenomenal quiddities are in play nor whether there are such quiddities at all makes any difference.[31] That suggests that (proto)phenomenal quiddities make no difference to physical causation. But on Russellian monism, such quiddities constitute consciousness. Plausibly, then, on the nomological-possibility interpretation, premise 2 is true: if the nomological contingency thesis is true, then consciousness is physically inefficacious.

But on the nomological-possibility interpretation, Russellian monists can (and should) reject premise 1 (Russellian monism -> the nomological contingency thesis). On their view, the laws governing (proto)phenomenal quiddistic grounding—that is, the laws

[31] The argument from the nomological contingency thesis to the conclusion that the actual physical laws are insensitive to facts about (proto)phenomenal quiddities is not entirely straightforward. For example, the argument depends on assumptions connecting the sensitivity of laws and counterfactual considerations, as well as assumptions about the sort of swapped- and absent-quiddity scenarios that are assumed to be nomologically possible. For the sake of argument, however, we will grant that the argument is sound.

expressing the categorical grounding relationship between (proto)phenomenal quiddities and dispositional properties—are among the laws of nature. After all, the relevant dispositional properties are among the most fundamental properties described by physical science. And those (proto)phenomenal quiddistic grounding laws (perhaps in conjunction with other laws of nature) rule out swapped- and absent-quiddity scenarios as impossible.[32]

For example, suppose that, according to the actual quiddistic grounding laws, R grounds negative charge, as in w1. In that case, w2, where not R but G grounds negative charge, is nomologically impossible. Such a world is ruled out by the actual quiddistic grounding laws: laws that, on the nomological contingency thesis, are assumed to obtain in all relevant worlds. Similar reasoning rules out other swapped- and absent-quiddity scenarios as nomologically impossible.[33] So, if the contingency thesis is interpreted as the nomological contingency thesis, then the Russellian monist can (and should) reject it, and with it premise 1 of the generic integration-failure argument.[34]

[32] Henceforth when referring to (proto)phenomenal quiddistic grounding laws we will omit "(proto)phenomenal" and assume this qualification is understood.

[33] Here is another example. Suppose that, according to the actual quiddistic grounding laws, R and only R grounds negative charge. Consider w3, where negative charge is disjunctively grounded in R or G. In w3, negative charge can be instantiated even in circumstances where R is not, in violation of the actual grounding laws.

[34] Are quiddistic grounding laws *physical* laws? This issue is largely terminological (Chalmers 2013). The term "physical" might be construed narrowly, in a way that excludes any reference to quiddistic information. In that case, the answer is no, quiddistic grounding laws are not physical. The term "physical" might instead be construed widely, in a way that has no such implication. In that case, the answer might well be yes.

Thus, the generic integration-failure argument fails. If it seems convincing, this is likely due to an equivocation between the two interpretations of the contingency thesis that we have distinguished. Premise 1 seems plausible because one assumes the thesis concerns metaphysical possibility. Premise 2 seems plausible because one assumes the thesis concerns nomological possibility. If either notion of possibility is employed consistently in both premises, then the Russellian monist can reasonably reject one premise or the other.

In response to our dilemma argument, a proponent of the integration-failure argument might suggest that the contingency thesis concerns neither nomological nor metaphysical possibility but another sort of possibility altogether. However, it is entirely unclear what that other sort would be. What sort of possibility would render both premises 1 and 2 plausible? We doubt that anything fits the bill. On the contrary, we contend that the argument likely derives its apparent force from an equivocation between nomological and metaphysical notions of possibility.

A second response would be that our argument fails to save the claim that on Russellian monism consciousness is adequately integrated into nature. On Russellian monism, it might be said, (proto)phenomenal quiddities enter into the causal story only indirectly, as categorical bases for fundamental physical dispositions. That, the response runs, does not suffice for genuine physical efficacy. To have such efficacy, quiddities would have to enter more directly into the causal chains of events that physics describes. And if (proto)phenomenal quiddities lack physical efficacy then, on Russellian monism, so does consciousness.

But that response fails for two reasons. One is that it does not save the generic integration-failure argument, in which the contingency thesis plays a central role. Rather,

the response replaces that argument with a separate argument, one that does not invoke the contingency thesis at all. If that separate argument were sound, then the generic integration-failure argument would not be needed, for (proto)phenomenal quiddities (and hence consciousness, on Russellian monism) would lack physical efficacy for reasons independent of the contingency thesis.[35]

Further, that point aside, the response risks making the criteria for physical efficacy overly narrow. Dispositional properties do not in general causally screen off their categorical bases (think of the relation of a vase's fragility, and the categorical property that grounds that disposition, to the event of the vase breaking).[36] Why think that they do so at the most fundamental level, where the dispositional properties are those described in physics and the categorical bases are (proto)phenomenal quiddities, as on Russellian monism? And if the objector should insist on her narrow conception of physical efficacy, then Russellian monists could retaliate by rejecting premise 3, that is, by arguing that physical efficacy in the objector's narrow sense is not required for adequately integrating consciousness into nature. If consciousness-constituting quiddities contribute to physical causation by categorically grounding physical dispositional properties, then that is

[35] Although strictly independent, on some views there will be close connections between the proposed response and the integration-failure argument. In particular, consider the Martin-Heil "powerful-qualities" view (Heil 2003, Martin 2008), on which dispositional properties are identical to their categorical grounds (a view with which Howell (2015: p. 27, fn. 16) expresses sympathy). On this view, there is plainly no room for quiddities to have physical efficacy independently of the dispositional features they ground. We discuss related views below, in the section on "Necessitarian Russellian monism".

[36] Recall Stoljar and Johnston's point: if dispositions are causally efficacious, then so too are their categorical grounds. See above, fn. 30. For the relevant notion of screening off, see LePore and Loewer 1987.

integration enough—and that is so even if categorical grounding does not suffice for physical efficacy in the objector's narrow sense.

A third response to our argument would be to reject the claim that causation is metaphysically contingent—a claim we used in objecting to the horn-1 version of premise 2. In other words, one might respond by invoking a necessitarian view about causation. But in that case, the Russellian monist could justifiably adopt a necessitarian view about her quiddistic grounding laws, thereby rejecting premise 1 (Russellian monism -> the metaphysical contingency thesis). So, the generic argument would still fail.[37]

Our objection to the generic integration-failure argument applies to any version of the integration-failure argument. The reason is simple: any version of the integration-failure argument will involve premises corresponding to premises 1 and 2 of the generic version. And any such pair of premises will either equivocate on different notions of possibility, rendering the argument invalid, or be impaled on one or other horn of the dilemma we brought against the generic version. With that in mind, let us turn to Howell's version.

