

The Inconceivability Argument

Brian Cutter

University of Notre Dame

Abstract

This paper develops and defends a new argument against physicalist views of consciousness: the *inconceivability argument*. The argument has two main premises. First, it is not (ideally, positively) conceivable that phenomenal truths are grounded in physical truths. (For example, one cannot positively conceive of a situation in which someone has a vivid experience of pink wholly in virtue of the movements of colorless, insentient atoms.) Second, (ideal, positive) inconceivability is a guide to falsity. I attempt to show that the inconceivability argument enjoys a significant advantage over the more familiar *conceivability* argument. One can reasonably endorse the inconceivability argument without endorsing the conceivability argument, but one cannot reasonably endorse the conceivability argument without also endorsing the inconceivability argument.

1. Introduction

Perhaps the most famous argument against physicalism is the conceivability argument. Let “P” stand for the conjunction of all physical truths, and let “Q” stand for some phenomenal truth, such as “someone is conscious,” or “someone is having a phenomenally reddish experience.” From here, the conceivability argument can be stated as follows:

The Conceivability Argument

- (C1) $P \& \sim Q$ is conceivable.
- (C2) If $P \& \sim Q$ is conceivable, then $P \& \sim Q$ is possible.
- (C3) If $P \& \sim Q$ is possible, then physicalism is false.

(C4) Therefore, physicalism is false.

I am no physicalist, but I have never found the conceivability argument entirely convincing, nor have I found it to capture the core intuitions underlying my own resistance to physicalism. My aversion to physicalism stems less from considerations of conceivability and more from considerations of *inconceivability*. Crudely put, it seems to me inconceivable—not just humanly inconceivable, but inconceivable in principle—that a vivid experience of pink should be wholly constituted by the movements of colorless, insentient atoms, or that a sharp stinging pain should occur wholly in virtue of the mechanical interactions of material bodies. In my view, if a proposition is inconceivable in principle, or ideally inconceivable, this is a good reason to believe that it's false. It follows that we have good reason to reject the physicalist thesis that experiences are wholly constituted by, or fully grounded in, physical processes. In this paper, I develop this crude line of thought into an argument against physicalism: the *inconceivability argument*.

We'll let Φ stand for an arbitrary collection of physical truths, and once again, we'll let Q stand for some phenomenal truth. To a first approximation, the inconceivability argument can be stated as follows:

The Inconceivability Argument

(I1) It is inconceivable that Q holds wholly in virtue of Φ .

(I2) If it is inconceivable that Q holds wholly in virtue of Φ , then it is not the case that Q holds wholly in virtue of Φ .

(I3) If Q doesn't hold wholly in virtue of any collection of physical truths, then physicalism is false.

(I4) Therefore, physicalism is false.¹

The conceivability argument and the inconceivability argument have a similar structure. Each has two main premises, an *epistemic* premise and a *transition* premise, plus a fairly innocuous final premise that functions as a logical bridge to the anti-physicalist conclusion. In the conceivability argument, the epistemic premise is a claim about conceivability, and the transition premise moves from conceivability to possibility. In the inconceivability argument, the epistemic premise is a claim about inconceivability, and the transition premise moves from inconceivability to falsity. The transition premise of the conceivability argument is motivated by the idea that conceivability is a guide to possibility. One very strong formulation of this idea holds that whatever is conceivable is possible, but (C2) can also be derived from weaker principles linking conceivability to possibility. The transition premise of the inconceivability argument is motivated by the idea that inconceivability is a guide to falsity. One very strong formulation of this idea holds that reality is *thoroughly intelligible*, in the sense that whatever is true is conceivable, at least in principle. In other words, there are no aspects of reality of which no possible mind could form a positive conception. But as we'll see, (I2) can also be derived from weaker principles linking inconceivability to falsity.

Both arguments can be classified under the broad heading of “epistemic gap” arguments—arguments that reason from an epistemic gap to an ontological gap—but they rely on different kinds of epistemic gap. We can say that there is a *conceivable-variation* gap between the A-truths and the B-truths when one can conceive of the A-truths holding without the B-truths, and there is an *inconceivable-grounding* gap when one cannot conceive of the B truths

¹ In case the validity of the argument isn't immediately obvious, note that, since Φ was arbitrary, the first two premises jointly entail the universal generalization in the antecedent of (I3): for each collection of physical truths, it's not the case that Q holds wholly in virtue of those physical truths.

being wholly constituted by, or holding entirely in virtue of, the A-truths. The conceivability argument derives an ontological gap from the first type of epistemic gap, while the inconceivability argument derives an ontological gap from the second type. The two types of epistemic gap often go together, but in my view, they sometimes come apart, and they likely differ in their metaphysical implications. That the two kinds of gap often come apart should be especially plausible if one thinks that conceivable-variation epistemic gaps between low-level facts and the high-level facts they ground are common outside the phenomenal domain.² To illustrate, suppose the high-level fact that there is a wiggling worm (call this fact W) holds in virtue of certain more fundamental facts about the movements of particles (call these facts M). Here it's plausible that there is no inconceivable-grounding gap. It's perfectly conceivable that there should be a wiggling worm wholly in virtue of the fact that there are particles moving in certain ways. But some think (and I'm inclined to agree) that there *is* a conceivable-variation gap here. That is, although W is wholly grounded in M, we can conceive of M&~W. (We might do so by imagining a mereological nihilist scenario, in which there are particles moving in certain ways, but they don't compose anything. In this imagined scenario, there is no worm, hence no wiggling worm.³) Those who agree that M&~W is conceivable are likely to be suspicious of inferences from conceivable-variation epistemic gaps to ontological gaps. But since there is no inconceivable-grounding gap here, a case of this kind would not cast doubt on the corresponding inference from inconceivable-grounding gaps. More generally, it will be my contention that the inconceivability argument altogether sidesteps some of the most significant concerns with other more familiar epistemic-gap style arguments.

² See, e.g., Block and Stalnaker (1999), Yablo (2002), Schaffer (2017; forthcoming)

³ Cf. Schaffer (2017: 5-6).

I hope to convince the reader that the inconceivability argument is a formidable argument in its own right, but the main goal of this paper is to defend the *comparative* thesis that the inconceivability argument enjoys a significant advantage over the conceivability argument. To be more precise, the claim is not merely that the former is preferable to the latter in some respects, which might be balanced out by compensating advantages on the other side. Rather, the intended comparative thesis is analogous to a “weak dominance” claim: the inconceivability argument is more rationally compelling than the conceivability argument, in the sense that one can reasonably endorse the inconceivability argument without endorsing the conceivability argument, but one cannot reasonably endorse the conceivability argument without also endorsing the inconceivability argument. A bit more precisely, I shall defend two main claims:

- (i) One can reasonably reject or suspend judgment on the premises of the conceivability argument while endorsing the inconceivability argument. In particular, there are reasonable grounds for rejecting or suspending judgment on the transition premise of the conceivability argument, but these considerations pose no threat to the inconceivability argument.
- (ii) One cannot reasonably reject or suspend judgment on the premises of the inconceivability argument while endorsing the conceivability argument. For the premises of the inconceivability argument (as well as the underlying claims that I take to motivate those premises) are logically entailed by the assumptions involved in the conceivability argument, taken in conjunction with a few very modest and plausible further assumptions.

After clarifying some key concepts in §2, I defend (i) and (ii) in §§3-5. In §3, I give reasons for doubting some of the assumptions involved in the conceivability argument. In

particular, there are powerful reasons, which have been recently highlighted by Jonathan Schaffer (2017; forthcoming), to think that the conceptual possibilities outstrip the metaphysical possibilities in a way that makes trouble for the conceivability argument. But as we'll see, these considerations pose no threat to the inconceivability argument. Those who are moved by these considerations (as I am) to reject or suspend judgment on the premises of the conceivability argument can happily endorse the inconceivability argument. Relatedly, I'll argue that standard versions of the phenomenal concept strategy, an important physicalist strategy for resisting epistemic-gap arguments like the conceivability argument, do not help one to resist the inconceivability argument. In §§4-5, I make the case for claim (ii). Here I show that the assumptions involved in the conceivability argument, taken in conjunction with a few very modest and plausible assumptions, entail the premises of the inconceivability argument (as well as the underlying claims that I take to motivate those premises). Along the way, I shall also indicate why I think the premises of the inconceivability argument are independently plausible, apart from any prior commitment to the conceivability argument. I suggest that the first premise can be motivated by a plausible "general inconceivability thesis" to the effect that phenomenal truths can't be conceived to be wholly constituted by any collection of truths purely about *structure and dynamics*, and the second premise can be motivated by plausible principles linking truth and conceivability.

The central comparative thesis of this paper, if true, would be an important result. The conceivability argument is widely thought to be the most significant and most powerful anti-physicalist argument. (Another contender is the knowledge argument,⁴ but it's widely recognized that the knowledge argument and the conceivability argument are very closely related—closely

⁴ Jackson (1982).

enough that they probably stand or fall together.) So if we can establish that the inconceivability argument enjoys a significant advantage over the conceivability argument in the manner described above, this would substantially strengthen the overall philosophical case against physicalism.

2. Physicalism, (In)conceivability, and Grounding

Let us first clarify the key concepts involved in the inconceivability argument, specifically the concepts of physicalism, (in)conceivability, and the “wholly in virtue of” relation (along with its converse: full grounding).

“Physicalism”: I’ll define physicalism (about consciousness) as the thesis that phenomenal truths are fully grounded in physical truths. “Physical truths” can be understood as truths of the sort discovered by the physical sciences. Our purposes won’t require a definition of “physical truth” much more precise than this rough-and-ready gloss, but one further point of clarification is in order. Contemporary discussions of Russellian monism (or “Russellian physicalism”) commonly distinguish between physical truths in the *narrow* sense, which are restricted to “structural” truths of the sort revealed by the physical sciences, and physical truths in the *broad* sense, which include all narrowly physical truths, plus truths about the supposed intrinsic, quidditative underpinnings of the narrowly physical truths. Although the Russellian monist affirms that phenomenal truths are wholly grounded in *broadly* physical truths, she denies that they are wholly grounded in any set of *narrowly* physical truths, and this denial is what distinguishes her view from so-called “standard physicalism.” In this paper, “physicalism” refers to standard physicalism, the view that phenomenal truths are fully grounded in *narrowly* physical truths, and unqualified uses of expressions like “physical truth” and “physical property” should

be understood in the narrow sense.⁵ Thus, Russellian monism does not qualify as a form of physicalism on my usage, nor do I take the inconceivability argument to threaten Russellian monism.

“(In)conceivable”: Inconceivability can be straightforwardly analyzed in terms of conceivability: p is inconceivable iff p is not conceivable. For our purposes, the relevant notion of conceivability (for both the inconceivability argument and the conceivability argument) is what Chalmers (2002b) calls *ideal positive* conceivability. Positive conceivability is distinguished from *negative* conceivability, where p is negatively conceivable just in case it cannot be conclusively ruled out *a priori*. In contrast, to say that p is *positively* conceivable is to say that one can form a coherent positive conception of a situation in which p holds. Positively conceivability, then, is roughly the same thing as *imaginability*, at least if we understand imagination in a broad and conceptually infused sense, a sense that is not limited to purely sensory imagination (the capacity to mentally picture arrangements of low-level sensible qualities like colors and shapes). In the relevant sense, when one imagines that p , one constructs a coherent mental representation of a situation—that is, a more or less specific arrangement of objects and properties—in which p holds true, where the mental representations involved can draw from our full stock of conceptual resources, not just low-level sensory representations. To imagine or positively conceive a situation in which p holds true, it is not enough to imagine a configuration of objects and properties that happens to entail p . One must also coherently imagine the situation *as* one that verifies p ; as Chalmers puts it, it must be that “reflection on the situation reveals it as a situation in which [p]” (152).⁶ To say that p is *ideally* positively

⁵ For the terminology of “standard” physicalism, see Stoljar (2015). This corresponds to what Goff (2017) calls “pure physicalism” and what Chalmers (2015) calls “narrow physicalism.”

⁶ Cf. Yablo (1993).

conceivable is to say that p is positively conceivable under ideal rational reflection, or for an ideally rational mind—a mind in full possession of all the concepts involved in p , and without any memory or processing limitations that would prevent it from clearly and distinctly imagining all details that may be relevant to a p -verifying scenario.

Appeals to ideal conceivability and inconceivability invite an epistemic challenge. Since we are not idealized agents, it is tempting to think that we can only know what's humanly conceivable and inconceivable, not what's ideally conceivable and inconceivable. Now, surely there are many questions about ideal conceivability that are beyond our ken. But the idea that we can *never* know whether a proposition is ideally conceivable is not plausible at all. As Chalmers (2010: 155) writes, "Although we are non-ideal, we can know that it is not ideally conceivable that $0=1$, and that it is ideally conceivable that someone exists."

Officially, I'll treat conceiving as a relation to propositions: one conceives that p , where "that p " designates a proposition. I'll follow the convention of embedding a sentence in corner brackets to refer to the proposition expressed by that sentence in the context of utterance. Thus, "<S>" serves as an abbreviation for "the proposition that S." When discussing the conceivability of propositions, it is useful to individuate propositions finely—at least up to cognitive equivalence. In other words, it's useful to assume that, if it's not *a priori* that (S iff S*), then <S> \neq <S*>. According to this "Fregean" conception of propositions, <Hesperus is bright> \neq <Phosphorus is bright>, since it's not *a priori* that Hesperus is bright iff Phosphorus is bright.⁷

"Wholly in virtue of": In (I1)-(I3), the phrase "wholly in virtue of" is meant to express the converse of *metaphysical grounding*, a relation of constitutive dependence that underwrites

⁷ This assumption is convenient, but not strictly necessary to the argument. We could adopt a more coarse-grained conception of propositions, such as a Russellian conception or a possible-worlds conception, but then we'd need to relativize conceivability to guises or modes of presentation. This would introduce tedious complications to our discussion that I'd prefer to avoid.

metaphysical explanation, which has lately received a great deal of attention from metaphysicians.⁸ Officially, I'll take (full) grounding to be a relation that holds between a collection of facts (the grounding facts) and an individual fact (the grounded fact). For our purposes, we can take facts simply to be true propositions.