Howell's argument

[37] See the section below on "Necessitarian Russellian monism." Given necessitarianism about causation, necessitarianism about quiddistic grounding laws might not be merely optional for Russellian monists. We have claimed that those laws are causal, at least insofar as they have causal implications. If our claim is correct, then necessitarianism about causation might entail necessitarianism about those laws.

Howell models his argument on Jaegwon Kim's well-known causal-exclusion argument.[38] Kim's reasoning, he argues, can be modified to show that, with respect to integrating consciousness into nature, Russellian monism fares no better than traditional dualism—a view that, as Kim, Howell, and Russellian monists agree, fares poorly in that regard.

Although Kim's argument is complex, the basic idea is fairly simple: if mental and physical properties are distinct, then they compete for causal influence—and mental properties lose, i.e., they are excluded from physical efficacy.[39] Russellian monism seems to circumvent the problematic competition. If consciousness-constituting (proto)phenomenal properties categorically ground the dispositional properties described by physics, as Russellian monism says, then the mental and the physical seem less like distinct properties and more like aspects of a single, unified property—aspects that play different but equally vital roles in physical causation. But Howell argues that, in the final analysis, Russellian monism merely relocates the competition that Kim identified. Competition between *properties* is replaced by competition between *aspects* of properties: (proto)phenomenal and dispositional aspects. But competition there is. And, Howell argues, the (proto)phenomenal aspects lose. That would undermine the claim that on Russellian monism consciousness is physically efficacious.[40]

[38] Kim 1989, 2000.

[39] As Howell recognizes, various responses to Kim's argument have been developed (he cites Yablo 1992, Bennett 2003, Shoemaker 2007, Ehring 2011, Wilson 2011, and the Heil and Mele (1993) collection; see also Burge 2007, List and Stoljar 2017). But Howell finds none satisfactory.

[40] Howell 2015, pp. 32-33.

As Howell explicates Russellian monism, it posits what he calls *RM properties*. He writes, "an RM property is a property that has a phenomenal categorical ground and some causal dispositions"[41] Positing these properties, he implies, is what is supposed to allow Russellian monists to say that consciousness is physically efficacious and thus to avoid Kim-style exclusion worries. He describes RM properties as having two aspects: one that "ground[s] phenomenal resemblance relations" and another that "ground[s] resemblance relations between causal profiles of RM properties."[42]

Howell's explication of Russellian monism raises at least two concerns. First, his explication might be taken to imply that *by definition* only one of the two aspects of RM properties has causal efficacy, namely, the aspect that "ground[s] resemblance relations between causal profiles of RM properties." Clearly, that would be a misunderstanding. Howell is trying show, not stipulate, that on Russellian monism (proto)phenomenality would lack physical efficacy. We will therefore refer to the two aspects he distinguishes simply as *(proto)phenomenal* and *dispositional* (or *physical dispositional*)—terminology he also uses, albeit not exclusively. Second, one might worry that by emphasizing aspects of properties, rather than the properties themselves, Howell smuggles in substantive assumptions that Russellian monists need not accept. For example, one might take his suggestion that Russellian monism posits two aspects of RM properties as implying adherence to a dualist view on which the two aspects are ontologically distinct entities—a

[41] Howell 2015, p. 32. "Protophenomenal" can be substituted for "phenomenal" in the definition of "RM property" and in Howell's argument to give versions that apply to Russellian panprotopsychism (Howell 2015, pp. 33-34).

[42] Howell 2015, p. 32.

view that a Russellian monist might well reject (for example, she might instead hold that there is but one aspect here, referred to under two different conceptions). But we will assume that his emphasis on aspects smuggles in no such assumptions.[43]

Howell claims, however, that the two aspects of RM properties are not only conceptually distinct but also modally separable. He does not merely assume that this modal-separability claim is true. Instead, he supports it by providing clear illustrations, namely, three of the scenarios we have been discussing: w1, in which R (phenomenal redness) grounds negative charge; w2, in which G (phenomenal greenness) grounds negative charge and R grounds negative spin; and w3, in which R or G governs negative

[43] The two concerns we raise in this paragraph are specific instances of a more general concern. In his article (Howell 2015), Howell does not say much about the metaphysics of aspects. What are aspects, exactly? And how do aspects relate to properties? For example, is an aspect a higher-order property, e.g., a property of an RM property? Or does an aspect relate to a property in something like the way a part relates to a whole? Howell does not say. He says more about aspects in his 2013 book, *Consciousness and the Limits of Objectivity*. But he does not cite that book in his 2015 article. Also, in his 2013 book aspects would appear to be intensional entities, and it is not clear that intensional entities could have the physical efficacy that, in his 2015 article, he ascribes to dispositional aspects of RM properties. Still, we will set all such issues aside. We do not believe they need be resolved in order to assess his criticisms of Russellian monism, as presented in his 2015 article. We should also note that we have adapted Howell's terminology to be more in line with our own. For example, we describe the two aspects of RM properties Howell distinguishes as "categorical" and "dispositional", whereas he sometimes writes as though the two aspects he distinguishes are both aspects of the categorical ground of a disposition (see, for example, Howell 2015, p. 29). We hope this does not create confusion. In any event, nothing substantive turns on it.

charge. In those scenarios, the two aspects come apart. And if Russellian monism were true, Howell suggests, those scenarios would be possible.

In Howell's view, the idea that the two aspects of RM properties are "distinct and separable" opens the door to Kim-style exclusion worries.[44] He writes,

> In the case of phenomenal causation, we want phenomenal properties to have causal power in virtue of their phenomenality. That [for Russellian monists] means that we want the properties to cause things in virtue of that which grounds the similarity between R in w1 and R in w2. But that doesn't appear to be the case since R in w1 and R in w2 are causally quite dissimilar.[45]

Here (with some small modifications)[46] is Howell's summary of his argument:

1. There are two distinct and separable aspects of RM properties, the (proto)phenomenal aspects and the dispositional aspects.
2. All physical events have sufficient causes in virtue of the dispositional aspects of RM properties.

[44] Howell 2015, p. 32.

[45] Howell 2015, p. 29.

[46] We replaced his descriptors for the two aspects of RM properties with the ones we prefer, for the reasons described above. And we added "(proto)" before "phenomenal", so that the argument applies to Russellian panprotopsychism, as he intends (2015, pp. 33-34). Howell's formulation—and thus ours—ignores complications that are not directly relevant. For example, his omits a premise ruling out the possibility of rampant overdetermination (Howell 2015, p. 24, fn. 9), and thus ours does too.