With these clarifications in view, (I1) amounts to the claim that psychophysical grounding claims of the form $\langle Q \text{ holds wholly in virtue of } \Phi \rangle$ are not ideally positively conceivable. It's reasonably clear that some grounding claims are conceivable in this sense, and some are not (even when the two "sides" of the grounding claim are jointly conceivable). For example, it's possible to conceive of a situation in which there is a party in virtue of the fact that teenagers are drinking, dancing, and listening to music,⁹ and it's possible to conceive of a situation in which there is a protest in virtue of the fact that there are students picketing, marching, and chanting. On the other hand, it's plausible that one cannot conceive of an object's being bright pink wholly in virtue of being spherical, and one cannot conceive of a person's being tall wholly in virtue of being smart. According to (I1), psychophysical grounding claims belong to the second class rather than the first.

3. Reasonable Doubts about the Conceivability Argument

The aim of this section is to defend part (i) of the main thesis of this paper, that is:

One can reasonably reject or suspend judgment on the premises of the conceivability argument while endorsing the inconceivability argument. In particular, there are reasonable grounds for rejecting or suspending judgment on the transition premise of the conceivability argument, but these considerations pose no threat to the inconceivability argument.

⁸ See especially the pioneering works of Fine (2001), Schaffer (2009), and Rosen (2010).

⁹ Cf. Goff (2017: 174).

As we've seen, the transition premise of the conceivability argument, (C2), reflects a commitment to the idea that conceivability is a guide to possibility. One strong and simple-minded expression of this idea says:

Simple-CP: If $\langle S \rangle$ is conceivable, then $\langle S \rangle$ is possible.¹⁰

But even proponents of the conceivability argument typically concede that Simple-CP is too strong. The most famous counterexamples are negations of Kripkean *a posteriori* necessities, such as “water \neq H₂O” and “Hesperus \neq Phosphorus.” Thus, proponents of the conceivability argument tend to agree that Simple-CP should be modified. Usually the modification takes the form of a restricted principle, according to which conceivable propositions that meet some further condition are possible. The restriction should exclude the Kripkean counterexamples to Simple-CP without excluding $P \& \sim Q$. The most influential principle along these lines, which is defended by Chalmers (2009b) and Jackson (1998), can be formulated as follows:

2D-CP: If $\langle S \rangle$ is conceivable, and the expressions that make up S have coinciding primary and secondary intensions, then $\langle S \rangle$ is possible.¹¹

Primary and secondary intensions are the two components of linguistic content within the semantic framework of epistemic two-dimensionalism. Without getting into the technical details of two-dimensional semantics,¹² the claim that (your token utterance of) an expression e has coinciding primary and secondary intensions is roughly equivalent to the claim that your use of e

¹⁰ The name “Simple-CP,” and the names for other CP principles below, are borrowed from Goff (2017).

¹¹ A closely related alternative principle defended by Goff (2017) is Transparency-CP: If $\langle S \rangle$ is conceivable, and the expressions that make up S express *transparent* concepts (i.e., concepts that fully reveal the natures of the entities they denote), then $\langle S \rangle$ is possible. As far as I can tell, everything critical I have to say about 2D-CP applies to Transparency-CP as well. In addition, the Transparency formulation of the conceivability argument requires the very strong premise that phenomenal concepts *fully* reveal the natures of their referents. Although I'm very sympathetic to a *partial* revelation thesis for phenomenal concepts, I am convinced by the argument in Pautz (2014: 179n) that they are not fully revelatory. For simplicity, I shall set aside the transparency formulation and focus on 2D-CP.

¹² For the technical details, see Chalmers (2004).

is not “twin-earthable.” We can say that your usage of an expression e is twin-earthable just in case you have a possible “twin” who uses e (or a counterpart expression) to refer to a different entity (property, object, kind, etc.) than you do, where a pair of possible individuals are twins just in case they are phenomenal and narrow functional duplicates, like Oscar and Twin Oscar in Putnam’s (1975) famous twin-earth thought experiment. The canonical example of a twin-earthable expression is “water,” which is used to refer to different natural kinds by Oscar and Twin Oscar. Now, it’s a familiar observation that the Kripkean counterexamples to Simple-CP all involve twin-earthable expressions, so they don’t constitute counterexamples to 2D-CP. Moreover, it can be plausibly argued that the expressions that make up $P \& \sim Q$ have coinciding primary and secondary intensions, so 2D-CP licenses the inference from the conceivability of $P \& \sim Q$ to its possibility.

Nonetheless, there are other counterexamples to 2D-CP, which have been vigorously pressed in recent work by Jonathan Schaffer (2017; forthcoming) (though Schaffer is certainly not the first to mention them). These counterexamples rely on the orthodox assumption that the notion of metaphysical necessity applies to paradigmatic “truths of metaphysics,” such as basic principles of composition, and the basic principles specifying the conditions under which sets exist. For example, it is widely held to be a necessary truth that, if Confucius exists, then his singleton exists. This necessary truth is thought to be underwritten by a grounding principle that holds of metaphysical necessity, perhaps a principle to the effect that, when something exists, its existence grounds the existence of its singleton. But, very plausibly, we can coherently conceive of a nominalist scenario in which there are no sets whatsoever, despite concrete reality being just as it actually is. Thus, claims like the following constitute plausible counterexamples to 2D-CP:

No-Set: For some x , there is no set whose only member is x .

On the assumptions just stated, this is conceivable but impossible, and it clearly satisfies the restriction clause of 2D-CP. After all, No-Set includes only logical and set-theoretic expressions, and these are paradigms of expressions with coinciding primary and secondary intensions.

Or consider another example, this time involving only the concrete realm. It is widely held that certain principles of composition are necessary. Hence, many will accept the necessity of a truth like: if there are some simples arranged chairwise, then those simples compose something. This necessary truth is thought to be underwritten by a grounding principle that holds of metaphysical necessity, a principle along the following lines: If there are some simples arranged chairwise (or, arranged in some more general way that subsumes chairwise arrangements as a special case), then the fact that simples are so arranged grounds the fact that their mereological sum exists. But, just as we can conceive of nominalist scenarios, it seems that we can conceive of mereological nihilist scenarios in which no composites exist, even while holding fixed the actual truths about the arrangements of simples. Thus, claims like the following constitute plausible counterexamples to 2D-CP:

No-Sum: There are some simples arranged chairwise, but they don't compose anything. On the assumptions just stated, No-Sum is conceivable but impossible. Moreover, it seems to satisfy the restriction clause of 2D-CP, since none of the expressions involved (logical and mereological expressions, plus the plural predicate "arranged chairwise") seem to be twin-earthable.

Examples like No-Set and No-Sum are significant for two reasons. First, they constitute plausible counterexamples to 2D-CP, which is the principle used to link conceivability to possibility in the most sophisticated versions of the conceivability argument. Second, each corresponds to a conceivable scenario that violates a plausible principle of grounding, thereby

lending support to the claim that at least *some* grounding principles, though they may hold of metaphysical necessity, are conceivably false. A physicalist can then respond to the conceivability argument by suggesting that the same goes for the principles of psychophysical grounding. This is precisely Schaffer's suggestion. According to Schaffer, when we conceive of a $P \& \sim Q$ scenario (say, a zombie scenario), we are conceiving of a scenario in which the metaphysically necessary principles of psychophysical grounding are falsified. This scenario is *conceptually* possible (i.e., coherently conceivable), since the basic principles of psychophysical grounding are not conceptually necessary. But these grounding principles are *metaphysically* necessary, so this scenario is not metaphysically possible. Now, if the psychophysical grounding principles were taken to be the *only* grounding principles that aren't conceptually necessary, Schaffer's view would look like special pleading on behalf of physicalism. But, as we've seen, there are other examples that fit the same pattern.