3. Therefore, the (proto)phenomenal aspects of RM properties make no unique causal contribution to the physical world.

Like the generic integration-failure argument, Howell's begins by associating Russellian monism with a claim that corresponds to the contingency thesis. He does this in his premise 1, with the claim that RM properties have two distinct and separable aspects. Also like the generic argument, he infers (in his premise 2 and conclusion) that on Russellian monism (proto)phenomenality, and thus consciousness, is physically inefficacious. And his overall moral is the same as the generic argument's conclusion: because on Russellian monism consciousness is physically inefficacious, the theory fails to live up to its promise to adequately integrate consciousness into nature. Presentational differences notwithstanding, Howell's argument would appear to be a version of the integration-failure argument, cast in terms of aspects of properties rather than properties themselves.[47]

Thus, one might expect Howell's argument to suffer from the same problem that we raised for the generic integration-failure argument. Indeed, this appears to be the case (though this requires a qualification, as we will explain at the end of this section). This can be seen by scrutinizing Howell's premise 1, which says that the two aspects of RM properties are distinct and separable. Here one can raise a question that corresponds to our question about the contingency thesis: in what sense are the two aspects separable? Are they metaphysically separable, as in the metaphysical contingency thesis? Or are they

[47] Although Howell presents his argument as a modification of Kim's exclusion argument, there are notable differences. In particular, Howell's centers on a modal-separability claim, and Kim's does not.

nomologically separable, as in the nomological contingency thesis? In other words, is the claim that the set of all metaphysically possible worlds includes worlds where the aspects have been swapped (or are absent)? Or is the claim that the set of all metaphysically possible worlds in which all actual natural laws obtain includes worlds where the aspects have been swapped (or are absent)? Call those *the metaphysical-separability interpretation* and *the nomological-separability interpretation* of Howell's integration-failure argument, respectively. In effect, on either interpretation his argument gets impaled on one or other horn of the dilemma that, we argued above, undermines the generic integration-failure argument.

Suppose first that the metaphysical-separability interpretation of Howell's argument is correct. With respect to w1-w3 (and presumably w4), his premise 1 would then say that those scenarios are metaphysically possible. Russellian monists might accept premise 1, thus interpreted. But in that case, they can reject his premise 2, the claim that all physical events have sufficient causes in virtue of the dispositional aspects of RM properties. They can base that rejection on the principle that causation is metaphysically contingent. For example, they can argue that in w1 negative charge has physical effects at least partly in virtue of R even though in w2 negative charge has those same effects at least partly in virtue of G (rather than R). This is so, they can argue, because the quiddistic grounding laws in w1 differ from those in w2: those laws differ with respect to which

(proto)phenomenal aspect of the relevant RM property plays the relevant categorical grounding role.[48]

In general, if causation is metaphysically contingent, then the assumption that the same categorical grounding role can be played by different (proto)phenomenal aspects across different metaphysically possible worlds does not entail that, in the actual world, the (proto)phenomenal aspect that actually plays that role lacks physically efficacy. The actual quiddistic grounding laws might be metaphysically contingent. But they still determine which (if any) (proto)phenomenal aspects actually do the categorical grounding. And on Russellian monism, it is precisely via such categorical grounding that the (proto)phenomenal contributes to physical causation. Thus, in effect, on the metaphysical-separability interpretation, Howell's integration-failure argument gets impaled on horn 1 of the dilemma we brought against the generic version.

Now suppose instead that the nomological-separability interpretation of Howell's argument is correct. With respect to w1-w3 (and presumably w4), his premise 1 would then say that those scenarios are nomologically possible, i.e., that they are not only metaphysically possible but also consistent with all actual natural laws. The nomological possibility of such scenarios would entail that the natural laws are sensitive to the dispositional aspects of RM properties but not to their (proto)phenomenal aspects, that is, to the grounded aspects and not to the aspects doing the grounding. And in that case, Howell's premise 2 (which says that all physical events have sufficient causes in virtue of

[48] Here and elsewhere when discussing Howell's argument, we use "quiddity" (and cognates) to refer not to properties but rather to aspects: (proto)phenomenal aspects of RM properties (which are taken to ground physical dispositional aspects).

the dispositional aspects of RM properties) would be plausible, given that physical causation is governed by the natural laws.

But on the nomological-separability interpretation, Russellian monists can (and should) reject Howell's premise 1. Again, on their view the natural laws include quiddistic grounding laws, and those laws rule out swapped-quiddity scenarios (and absent-quiddity scenarios) as impossible. For example, suppose that according to the actual quiddistic grounding laws R grounds negative charge, as in w1. In that case, in all worlds where the actual natural laws obtain, R grounds negative charge. It follows that w2, where not R but G grounds negative charge, would not be among those worlds. In other words, w2 would be nomologically impossible, contrary to premise 1, on the nomological-separability interpretation. To put the point differently, it is simply part of the Russellian monist view that some natural laws are sensitive to which (proto)phenomenal quiddities are in play. Therefore, it would be question-begging for Howell to claim that there are nomological possibilities that are inconsistent with that sensitivity. In effect, on the nomological-separability interpretation, Howell's argument gets impaled on horn 2 of the dilemma we brought against the generic argument.

Thus, the dilemma we raised for the generic argument would appear to apply equally to Howell's argument, *mutatis mutandis*. His argument fails both on the metaphysical-separability interpretation and on the nomological-separability interpretation. That result illustrates and reinforces the claim we made at the end of the previous section: any version of the integration-failure argument will be vulnerable to the dilemma we raised against the generic version.

Now for the qualification we alluded to above (six paragraphs back). In presenting his argument, Howell makes some moves we have not mentioned. We do not believe these omissions affect our criticisms. But we will briefly discuss two that might seem to.[49]

In a section on "Categorical bases and causal grounds," Howell writes:

> Perhaps the very same categorical property C can play different roles r1 and r2 in different worlds w1 and w2, but this just means that C must have different features in w1 and w2 in virtue of which it can play those roles.[50]

Call that claim (if "categorical property C can play different roles r1 and r2 in different worlds w1 and w2," then "C must have different features in w1 and w2 in virtue of which it can play those roles") *Different Roles Means Different Features.*

Might *Different Roles Means Different Features* help Howell answer our charge that his argument falls prey to our dilemma?[51] If so, it is not obvious how. *Different Roles Means Different Features* has no bearing on horn 2 of our dilemma, where we argue that the Russellian monist can (and should) deny the premise that Russellian monism entails the nomological contingency thesis. Denying that premise does not require assuming that the same categorical property plays different roles in different worlds, let alone that a

[49] One might take our omissions to count against our interpretation of Howell's argument as giving a version of the integration argument. But even if that is right, we see no interpretation that fares any better against our criticisms.

[50] Howell 2015, p. 30. Howell seems to use "features" and "aspects" interchangeably.