The counterexamples to 2D-CP relied on assumptions that might be questioned. The first case, for example, relied on the assumption that there are sets, and the second relied on the assumption that there are composites. These assumptions will be rejected by nominalists and mereological nihilists, respectively. But nominalism and nihilism don't really help in avoiding counterexamples to 2D-CP, at least given the standard assumption nominalism and nihilism are necessary if true. Given nominalism or nihilism, we could generate counterexamples to 2D-CP simply by reversing the examples, changing No-Set to "for some x , there is a set that includes x ," and changing No-Sum to "there are some simples arranged chairwise, and they compose something." These claims are, after all, pretty clearly conceivable, and they satisfy the restriction clause of 2D-CP. Let us therefore set aside the nominalist and nihilist response.

Alternatively, one might respond to the cases above by claiming that No-Set and No-Sum are possible after all (though false in actuality). In effect, this is to claim that the basic principles of set formation and composition are contingent. But this view is widely rejected. At the very least, it would be nice (for opponents of physicalism) to have an argument that didn't require such contentious metaphysical assumptions. Alternatively, one might respond to these cases by adopting a kind of metaphysical deflationism. The most straightforward version of deflationism holds that it is an analytic or conceptual truth that (for example) if Confucius exists, then his singleton does as well, or that if there are simples arranged chairwise, then those simples compose something. In my view, this kind of deflationism is implausible. I don't see how it could be analytic that, if some simples exist, and are arranged in a certain way, then *another* thing exists as well, a thing that is numerically distinct from each of the simples. As Schaffer (2017: 23) remarks, "the most that can be analytic are the conditions under which a concept is satisfied. It cannot be analytic which concrete individuals the world gives out." In any case, this kind of metaphysical deflationism (along with other more sophisticated brands of deflationism, such as the ontological anti-realism defended by Chalmers (2009a) and Hirsch (2010)) is controversial enough that it would be nice for opponents of physicalism to have an argument that doesn't rely on it.

Note that the examples above do nothing to threaten the *inconceivability* argument, and in particular, they do nothing to threaten the *inconceivability-falsity* linking principles that motivate (I2). We'll discuss such principles in detail in §5, but for now, let it suffice to note that the cases above aren't even of the right general form to threaten an *inconceivability-falsity* linking principle (or, relatedly, an *inconceivability-impossibility* linking principle). The whole point of these cases is to show that conceivability outstrips possibility. This is roughly the

opposite of what we'd need to refute a principle linking inconceivability to falsity or impossibility. To refute an inconceivability-impossibility link, we'd need to show that the metaphysically possible scenarios outstrip the conceivable scenarios, and to challenge an inconceivability-falsity link, we'd need to show in particular that the actual world is among the metaphysical possibilities that fall outside the domain of the conceivable. If we wanted to use the examples above to challenge the principles motivating (I2), we'd need to say, for example, that it is true yet somehow *inconceivable* that the existence of Confucius grounds the existence of his singleton, or that the existence of simples arranged chairwise grounds the existence of a chair. But this is not plausible at all.

For closely related reasons, the transition premise of the inconceivability argument doesn't seem to be threatened by standard versions of the phenomenal concept strategy, an important physicalist strategy for resisting the transition premise of the conceivability argument and other closely related anti-physicalist arguments. Phenomenal concept strategists offer a physicalist-friendly account of the nature of phenomenal concepts that is supposed to explain why phenomenal concepts (despite having purely physical referents) lack *a priori* conceptual connections to our physical concepts, and why we can coherently imagine the phenomenal truths varying independently of the physical truths (despite the impossibility of such independent variation).¹³ Now, an account of phenomenal concepts that expands the domain of the conceivable beyond the domain of the possible is suitable for resisting (C2), the transition premise of the conceivability argument. But it does not seem to be suitable for resisting (I2), the transition premise of the inconceivability argument. To resist (I2), we'd need to do roughly the opposite; we'd need to *contract* the domain of the conceivable to exclude certain real

¹³ The phenomenal concept strategy is defended in various forms by Loar (1990), Papineau (2002), Tye (2003), Balog (2012), and many others.

possibilities, including some actual truths. Perhaps a physicalist-friendly account of phenomenal concepts will one day be given that delivers this result. But standard versions of the phenomenal concept strategy are designed to undermine claims like (C2) by pushing toward a more *expansive* vision of what's conceivable, so they seem to be particularly unsuited to helping the physicalist resist (I2). (Instead, the liberal view of conceivability suggested by extant versions of the phenomenal concept strategy goes better with the denial of (I1). However, I'll argue in §4 that the liberal views of conceivability that are most naturally invoked to resist (I1) are extremely implausible on reflection.)

It might be suggested that the phenomenal concept strategy can help the physicalist resist (I2) if we accept something like the following principle linking conceivability to apriority:

In order to conceive <A grounds B>, one must (in addition to conceiving A and B), somehow perceive an *a priori* connection between A and B.

This principle, together with the absence of an *a priori* connection between Φ and Q (which phenomenal concept strategists purport to explain), would account for the inconceivability of < Φ grounds Q>. But this principle is undermined by the examples above. For example, this principle would yield the implausible result that it's inconceivable that sets are grounded in their members, or that sums are grounded in their parts.

The phenomenal-concept strategist might suggest that there is not merely a failure of *a priori* entailment from physical to phenomenal concepts, but a more radical kind of conceptual breach, a breach that does not exist between the concepts of sets and their members, or composites and their simple parts. This conceptual breach—the fact that physical and phenomenal concepts in some sense belong to utterly distinct conceptual families—might be

invoked to explain the inconceivability of $\langle \Phi \text{ grounds } Q \rangle$.¹⁴ Perhaps some explanation along these lines can be given. In any case, I allow that there must be *some* true psychological explanation of the inconceivability of $\langle \Phi \text{ grounds } Q \rangle$ in terms of the structures and roles of the relevant concepts. But giving a psychological explanation of inconceivability would not suffice on its own to undermine the metaphysical significance of inconceivability.¹⁵ Consider, by analogy, some other inconceivable grounding claims, like “something is pink wholly in virtue of being perfectly spherical,” or “a traffic jam occurs wholly in virtue of the fact that a bearded man utters ‘abra kadabra.’” (Such examples will be discussed further in §4.) Or consider some inconceivable non-grounding claims, like “something is red and green all over at the same time,” “something is longer than itself,” or “there are two distinct things, each a part of the other.” The fact that such claims are inconceivable is, very plausibly, a reason to think they are false (and, indeed, necessarily false). This is so even if there is a psychological explanation (as there presumably is) for why such claims are inconceivable.

For the reasons given above, it seems to me that there are significant reasons for doubting the conceivability-possibility linking principles that motivate the transition premise of the conceivability argument, but these considerations do not cast doubt on the premises of the inconceivability argument. One could, therefore, endorse the inconceivability argument while reasonably rejecting or suspending judgment on the conceivability argument.

4. The Epistemic Premise

I’ve argued that one can endorse the inconceivability argument while reasonably rejecting or suspending judgment on the premises of the conceivability argument. In this section

¹⁴ Thanks to an anonymous referee for this suggestion.

¹⁵ Of course, a similar point holds for psychological explanations of conceivability, as others have noted in debates about the conceivability argument (Chalmers 1999: 486-9).

and the next, I argue for the second part of the main thesis of this paper: one cannot reasonably reject or suspend judgment on the premises of the inconceivability argument while endorsing the conceivability argument. A bit more specifically, I show that the assumptions involved in the conceivability argument entail the premises of the inconceivability argument, at least given a few very modest and plausible assumptions that proponents of the conceivability argument ought to accept. Along the way, I'll also provide reasons for accepting the premises of the inconceivability argument that don't rely on a prior commitment to the conceivability argument. In this section, we'll focus on the epistemic premise, (I1). In the next section, we'll turn to the transition premise, (I2).

The epistemic premise of the inconceivability argument says that the following grounding claim is inconceivable:

Psychophysical Grounding: Q holds wholly in virtue of Φ .

Once again, Q is a phenomenal truth, like <someone is conscious> or <someone is having a phenomenally reddish experience>, and Φ is an arbitrary class of physical truths. It's not hard to show that typical proponents of the conceivability argument are committed to the inconceivability of phenomenal grounding. (As before, I'll work with a version of the conceivability argument that motivates (C2) with 2D-CP, though similar points would apply to those that rely on Simple-CP or other reasonable variations on these principles.)

Assume that Psychophysical Grounding *is* conceivable. From this assumption together with 2D-CP, we get the possibility of psychophysical grounding:

(\diamond PG) Possibly, Q holds wholly in virtue of Φ .

At least, (\diamond PG) follows on the assumption that the expressions in Psychophysical Grounding have coincident primary and secondary intensions. The 2D conceivability argument is already

committed to this assumption for the expressions in Q and Φ , since these are a subset of the expressions in $P \& \sim Q$, and it is clear that “wholly in virtue of” satisfies this condition as well. Now, following orthodoxy, I assume that (full) grounding entails metaphysical necessitation. In other words, I assume:

Ground Necessitarianism: Necessarily, if B holds wholly in virtue of A_1, \dots, A_n , then necessarily: if $A_1 \dots A_n$ hold, then B holds.

Ground Necessitarianism and (\diamond PG) jointly imply:

\diamond -Entailment: Possibly, it’s necessary that (if $\wedge\Phi$, then Q).¹⁶

Relying on the standard assumption that metaphysical modality conforms to the logic S5, this implies:

Entailment: Necessarily, if $\wedge\Phi$, then Q ,

which is inconsistent with the claim that $P \& \sim Q$ is possible, a direct implication of (C1) and (C2).

(Recall that P is the conjunction of all physical truths, so Φ consists only of conjuncts of P .)

Hence, proponents of the conceivability argument must deny our initial assumption that Psychophysical Grounding is conceivable (at least given S5 and Ground Necessitarianism—and I’m not aware of any proponents of the conceivability argument who would challenge these). In other words, they must accept (I1).¹⁷

To be fair, Ground Necessitarianism is somewhat controversial, despite its orthodox status.¹⁸ Still, the assumption is dialectically fair, not only because it is widely accepted and (I think) highly plausible, but also because the conceivability argument already presupposes

¹⁶ “ $\wedge\Phi$ ” stands for the conjunction of all members of the class Φ .

¹⁷ Closely related points are made by Marton (1998), Yablo (1999), and Sturgeon (2000), who give parody conceivability arguments *for* physicalism that rely on the premise that physicalism about consciousness is conceivable. However, as I argue below, the conceivability premise of the parody argument (understood as a claim about *positive* conceivability) is implausible (cf. Chalmers 2009b: 332).

¹⁸ For dissent from the orthodoxy, see Leuenberger (2014) and Skiles (2015).

something like Ground Necessitarianism, at least if its conclusion is meant to rule out physicalism as defined in §2. After all, if Ground Necessitarianism were false, then establishing the possibility of $P \& \sim Q$ wouldn't suffice to establish that Q isn't fully grounded physical truths, and hence wouldn't suffice to establish the falsity of physicalism. In other words, without Ground Necessitarianism (or at least the specific instance of Ground Necessitarianism invoked above), it would be unclear why we should accept premise (C3) of the conceivability argument. By contrast, proponents of the inconceivability argument can go either way, or simply remain neutral, on the question of Ground Necessitarianism. Thus, rejecting Ground Necessitarianism would only strengthen the case for my thesis that the inconceivability argument enjoys a significant advantage over the conceivability argument.

For the reasons just given, proponents of the conceivability argument seem to be committed to accepting (I1). However, in my view, (I1) is plausible independently of any prior commitment to the conceivability argument. In what follows, I provide independent motivation for (I1) and defend it against objections stemming from certain liberal views about conceivability.

4.1 In Defense of (I1)

According to premise (I1), instances of Psychophysical Grounding are not positively conceivable. For example, we cannot imagine someone having a vivid experience of bright pink wholly in virtue of the movements of colorless insentient atoms, or in virtue of having neuronal cells that exhibit a certain firing pattern, or in virtue of having parts that mechanically interact in certain ways or perform certain computations. Nor is this inability merely a consequence of our contingent psychological limitations—such as our inability to imagine scenarios in sufficiently rich detail, or our inability to distinctly imagine all the moving parts of a complex physical

system. So idealizing away these limitations would not make a difference to our inability to conceive of such a thing.

Or so I claim. To warm up intuitions, it will be helpful to consider, by way of contrast, a range of grounding claims that clearly *are* conceivable:

Traffic Jam: Some cars slow down in response to an accident, causing the cars behind them to slow down, which causes cars further back to slow down almost to a stop for miles back. Wholly in virtue of these facts, a traffic jam occurs.

Death: Bleeding profusely, the rate of oxygen transfer to Tom's cells slows dramatically. His respiratory and cardiac functions cease, and his central nervous system stops functioning entirely. Wholly in virtue of these facts, Tom dies.

In each of these cases, we can conceive of the putative derivative circumstance (e.g., the man dying, the traffic jam occurring), and the putative underlying facts, and moreover, we can conceive of the former being grounded in, or constitutively depending on, the latter. Even if cognitive limitations prevent us from forming a fully adequate positive conception of the scenarios above (for example, maybe we cannot clearly imagine the behavior of all the individual cars in a traffic jam), it's clear that these shortcomings could be overcome by idealization. There is little doubt that these scenarios are ideally positively conceivable.

Are psychophysical grounding claims similarly conceivable? Before I make the case for a negative answer to this question, let us briefly consider three liberal views of conceivability that might be invoked in defense of an affirmative answer. According to the first liberal view, just about anything can be conceived to ground anything, at least if the putative ground and grounded fact are themselves jointly conceivable. But this liberal view of conceivability is extremely

implausible on reflection. Perhaps just about anything can be conceived to *cause* anything. For example, it's conceivable that a traffic jam is caused (perhaps even "directly" caused) by a bearded man saying "abra kadabra." But the corresponding grounding claim seems clearly *not* to be conceivable.