[51] Thanks to an anonymous referee for *Noûs* for bringing our attention to this issue.

categorical property could do that without having different features in virtue of which it does. Likewise, on horn 1 of our dilemma, we do not assume that the same categorical property plays a different role in different worlds. Instead we appeal to one of Howell's own cases, in which the same role is played by the different categorical aspects in different worlds (R plays the role of grounding negative charge in w1, whereas G plays that same role in w2). *Different Roles Means Different Features* does not seem relevant here either.

To be sure, there is a tension between *Different Roles Means Different Features* and our horn 1 argument. In giving our horn 1 argument, we are defending *non-necessitarian* Russellian monism: the version of Russellian monism on which causation is taken to be metaphysically contingent. On causal non-necessitarianism, whether a given categorical property C grounds this or that disposition might depend not on C's features but instead on which causal laws obtain in a given world.[52] That might result in C grounding different dispositions in different worlds even though C does not have different features in those worlds: the difference in which disposition C grounds might derive from differences in the laws obtaining at the different worlds. That result seems to conflict with *Different Roles Means Different Features.*[53]

[52] Here we assume that Cambridge properties such as *being such that law L obtains* are not features of C. If one rejects that assumption, then the non-necessitarian can accept *Different Roles Means Different Features.* But in that case *Different Roles Means Different Features* would even more clearly fail to help Howell defend his argument against our criticisms.

[53] Causal non-necessitarianism allows situations such as the following one. The laws in world wA are such that categorical property C grounds disposition D1, i.e., such that C plays the D1 role. The laws in a second world wB differ, such that C grounds disposition D2, i.e., such that C plays the D2 role, where D2 and D1

But what does that show? *Different Roles Means Different Features* is tantamount to a flat-out denial of a central clam of causal non-necessitarianism, that the causal efficacy of a categorical property in a world might depend on which laws obtain in that world rather than on features of that property. So, it would be dialectically unacceptable to simply assert *Different Roles Means Different Features* in objecting to our horn 1 argument. Of course, one might give independent reasons to reject causal non-necessitarianism. But in that case, the Russellian monist can respond by adopting a necessitarian view, and thus rejecting the contingency thesis (we discuss this in the next section).

Howell invokes other considerations, in addition to *Different Roles Means Different Features*, that we have not mentioned. But none come any closer to insulating his argument from our criticisms. We will discuss one more. Howell claims that the modal separability of the quiddistic and dispositional aspects of RM properties "helps us to see quite clearly" that it is only the dispositional aspect that is "really causally relevant."[54] When he makes that claim, he draws an analogy that might seem to strengthen his position. He remarks that, "We can see it just as easily as we can tell that it is not the redness of a brick that explains why it breaks a window."

But Howell's analogy does not go through. The *metaphysical* possibility of worlds in which the brick breaks the window but is not red would *not* show that its redness does

---

are distinct. So, C plays different roles in different worlds (the D1 role wA, the D2 role in wB), but contrary to *Different Roles Means Different Features*, C does not have different features in those worlds in virtue of which it plays those roles. C can do that because those worlds have different causal laws, and different causal laws can turn the same property, or aspect, to different causal purposes across worlds.

[54] Howell 2015, p. 31.

not explain, or does not help cause, the window's breaking (assuming that causation is metaphysically contingent). What might show that would be the corresponding *nomological* possibility. There is such a nomological possibility if the relevant causal laws are indifferent to the brick's color, as is plausible.[55] But in that respect, the brick's color is disanalogous to the Russellian monist's quiddities: as we have explained, the non-necessitarian Russellian monist can (and should) *deny* that the actual causal laws, including those governing quiddistic grounding, are indifferent to which quiddities ground which dispositions. She need not (and should not) accept the modal separability claim if that claim is construed in terms of nomological possibility. Her non-necessitarianism concerns metaphysical, not nomological, possibility.

Thus, our dilemma applies once again. If the modal separability Howell has in mind concerns *metaphysical* possibility, then his modal separability premise does not establish his causal conclusion that the quiddistic aspects of RM properties are physically inefficacious—just as a corresponding modal separability premise, if it concerns metaphysical possibility, does not establish his causal conclusion that the redness of a window-breaking brick does not help cause the window to break. If instead he has *nomological* possibility in mind, then the Russellian monist can (and should) deny Howell's modal separability premise regarding the dispositional and quiddistic aspects of RM properties. By contrast, in the case of the brick, it is not plausible to deny that there

[55] This is plausible in the sort of ordinary window-breaking case that Howell plainly has in mind. There are, of course, conceivable cases in which a brick's redness does help cause it to break a window, e.g., where I am disposed to throw all and only *red* bricks I find at windows, or where the microscopic differences in surface texture between blue and red bricks turns out to matter to the fragility of a particular pane of glass.

are nomologically possible worlds in which the brick breaks the window and is not red. Thus, Howell's analogy fails to support his case for the physical inefficacy of the quiddistic.

Necessitarian Russellian Monism

Two sections back, we said that Russellian monists could reject the metaphysical contingency thesis by adopting a necessitarian view on which quiddistic grounding laws hold with metaphysical necessity.[56] Adopting this view would block the generic integration-failure argument at its first step (Russellian monism -> the contingency thesis). Russellian monists might respond to Howell's integration-failure argument in the same way: they might deny that the (proto)phenomenal and dispositional aspects of RM properties are modally separable, maintaining instead that the aspects are connected by metaphysical necessity. Howell considers this strategy, but he objects that necessitarianism is not a viable option for Russellian monists. In this section, we will discuss his arguments for that conclusion. First, however, we will say more about what necessitarian Russellian monism is and how it would block the generic (and any) integration-failure argument.

Necessitarian Russellian monism comes in three varieties, holding either that: (i) any possible world that is a dispositional duplicate of the actual world is a (proto)phenomenal duplicate; or (ii) any possible world that is a (proto)phenomenal

[56] Necessitarian Russellian monism is discussed in Chalmers 2013 and defended in Mørch 2014 (a panpsychist version) and Coleman 2015 (a panprotopsychist version).

duplicate of the actual world is a dispositional duplicate; or both (i) and (ii).[57] Howell focuses on the variety that upholds only (ii). Thus he writes,

> Such a 'necessitarian' Russellian Monism might in fact dodge the exclusion argument. Whether or not the base is phenomenal or protophenomenal, if the relationship between the [dispositional] and phenomenal features of the base is intimate enough—and metaphysical necessitation from the phenomenal to the [dispositional] probably qualifies—the exclusion argument doesn't succeed.[58]

In fact, metaphysical necessitation from the (proto)phenomenal to the dispositional (which corresponds to (ii)) would only partly qualify as supplying the requisite intimacy, that is, as ruling out all swapped- and absent-quiddity scenarios as metaphysically impossible. Such metaphysical necessitation would rule out some swapped-quiddity scenarios as impossible. For example, such necessitation would guarantee that if (proto)phenomenal quiddity R actually grounds negative charge, as in w1, then there is no possible world such as w2, in which R is instantiated but does not ground negative charge. But metaphysical necessitation from the (proto)phenomenal to the dispositional is consistent with the claim that absent-quiddity scenarios, in which basic dispositional properties lack categorical grounding altogether, are possible. The latter possibility claim might well suffice for the purposes of Howell's integration-failure argument. Also, there

[57] We assume, with Howell (2015, p. 36), that the relevant sort of possibility here is metaphysical.