Strange Traffic Jam: A bearded man says "abra kadabra." Wholly in virtue of this fact, there is a traffic jam.

Similarly, consider the following strange grounding claims:

Strange Death: A gust of wind blows out a candle. Wholly in virtue of this fact, a man on the other side of the world dies.

Pink Lump: A lump of matter is spherical. Wholly in virtue of this fact, the lump is bright pink.

If we interpret these as claims about *constitutive* dependence rather than mere *causal* dependence, they seem clearly not to be conceivable. At the very least, it seems clear that they are not *humanly* conceivable. Moreover, their inconceivability does not seem to be due to any conceptual limitations or cognitive limitations on our part, such as an inability to imagine scenarios in sufficiently rich detail. For example, the difficulty we have in positively conceiving Pink Lump has nothing to do with the complexity of the putative ground or grounded fact, nor do we lack any of the relevant concepts. So it's reasonable to suppose that the inconceivability of these propositions would be preserved under idealization.

Even if it is false that anything can be conceived to ground anything, one might think that anything (or nearly anything) can be conceived to ground phenomenal truths in particular. But this somewhat weaker liberal view of conceivability is also implausible on reflection. Consider the following variation on Pink Lump:

Pinkish Experience: A lump of matter is spherical. Wholly in virtue of this fact, the lump undergoes a vivid experience of bright pink.

It seems to me that the inconceivability of Pinkish Experience is just as self-evident as the inconceivability of Strange Traffic Jam, Strange Death, and Pink Lump.

Let's turn now to the positive intuitive case for (I1). The intuitive considerations that support (I1) can be usefully compared to those that support (C1). Assume for a moment a naïve Democritean view of physics, according to which the physical domain can be exhaustively characterized in terms of the motions of Democritean atoms—tiny colorless, insentient billiard balls. In this setting, (C1) can be supported by powerful intuitions like the following:

Conceivability Intuition: For any set of truths purely about the motions of Democritean atoms, one can in principle (positively) conceive of those motions occurring in the absence of a vivid experience of pink.

In the same way, (I1) can be supported by corresponding intuitions about *inconceivability*.

Inconceivability Intuition: For any set of truths purely about the motions of Democritean atoms, one cannot conceive of a vivid experience of pink being fully constituted by, or occurring wholly in virtue of, those motions.

To my mind, Conceivability Intuition and Inconceivability Intuition have roughly equal force, and both are highly plausible. Of course, physics has moved well beyond Democritus, but it's plausible that none of the novelties of modern physics do much to blunt the force of intuitions like these.

Let's see if we can reformulate these intuitions in a way that sits more comfortably with modern physical theory. Recall that, in §2, we defined physicalism as the view that phenomenal truths are grounded in physical truths, where “physical truths” here means *narrowly* physical truths

(truths of the sort revealed by the physical sciences) rather than *broadly* physical truths (which comprise narrowly physical truths as well as their supposed quidditative underpinnings). There I mentioned that narrowly physical truths are standardly assumed to be restricted to structural truths—or, more precisely, truths about “structure and dynamics”—and I shall follow this assumption in this paper. The idea that the physical sciences tell us only about structure and dynamics has roots in the work of Russell (1927) and Eddington (1928), and receives its canonical modern statement in the work of Chalmers (2002a). Concerning microphysics in particular, Chalmers writes:

A microphysical description of the world specifies a distribution of particles, fields, and waves in space and time. These basic systems are characterized by their spatiotemporal properties, and properties such as mass, charge, and quantum wavefunction state. These latter properties are ultimately defined in terms of spaces of states that have a certain abstract structure (e.g., the space of continuously varying real quantities, or of Hilbert space states), such that the states play a certain causal role with respect to other states. We can subsume spatiotemporal descriptions and descriptions in terms of properties in these formal spaces under the rubric of structural descriptions. The state of these systems can change over time in accord with dynamic principles defined over the relevant properties. The result is a description of the world in terms of its underlying spatiotemporal and formal structure, and dynamic evolution over this structure.

For Chalmers, the “structure” in “structure and dynamics” is supposed to be limited, roughly, to *spatial* (or spatiotemporal) structure, *causal/nomic* structure, and *mathematical/logical* structure. “Dynamics” concerns the manner in which these structural truths change over time. Hence, truths about structure and dynamics are, roughly, truths that specify the

spatial configuration of a system, the causal relations among its elements, the mathematical/logical properties of the system (for example, the mathematical structure of the abstract state spaces to which its elements are assigned locations), and the way these structures evolve over time.

This suggests a reformulation of Conceivability Intuition and Inconceivability Intuition. It's common to motivate (C1) with a modernized version of Conceivability Intuition along the following lines:¹⁹

General Conceivability Thesis: If A is a truth purely about structure and dynamics, then A&~Q is conceivable.

The General Conceivability Thesis can be motivated by a series of intuitive steps in something like the following manner: Starting with spatial structure, it's intuitively plausible that one can conceive of any purely *spatial* configuration holding in the absence of consciousness. From here, one might plausibly claim that the situation isn't relevantly changed by adding in dynamics, specifying how the spatial configuration of the system changes over time. In the next step, one could claim that the absence of consciousness remains conceivable no matter what information we add about the system's purely causal and mathematical structure. Through some such series of intuitive steps, we might conclude that the General Conceivability Thesis is true, and thus so is (C1). Reasoning along these lines may not provide *conclusive* support for the General Conceivability Thesis; I don't think we can be *certain* that there are no as-yet-unimagined structural truths about brain activity that can't be coherently imagined to hold without consciousness.²⁰ But reasoning along these lines nonetheless seems to provide substantial intuitive support for the General Conceivability Thesis.

¹⁹ See, e.g., Chalmers (2002a) and Alter (2016).

²⁰ The hypothesis that there are such structural truths is explored and defended in Stoljar (2006).

In the same way, we can motivate (I1) by appeal to a corresponding General

Inconceivability Thesis:

General Inconceivability Thesis: If A_1, \dots, A_n are truths purely about structure and dynamics, then it is not conceivable that Q holds wholly in virtue of A_1, \dots, A_n .

As with the General Conceivability Thesis, we can work up to the General Inconceivability Thesis by a series of intuitively plausible steps, with incremental structural enrichments at each step. Start with spatial structure. It's very plausible that, for any spatial configuration one might wish to describe, it's not conceivable that there should be an experience of bright pink wholly in virtue of that spatial configuration. Note that this is a fairly modest and plausible generalization from the inconceivability of the Pinkish Experience case, in which an experience of pink is grounded in a spherical spatial configuration.

From here, we can ask whether the situation changes in any relevant way after we add in *dynamics*—that is, after we specify how the spatial configuration of the system changes over time. Again, it is intuitively very plausible that this enrichment doesn't change the situation. Just as it's inconceivable that an experience of pink should be constituted by any purely spatial configuration, it seems to be inconceivable that an experience of pink should be wholly constituted by any dynamically evolving spatial configuration.

From here, we can ask whether the situation changes after we add information about causal structure—for example, information about how earlier spatial configurations causally impinge on later spatial configurations. Again, this doesn't seem to change the situation in any relevant way. If a dynamically evolving spatial arrangement on its own can't be conceived to fully ground an experience of pink, then the same plausibly holds for a dynamically evolving spatial configuration that exhibits a certain causal structure. Nor does the

situation seem to change if we enrich the structural description to include mathematical/logical structure—for example, by specifying that each component of the system occupies a certain location in a state space with a certain abstract mathematical structure. It's plausible that no purely abstract structural truths of this kind can be positively conceived to constitute or ground a vivid experience of bright pink.