[58] Howell 2015, pp. 35-36. We replaced "causal" with "[dispositional]" for reasons discussed in the previous section.

are swapped-quiddity scenarios that are consistent with metaphysical necessitation from the (proto)phenomenal to the dispositional. For example, suppose R actually grounds negative charge, as in w1. That supposition is consistent with the claim that w3, in which either R or G grounds negative charge, is metaphysically possible, even given metaphysical necessitation from the (proto)phenomenal to the dispositional.

So, to completely dodge Howell's (or any) integration-failure argument by invoking necessitarianism, Russellian monists would have to endorse metaphysical necessitation in both directions, from the (proto)phenomenal to the dispositional and *vice versa*. That is, they would have to hold not only (ii) but (i) as well. The conjunction of (i) and (ii) would rule out all swapped- and absent-quiddity scenarios as metaphysically impossible. Because the necessitarian response to the integration-failure argument requires that the Russellian monist adopt both (i) and (ii), we will use the term "necessitarian Russellian monism" to refer to that biconditional form of the view.[59]

Let us now turn to Howell's objections to the necessitarian strategy. He presents two. The first is that adopting necessitarianism would conflict with the Russellian monist's "acceptance of…zombie-style conceivability arguments that pushed her to Russellian Monism in the first place."[60] That concern is natural enough. Necessitarian Russellian monism is incompatible with a premise those arguments invoke: the premise that zombie

[59] One might (but need not) base such a biconditional necessitarianism on an identity theory, according to which basic dispositional properties are identical to (proto)phenomenal properties. Such an identity theory has been developed partly on the basis of the Martin-Heil "powerful-qualities" view (Heil 2003, Martin 2008). See Gundersen 2015, Pereboom 2016, n.d., Morris n.d. Cf. Coleman 2015, Carruth 2015.

[60] Howell 2015, pp. 36-37.

worlds—consciousness-free worlds that are physical and functional duplicates of the actual world—are metaphysically possible.[61]

However, it is not really a problem that necessitarian Russellian monists must reject zombie worlds as impossible. Even *non*-necessitarian Russellian monists might well reject zombie worlds as impossible.[62] By definition, a zombie world would duplicate the actual world in all physical respects. That might (or might not) imply duplicating the actual categorical bases for all dispositional properties—bases that Russellian monists identify as consciousness-constituting (proto)phenomenal quiddities.[63] If there is that implication then, on Russellian monism, it would follow that a complete physical duplicate of the actual world would have to contain consciousness, contrary to the definition of a zombie world.

More generally, Howell's first objection to necessitarian Russellian monism oversimplifies the Russellian monist's attitude toward zombie-style conceivability

[61] This incompatibility can easily be demonstrated. Any physical and functional duplicate of the actual world is also a dispositional duplicate of the actual world. By the *necessitarian* part of necessitarian Russellian monism, there is a necessary entailment from the dispositional to the (proto)phenomenal. By the *Russellian monist* part of the theory, the macrophenomenal consists in the (proto)phenomenal, perhaps structured by the dispositional. Therefore, on necessitarian Russellian monism, any physical and functional duplicate of the actual world is a macrophenomenal duplicate of the actual world—a result that contradicts the premise that a zombie world is metaphysically possible.

[62] Chalmers 2003, Alter and Nagasawa 2012,

[63] Whether there is this implication depends on how the physical is construed. For example, on broad construals (such as Stoljar's (2001) "object-based conception") the physical includes categorical bases of basic physical dispositions, whereas on narrow construals (such as Stoljar's (2001) "theory-based conception") the physical does not include those categorical bases.

arguments. She takes those arguments to threaten traditional physicalism. But she need not endorse all versions of them. In particular, she need not endorse versions that invoke the premise that zombie worlds are metaphysically possible. She would reject those versions if she holds that duplicating all actual physical properties entails duplicating (proto)phenomenal quiddities. She would nevertheless accept the following amended premise: zombie worlds are metaphysically possible if (*per impossible* perhaps) there are no (proto)phenomenal quiddities. A zombie-style conceivability argument that invoked that amended premise would still threaten traditional physicalist views, in so far as those views entail that zombie worlds are metaphysically impossible even assuming that there are no (proto)phenomenal quiddities. So, contrary to Howell's claim, rejecting zombie worlds as metaphysically impossible is perfectly consistent with the Russellian monist's attitude toward zombie-style conceivability arguments.[64]

Howell's second objection to the necessitarian response is that adopting necessitarianism would undercut the Russellian monist's claim that her theory has advantages over traditional theories. This objection too can be explained in relation to the Russellian monist's attitude towards zombie-style conceivability arguments. Again, the necessitarian Russellian monist holds that zombie worlds are metaphysically impossible. Why, then, does her theory have any theoretical advantages over more traditional theories that make the same claim? If she argues that zombie worlds are only prima facie and not

[64] Notably, significant formulations of zombie-style conceivability arguments reflect the subtleties of the Russellian monist's attitude. For example, instead of "Zombie worlds are metaphysically possible," some formulate the relevant premise as "Zombie worlds are metaphysically possible or Russellian monism is true" (or something similar). Chalmers 2002, 2010, ch. 6, Pereboom 2011.

ideally conceivable, "then she appears to be making the same sort of move as the type A physicalist with no more plausibility."[65] If instead she posits "necessities that hold despite conceivability"—e.g., if she maintains that zombie worlds are metaphysically impossible despite being ideally conceivable—then "she has to allow the same answer for the type B physicalist and the property dualist."[66] Thus, Howell concludes, "necessitarian Russellian Monism might be conceptually coherent, but it is unmotivated."[67]

Howell's second objection is partly correct. It is not clear that Russellian monism provides distinctive reasons for positing "necessities that hold despite conceivability," that is, reasons that type-B physicalists and (necessitarian) dualists could not avail themselves of. It is also true that the Russellian monist who denies that zombie worlds are ideally

[65] Howell 2015, p. 37. Roughly, prima facie conceivability is conceivability at first glance, and ideal conceivability is conceivability on ideal rational reflection. For a detailed discussion, see Chalmers 2002. The alphabetic taxonomy comes from Chalmers (2003). Type-A physicalism says roughly that all phenomenal truths are a priori entailed by the complete physical truth (and thus zombie worlds are not ideally conceivable). Type-B physicalism says roughly that though some phenomenal truths are not a priori entailed by the complete physical truth (and thus zombie worlds are ideally conceivable), all phenomenal truths are nevertheless metaphysically necessitated by the complete physical truth.