Thus we come by gradual steps to the General Inconceivability Thesis, which implies (I1) (at least given our earlier stipulation that “physical truths” are to be understood as restricted to truths about structure and dynamics). Those who wish to reject the General Inconceivability Thesis (and hence, those who wish to reject (I1)) must reject one of the steps in the reasoning above. Now, surely some readers will find this line of reasoning more compelling than others. I think that a good dose of epistemic humility is called for when dealing with questions about ideal conceivability, so I would certainly concede that the reasoning above falls well short of *conclusively* establishing (I1). Still, I think this line of reasoning provides a very substantial degree of intuitive support for the General Inconceivability Thesis, roughly on a par with the intuitive support for the corresponding General Conceivability Thesis by which (C1) is justified.

One potential worry about the structure-and-dynamics argument for (I1) is that we may find it difficult to conceive of a system purely in terms of structure and dynamics. Quite apart from the question of grounding phenomenal states, it can be hard to imagine a thing that takes up space, with parts that spatially relate to one another and change over time in certain ways, without some sort of qualitative “filling” to give the thing substance. If we find it difficult even to conceive of a system that is pure structure and dynamics, then arguably our inability to conceive such a system *and* conceive of it as grounding phenomenal states should not give us

much confidence in (I1).²¹ To address this worry, we must distinguish (i) the task of conceiving of a situation in which some propositions purely about structure/dynamics hold true (e.g., a situation in which there is a cube next to a sphere) from (ii) the task of conceiving of a situation that is purely structural/dynamical—that is, a situation that is exhausted by facts about structure and dynamics. Whether or not we can (or can easily) perform the second task, we can surely perform the first. We can, for example, conceive of a situation in which a cube is next to a sphere, even if we cannot do so without imagining the shapes with some qualitative “filling,” such as Lockean solidity or grayness. It is task (i) that is most relevant to the structure-and-dynamics argument for the General Inconceivability Thesis. The latter implies (*inter alia*) that we cannot conceive of a situation in which someone has an experience of pink wholly in virtue of the fact that there is a cube next to a sphere. Our inability to conceive this grounding claim is not explained by any inability to conceive <there is a cube next to a sphere>, since the latter is readily conceivable.

It’s worth emphasizing, however, that even if one is not convinced by the structure-and-dynamics argument for (I1), this would not undermine the central project of this paper, which (recall) is to defend the *comparative* thesis that the inconceivability argument is strictly more rationally compelling than the conceivability argument. If I can convince the reader that the inconceivability argument is a compelling argument in its own right, even better. But if not, this wouldn’t threaten the main thesis.

It’s also worth noting that proponents of the conceivability argument are committed to the General Inconceivability Thesis, at least if they motivate (C1) by appeal to the General Conceivability Thesis mentioned above. A violation of the former would involve a

²¹ Thanks to an anonymous referee for raising this concern.

collection of structural truths X such that X conceivably grounds Q . Taking conceivability as a guide to possibility, it would follow that X possibly grounds Q , and thus (given Ground Necessitarianism and S5) that X necessitates Q . But proponents of the conceivability argument who use the General Conceivability Thesis to motivate (C1) can't accept this consequence. (They will hold that $X \& \sim Q$ is possible, since the General Conceivability Thesis implies that $X \& \sim Q$ is conceivable.) Thus, typical proponents of the conceivability argument are not only committed to accepting the *truth* of (I1). They are also committed to accepting the more basic principle that motivates (I1): the General Inconceivability Thesis.

5. The Transition Premise

Let's turn now to the transition premise of the inconceivability argument:

(I2) If it's inconceivable that Q holds wholly in virtue of Φ , then it's not the case that Q holds wholly in virtue of Φ .

In this section, I'll begin by arguing for the independent plausibility of (I2), giving an argument for (I2) that doesn't presuppose any commitment to the conceivability argument. Then I'll show that proponents of the conceivability argument have especially strong reason to accept (I2). A bit more precisely, I'll argue that they have especially strong reason to accept the principles by which I motivate (and which entail) (I2), since given their other commitments, it would be especially implausible (or, in some cases, inconsistent) for them to reject these principles. (There is no need to argue that proponents of the conceivability argument are committed to the *truth* of (I2). That much is trivial. (I2) is a material conditional whose consequent they accept.)

Just as (C2) is motivated by the idea that conceivability is a guide to possibility, (I2) is motivated by the idea that inconceivability is a guide to falsity. Again, a strong formulation of

this idea would be a universal principle to the effect that whatever is inconceivable is false. Or, equivalently:

Intelligibility: If p is true, then p is conceivable.

The Intelligibility thesis is so-called because it articulates the idea that reality is, in a certain well-defined sense, thoroughly *intelligible*. It says that if things are a certain way, then it is possible in principle for a mind to form a clear positive conception of a scenario in which things are that way, so there are no aspects of reality that are closed off to all possible minds.

I think Intelligibility is plausible. One might even argue that something like Intelligibility is presupposed by scientific inquiry, at least as a working hypothesis (or perhaps just the most ambitious forms of scientific inquiry, such as the search for a final “theory of everything”). However, there is one class of potential counterexamples that I don’t think we can confidently rule out. In particular, it seems to be an epistemic possibility that some fundamental properties are inconceivable, in the sense that even an idealized mind cannot positively conceive of their instantiation.²² Fortunately, inconceivable truths of this kind do not pose a serious threat to (I2). To motivate (I2), we don’t need anything as strong as Intelligibility in its unrestricted form. We can rely on *restricted* versions of Intelligibility that are compatible with the existence of fundamental inconceivable properties. In my view, the best motivation for (I2) comes from two principles, each of which amounts to a restricted form of Intelligibility:

Physical Intelligibility: If p is a physical truth, then p is conceivable.

Ground Intelligibility: If p is a grounding truth where “both sides” of p are conceivable, then p is conceivable. In other words, if we have a truth of the form $\langle B$ holds

²² This possibility is raised briefly in Chalmers (2002b: 186-7). See also Chalmers’ (2012: 351-2) discussion of the possibility of “ungraspable thick quiddities,” whose instantiation would arguably entail the falsity of Intelligibility.

wholly in virtue of A> such that A and B are individually and jointly conceivable, then <B holds wholly in virtue of A> is conceivable.²³

Physical Intelligibility is a restriction of Intelligibility to physical truths. Ground Intelligibility is a restriction of Intelligibility to grounding truths both sides of which are conceivable. From these two principles, (I2) follows. Ground Intelligibility implies that <Q holds wholly in virtue of Φ > is conceivable if Φ and Q are conceivable, and Physical Intelligibility guarantees that Φ is conceivable. Since there is no question about the conceivability of Q, we can conclude from here that (I2) is true.