[66] Howell 2015, p. 37. That answer sounds more suited to type-B materialism than to property dualism. Property dualists typically accept the claim that zombie worlds are metaphysically possible (Gertler 2007). But they need not, and Howell sees an advantage in their rejecting that metaphysical-possibility claim: "Necessary connections between the phenomenal and the physical could help the property dualist avoid epiphenomenalism since it is difficult to make a case for competition between properties that necessitate one another" (Howell 2015, p. 37, n. 28).

[67] Howell 2015, p. 37.

conceivable makes "the same sort of move as the type A physicalist." But Howell adds "and with no more plausibility," and that addition is unwarranted. The Russellian monist who denies that zombie worlds are ideally conceivable—call that view *type-A Russellian monism*—and the type-A materialist differ markedly in how they justify the claim that zombie worlds are not ideally conceivable. Arguably, the type-A Russellian monist's justification is more plausible than the type-A materialist's, contrary to Howell's added claim.

Consider a prototypical version of type-A materialism: analytic functionalism.[68] On this view, the conclusion that zombie worlds are not ideally conceivable traces to a doctrine about the meaning of phenomenal terms, such as "consciousness" and "pain": the doctrine that such terms can be functionally analyzed, that is, that such terms can be fully analyzed in functional terms alone. Analytic functionalists use that doctrine to argue that, because the actual world contains consciousness, so must any functional duplicate of the actual world. Because zombie worlds are functional duplicates of the actual world by definition, it follows that the zombie worlds are not ideally conceivable.

But the doctrine that phenomenal terms are functionally analyzable is widely regarded as implausible. Not all type-A materialists accept that particular doctrine. But all type-A materialists are committed to a doctrine that is also widely regarded as implausible: the doctrine that all macrophenomenal truths are a priori entailed by the complete dispositional (or the complete structural-and-dynamic) truth.[69] Type-A Russellian monists are not committed to either doctrine. They are committed to the claim that all

[68] Armstrong 1968, Lewis 1972.

[69] Chalmers 1996, 2003, 2010.

macrophenomenal truths are a priori entailed by a conjunction of the complete (proto)phenomenal quiddistic truth and the complete dispositional (or the complete structural-and-dynamic) truth. But the latter a priori-entailment claim is not widely regarded as implausible. And it is far from clear that it *is* implausible.

Even so, one might suspect that, at the end of the day, type-A Russellian monism is hardly more plausible than type-A materialism. After all, on either view zombie worlds are not ideally conceivable. That inconceivability claim might seem counterintuitive regardless of whether it is driven by any doctrine that is specific to type-A materialism, such as the doctrine that phenomenal terms can functionally analyzed.

But it is unclear how much weight to put on that intuition. To be well positioned to assess the claim that zombie worlds are ideally conceivable, or to trust the intuition that they are, we would need to have at least a fairly good understanding of all properties involved in the zombie-world hypothesis, including properties that are instantiated in the actual world but not in zombie worlds. Yet for Russellian monists, who take (proto)phenomenal quiddities to ground basic physical dispositions, the properties involved in that hypothesis include protophenomenal or microphenomenal properties. And our ignorance with respect to both of the latter sorts of property is considerable. Arguably, from the Russellian monist's perspective, we do not understand either sort well enough to be well positioned to assess the claim, or trust the intuition, that zombie worlds are ideally conceivable.

Consider protophenomenal properties. These are described as non-phenomenal, non-structural properties that play three roles: grounding basic dispositional properties described by fundamental physics; constituting macrophenomenal properties, perhaps

when arranged in the right structure; and being such that there is an a priori entailment from protophenomenal and dispositional truths to macrophenomenal truths. That description leaves much about protophenomenal properties undetermined. On most versions of Russellian panprotopsychism, we know nothing positive about what they are like intrinsically. Nor do we know what enables them to play the roles they are supposed to play. They are theoretical posits, with a nature that is left largely unspecified.[70]

Our understanding of the microphenomenal properties posited by Russellian panpsychists is little better. We do not know how they combine to constitute macrophenomenal properties.[71] Nor do we know much about the natures of microphenomenal properties. For example, consider the microphenomenal properties that, on Russellian panpsychism, constitute what it is like to be a quark. Those might well differ radically from any macrophenomenal properties with which we are familiar—despite both sorts of property being phenomenal. Indeed, it would be surprising if that were not the case.[72]

Given the extent of our ignorance regarding protophenomenal and microphenomenal properties, the Russellian monist could reasonably argue that we are not well positioned to assess the claim, or trust the intuition, that zombie worlds are ideally conceivable. She could argue that our ignorance concerning protophenomenal and microphenomenal properties explains why, even if there are a priori connections between

[70] Chalmers 2016, Pereboom 2011. However, see Pereboom 2011, Coleman 2015 for proposals about their nature.

[71] This is part of what is known as the combination problem (Seager 1995, Chalmers 2016, Coleman 2016).

[72] Chalmers 1996, pp. 293-97, Rosenberg 2004, p. 95.

the dispositional and the quiddistic, we would not recognize them. Perhaps, she could argue, if we come to better understand protophenomenal or microphenomenal properties, zombie worlds will then seem no more conceivable than worlds containing round squares.

Howell might respond that the foregoing ignorance-based argument would license the type-A materialist to give a parallel ignorance-based argument. However, our ignorance-based argument appeals specifically to (proto)phenomenal quiddistic ignorance: ignorance with respect to information about (proto)phenomenal quiddities. It is this, the Russellian monist can claim, that might make it seem as though zombie worlds are ideally conceivable even if they are not. The type-A materialist, who does not posit (proto)phenomenal quiddities, cannot make that move.[73] More generally, Russellian monism has quiddistic resources that traditional views lack, and this undermines the parity-of-reasons arguments Howell gives.[74]

---

[73] Type-A materialists might try to attribute the apparent conceivability of zombie worlds to ignorance of yet-to-be-discovered non-quiddistic truths (Stoljar 2006). However, Chalmers (2003) and Alter (2016) argue that any such effort is bound to fail. Even so, the dialectical force of the Russellian monist's ignorance-based argument should not be overstated. Russellian monism is supposed to provide a framework in which the hard problem of consciousness can be investigated more fruitfully than on traditional physicalism and traditional dualism. The plausibility of that claim is weakened to the extent that the Russellian monist's response to zombie-style conceivability arguments relies on the hope that were we to better understand the nature of protophenomenal or microphenomenal properties, we would see why zombie worlds are inconceivable. Also, we should note that some versions of Russellian monism rely less than others on quiddistic ignorance. For a version that relies comparatively little on such ignorance see, for example, Coleman 2015, 2016.

[74] There is a closely related argument that deserves mention. The combination problem (see above, fn. 70) has been formulated as a zombie-style conceivability argument (Goff 2009, Chalmers 2013) against certain

Thus, a necessitarian response to the integration-failure argument is not ruled out by the two objections Howell adduces against it.[75]

Conclusion

At the outset, we raised two questions: must Russellian monists accept the contingency thesis? What are its implications for their view? We then examined answers to those questions supplied by the integration-failure argument. That argument begins with the premise that Russellian monism entails the contingency thesis, on which swapped- and absent-quiddity scenarios are possible; infers that, were Russellian monism true, consciousness would be physically inefficacious; and concludes that on Russellian monism consciousness is not adequately integrated into nature. That conclusion would have serious implications for Russellian monism. Russellian monists claim as a main advantage of their theory that on it consciousness is adequately integrated into nature. The integration-failure

versions of Russellian monism. On that formulation, it is claimed, a minimal physical duplicate of the actual world might lack consciousness even if enriched by any arrangement of (proto)phenomenal quiddities that still lacks consciousness. We cannot address that argument here (but see Coleman 2016, Goff 2016).

[75] *Should* Russellian monists be necessitarians? This is another issue, which might turn on complex matters such as the nature of causal laws and the relationship between causal laws and causal powers. We plan to address this complex issue elsewhere. However, we will note that Russellian monists would appear to have at least one good reason to be necessitarians. If Russellian monists reject necessitarianism, then they will likely have to posit contingent grounding laws—and that raises the concern that such laws will seem just as arbitrary as the interactionist dualist's psychophysical laws are widely thought to be (cf. Pereboom 2011, p. 115).

argument, if sound, would undermine their claim. Russellian monism would then be left in the same boat as the traditional dualist theories it is supposed to supersede.

However, we examined two versions of the integration-failure argument, a generic version and Howell's version, and we argued that neither succeeds. Indeed, we argued that no version of the argument could succeed. If our arguments are sound, then the questions we raised at the outset can be answered as follows. Russellian monism is consistent with a contingency thesis framed in terms of metaphysical possibility. But such a metaphysical contingency thesis does not entail that consciousness lacks physical efficacy. That physical inefficacy claim might well follow from a contingency thesis framed in terms of nomological possibility. But the Russellian monist can (and should) reject such a nomological contingency thesis. On either interpretation, the contingency thesis is no threat to Russellian monism's touted integrationist virtues.

Further, we argued, Russellian monists could (though need not) reject not only the nomological contingency thesis but the metaphysical contingency thesis as well. In other words, they could adopt a necessitarian view on which swapped- and absent-quiddity scenarios are metaphysically impossible. Howell anticipates that response and objects that necessitarianism is not a viable option for Russellian monists. However, we argued, his objection is based on misunderstandings about what Russellian monism says and how it is motivated. *Contra* Howell, necessitarianism is a viable option for Russellian monists.

We do not claim to have fully vindicated the claim that Russellian monism is a breakthrough theory of consciousness and its place in nature. But if our arguments are sound, then that claim is safeguarded from threats stemming from the alleged contingency

of the relationship between (proto)phenomenal quiddities and the dispositional properties that, according to Russellian monism, they ground.[76]

References

Alter, T. 2016. The structure and dynamics argument against materialism. *Noûs* 50: 794-815.

Alter, T., and Y. Nagasawa. 2012. What is Russellian monism? *Journal of Consciousness Studies* 19: 67-95.

Armstrong, D. 1968. *A Materialist Theory of Mind*. London: Routledge and Kegan Paul.

Bennett, K. 2003 Why the exclusion problem seems intractable and how, just maybe, to tract it. *Noûs* 37: 471–97.

[76] We presented this paper at The University of Alabama, University of Arizona, The 2019 Central Division Meeting of the American Philosophical Association, *Russellian Monism: Time for the Details* at the Central European University, and Charles University Prague. We thank those who attended for helpful discussions. For further helpful comments and discussions, we wish to thank Chris Brown, David Chalmers, Philip Goff, Eric Hiddleston, Amy Kind, Adam Pautz, Derk Pereboom, both anonymous reviewers for *Noûs,* and most especially Robert Howell.

Burge, T. 2007. Postscript to "Mind-body causation and explanatory practice". In his *Foundations of Mind,* New York: Oxford University Press: 363-82.

Carruth, A. 2015. Powerful qualities, zombies and inconceivability. *Philosophical Quarterly* 66 (262): 25-46.

Chalmers, D. J. 2016. The combination problem for panpsychism. G. Brüntrop and L. Jaskolla (eds.) *Panpsychism*. New York: Oxford University Press: 179-214.

Chalmers, D. J. 2013. Panpsychism and panprotopsychism. Amherst Lecture in Philosophy: http://www.amherstlecture.org/index.html. Reprinted in T. Alter and Y. Nagasawa (eds.) *Consciousness in the Physical World: Perspectives on Russellian Monism.* New York: Oxford University Press, 2015: 246-76.

Chalmers, D. J. 2012. *Constructing the World.* New York: Oxford University Press.

Chalmers, D. J. 2010. *The Character of Consciousness.* New York: Oxford University Press.

Chalmers, D. J. 2003. Consciousness and its place in nature. In S. Stich and T. Warfield (eds.) *Guide to the Philosophy of Mind*. Cambridge: Blackwell. Reprinted in D. J. Chalmers (ed.) *Philosophy of Mind: Classical and Contemporary Readings.* New York: Oxford University Press, 2002: 247-72.

Chalmers, D. J. 2002. Does conceivability entail possibility? In T. Gendler and J. Hawthorne (eds.) *Conceivability and Possibility*. New York: Oxford University Press: 145–200.

Chalmers, D. J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

Churchland, P. S. 1986. *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. Cambridge, MA: MIT Press.

Churchland, P. M. 1984/1988. *Matter and Consciousness,* revised edition. Cambridge, MA: MIT Press.

Coleman, S. 2016. Panpsychism and neutral monism: How to make up one's mind. In G. Brüntrop and L. Jaskolla (eds.) *Panpsychism: Contemporary Perspectives*. New York: Oxford University Press: 249-282.

Coleman, S. 2015. Neuro-Cosmology. In P. Coates and S. Coleman (eds.) *Phenomenal Qualities: Sense, Perception, and Consciousness*. Oxford: Oxford University Press: 66-102.

Ehring, D. 2011. *Tropes: Properties, Objects, and Mental Causation*. Oxford: Oxford University Press.

Gertler, B. 2007. In defense of mind-body dualism. In J. Feinberg and R. Shafer-Landau (eds.) *Reason and Responsibility*, 13th edition. Boston: Wadsworth: 285–97.

Goff, P. 2017. *Consciousness and Fundamental Reality*. New York: Oxford University Press.

Goff, P. 2016. The phenomenal bonding solution to the combination problem. In G. Brüntrop and L. Jaskolla (eds.) *Panpsychism: Contemporary Perspectives*. New York: Oxford University Press: 283-304.

Goff, P. 2015. Against constitutive Russellian monism. In T. Alter and Y. Nagasawa (eds.) *Russellian Monism*. New York: Oxford University Press: 370-400.

Goff, P. 2010. Ghosts and sparse properties: Why the physicalist has more to fear from ghosts than zombies. *Philosophy and Phenomenological Research* 81: 119-39.

Goff, P., 2009, Why panpsychism doesn't help us explain consciousness," *Dialectica* 63: 289-311.

Gundersen, S., 2015, Russellian monism and epistemic pessimism. *SATS*, 16: 27-48.

Hawthorne, J., 2002, Advice for physicalists. *Philosophical Studies* 109:17-52.

Hart, W. D. 1988. *The Engines of the Soul*. Cambridge, U.K.: Cambridge University Press.

Heil, J. 2003. *From an Ontological Point of View*. New York: Oxford University Press.

Heil, J., and Mele, A. 1993. *Mental Causation*. Oxford: Oxford University Press.

Howell, R. J. 2015. The Russellian monist's problems with mental causation. *Philosophical Quarterly* 65: 22-39.

Howell, R. J. 2013. *Consciousness and the Limits of Objectivity: The Case for Subjective Physicalism*. Oxford: Oxford University Press.

Huxley, T. 1784. On the hypothesis that animals are automata, and its history. *Fortnightly Review* 95: 555-80. Reprinted in his *Collected Essays*. London: Macmillan, 1893.

Jackson, F. 1982: Epiphenomenal qualia. *Philosophical Quarterly* 32: 127-36.

Johnston, M. 1992. How to speak of the colors. *Philosophical Studies*
68: 221-63.

Kim, J. 2000. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge: MIT Press.

Kim, J. 1989. The myth of non-reductive materialism. *Proceedings and Addresses of the American Philosophical Association* 63**:** 31–47**.**

Kind, A. 2015. Pessimism about Russellian monism. In T. Alter and Y. Nagasawa (eds.) *Consciousness in the Physical World: Perspectives on Russellian Monism.* New York: Oxford University Press, 2015: 401-21.

Kripke, S. 1972. Naming and necessity. In Harman and Davidson 1972, pp. 253-355. Reprinted in G. Harman, D. Davidson (eds.) *The Semantics of Natural Language*. Dordrecht: Reidel.

Le Pore, E., and Loewer, B. 1987. Mind matters. *The Journal of Philosophy* 84: 630-642.

Lewis, D. 1972. Psychophysical and theoretical identifications. *Australasian Journal of Philosophy* 50: 249-58.

List, C. and Stoljar, D. 2017. Does the exclusion argument put any pressure on dualism? *Australasian Journal of Philosophy* 95: 96-108.

Martin, C. B. 2008. *The Mind in Nature*. Oxford: Clarendon Press.

McClelland, T. 2013. The neo-Russellian ignorance hypothesis: A hybrid account of phenomenal consciousness. *Journal of Consciousness Studies*, 20 (3–4):125–151.

Montero, B. G. 2015. Russellian physicalism. In T. Alter and Y. Nagasawa (eds.) *Consciousness in the Physical World: Perspectives on Russellian Monism.* New York: Oxford University Press: 209-223.

Montero, B. G. 2010. A Russellian response to the structural argument against physicalism. *Journal of Consciousness Studies* 17: 70-83.

Mørch, H. H. 2014. *Panpsychism and Causation: A New Argument and a Solution to the Combination Problem.* Dissertation, University of Oslo.

Morris, K. Russellian monism and the problem of consciousness. M.S.

Morris, K. 2016. Russellian physicalism, bare structure, and swapped inscrutables. *Journal of Consciousness Studies* 23: 180-98.

Papineau, D. 2002. *Thinking about Consciousness.* New York: Oxford University Press.

Pautz, A. A dilemma for Russellian monists about consciousness. Brown University ms.

Pereboom, D. 2016. Illusionism and anti-functionalism about phenomenal consciousness. *Journal of Consciousness Studies* 23: 172-85.

Pereboom, D. Russellian monism, qualitative inaccuracy, and the meta-problem of consciousness. M.S.

Pereboom, D. 2011. *Consciousness and the Prospects of Physicalism*. New York: Oxford University Press.

Robinson, W. S. 2018. Russellian monism and epiphenomenalism. *Pacific Philosophical Quarterly* 99: 100-17.

Robinson, D. 1993. Epiphenomenalism, laws, and properties. *Philosophical Studies* 69: 1-34.

Rorty, R. 1965. Mind-body identity, privacy, and categories. *Review of Metaphysics* 19: 24-54.

Rosenberg, G. 2004. *A Place for Consciousness: Probing the Deep Structure of the Natural World*. New York: Oxford University Press.

Seager, W. 1995. Consciousness, information, and panpsychism. *Journal of Consciousness Studies* 2: 272-88.

Shoemaker, S. 2007. *Physical Realization*. Oxford: Oxford University Press.

Smart, J. J. C. Sensations and brain processes. *Philosophical Review* 68: 141-56.

Stoljar, D. 2015. Russellian monism or Nagelian monism? In T. Alter and Y. Nagasawa (eds.) *Consciousness in the Physical World: Perspectives on Russellian Monism.* New York: Oxford University Press: 324-345.

Stoljar, D. 2006. *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness.* New York: Oxford University Press.

Stoljar, D. 2001. Two conceptions of the physical. *Philosophy and Phenomenological Research* 62: 253–81.

Strawson, G. 2006. Realistic monism: why physicalism entails panpsychism. In A. Freedman (ed.) *Consciousness and its Place in Nature.* Exeter: Imprint Academic: 3-31.

Wilson, J. M. 2011. Non-reductive realization and the powers-based subset strategy. *The Monist (Issue on powers)* 94: 121–54.

Yablo, S. 1992. Mental causation. *Philosophical Review* 101: 245-80.