Physical Intelligibility is very plausible. Recall that we are assuming that physical truths are restricted to structural truths. Even if there are fundamental truths that are somehow conceptually inaccessible to any possible mind, this doesn't seem to be the case for structural truths. Of course, there are inconceivable structural *propositions*—for example, those that describe incoherent spatial structures (e.g., spatial arrangements that violate triangle inequality), or those that describe incoherent mathematical structures (e.g., a continuous state space with finitely many points). But Physical Intelligibility is restricted to *truths*, and presumably no inconceivable structural propositions of this kind are true.

Apart from the general plausibility of Physical Intelligibility, proponents of the conceivability argument are clearly committed to it in any case. The first premise of the conceivability argument is that $P \& \sim Q$ is conceivable, so they can hardly deny that P itself is conceivable. And if P is conceivable, then Physical Intelligibility is true, for P comprises all physical truths.

²³ At the risk of pedantry: “A” officially ranges over collections of propositions rather than just individual propositions, so to conceive A is to jointly conceive its members, and for A to hold is for its members to hold jointly.

Ground Intelligibility is also very plausible. This thesis is not threatened by the idea that our world might include fundamental inconceivable features. What we would need to threaten Ground Intelligibility is a case with roughly following structure: There's a conceivable truth A, for example, <there are three pebbles sitting equidistant from one another>. And there is another conceivable truth B, which holds wholly in virtue of A. But this grounding truth—that B holds wholly in virtue of the fact that there are three pebbles sitting equidistant from one another—is inconceivable in principle. I think it's very implausible that there are truths of this kind. I am aware of no remotely plausible examples outside the phenomenal domain that exhibit this structure, and positing inconceivable grounding truths of this kind *only* for cases of psychophysical grounding looks like special pleading. (Note the contrast with the conceivability-possibility linking principles used to motivate (C2). In §3, we saw that there are plausible examples outside the phenomenal domain in which these principles are violated, so the physicalist can reject these without special pleading.)

Apart from the general plausibility of Ground Intelligibility, it would be especially awkward for proponents of the conceivability argument to reject it. In order to reject it, they would need to reject at least one of the following two principles:

Conceivability Excluded Middle: If A and B are (individually and jointly) conceivable, then either it's conceivable that A holds and grounds B or it's conceivable that A holds and doesn't ground B.

Strong Ground Necessitarianism: If A grounds B, then necessarily, if A holds, then A grounds B.

For, if we take conceivability as a guide to possibility (as proponents of the conceivability argument do), these two claims imply Ground Intelligibility. To see this, let <A grounds B> be

any grounding truth such that A and B are (individually and jointly) conceivable. By Conceivability Excluded Middle, one of the following is conceivable:

- (a) A holds, and grounds B.
- (b) A holds, and does not ground B.

The conceivability of (b) can be ruled out. Using conceivability as a guide to possibility, the conceivability of (b) would yield the possibility of (b), and the possibility of (b) together with Strong Ground Necessitarianism directly entails that A does not ground B, which contradicts our assumption that $\langle A \text{ grounds } B \rangle$ is true. Hence, (a) is conceivable, which establishes Ground Intelligibility.²⁴

It would be very bizarre to deny Conceivability Excluded Middle. Assuming A and B are conceivable, then if it's in principle impossible to conceive of A grounding B, it should be possible in principle to conceive of A holding *without* grounding B. Or, equivalently: if it's impossible in principle to conceive of A holding without grounding B, then it's surely possible to conceive of A holding *while* grounding B. (For example, if it's impossible to conceive of a scenario in which [a lump is pink wholly in virtue of the fact that it is spherical], then it should be possible to conceive of a scenario in which [a lump is spherical, but its sphericity doesn't ground its being pink]. Or, if it's impossible to conceive of a scenario in which [cars behave in a certain way, but this car behavior doesn't ground the occurrence of a traffic jam], then it should be possible to conceive of a scenario in which [cars behave in this way, and this car behavior *does* ground the occurrence of a traffic jam.]) Even a physicalist who rejects Ground Intelligibility (say, because she accepts that psychophysical grounding claims are inconceivable truths) will presumably accept Conceivability Excluded Middle.

²⁴ For simplicity, the argument here relies on Simple-CP. Of course, we can reach a similar conclusion with restricted forms of Simple-CP, such as 2D-CP, if we add a corresponding restriction to Ground Intelligibility.

Strong Ground Necessitarianism is also a very reasonable assumption, especially for those who already endorse the conceivability argument. Strong Ground Necessitarianism is a slight strengthening of Ground Necessitarianism. While those who reject Ground Necessitarianism will of course reject the stronger thesis, among those who accept Ground Necessitarianism, virtually no one disputes Strong Ground Necessitarianism.²⁵ Indeed, Strong Ground Necessitarianism is a straightforward consequence of Ground Necessitarianism together with the very widely accepted thesis that grounding is an “internal” relation in the following sense: necessarily, if A grounds B, then necessarily, if A holds and B holds, then A grounds B.²⁶ We’ve seen already that proponents of the conceivability argument are committed to Ground Necessitarianism, or something in the vicinity, so they will presumably be sympathetic to Strong Ground Necessitarianism as well.²⁷

Thus, not only are proponents of the conceivability argument trivially committed to the *truth* of (I2); they also have especially strong reasons to accept the underlying principles that motivate (I2): Physical Intelligibility and Ground Intelligibility. I’ve suggested that these principles are independently plausible, but they should be especially attractive to proponents of the conceivability argument. As we’ve seen, they are already committed to Physical

²⁵ One of the few exceptions is Litland (2015), who proposes a counterexample to Strong Ground Necessitarianism that is consistent with Ground Necessitarianism. His counterexample is a complicated and *recherché* case involving contingent self-reference. Even if the case is successful, it seems to me to be an isolated example that is far enough from the concerns of this paper that it can be harmlessly ignored.

²⁶ The internality of grounding is accepted by Rosen (2010), Bennett (2011), Fine (2012), deRosset (2013), Trogon (2013), and Dasgupta (2014), among many others.

²⁷ Another route to Strong Ground Necessitarianism goes by way of the attractive (to my mind) assumption that the most fundamental notion of grounding is what Fine (2012) calls “non-factive” grounding, which can hold among propositions independently of whether those propositions are true. We’ve been working with the factive notion of grounding, but as Fine observes, factive grounding can be analyzed in terms of non-factive grounding in a natural way: A factively grounds B =: A holds and A non-factively grounds B. From this definition together with the natural assumption that the non-factive grounding truths are non-contingent, Strong Ground Necessitarianism follows immediately.

Intelligibility, and denying Ground Intelligibility would be especially awkward for them, given their prior commitments to conceivability as a guide to possibility and Ground Necessitarianism.

6. Conclusion

The two key premises of the inconceivability argument, (I1) and (I2), are independently plausible: (I1) can be motivated by a plausible general inconceivability principle modeled after the principles used to motivate the epistemic premise of the conceivability argument, and (I2) can be motivated by plausible restricted Intelligibility principles (Physical Intelligibility and Ground Intelligibility). Apart from the independent plausibility of (I1) and (I2), I've argued that proponents of the conceivability argument are in any case committed to accepting them, and have especially strong reason to accept the underlying principles that motivate these premises as well. Conversely, I've argued that proponents of the inconceivability argument can reasonably reject or suspend judgment on premises of the conceivability argument. In particular, there are reasonable grounds for doubting the transition premise of the conceivability argument, grounds that pose no threat to the inconceivability argument. One can, therefore, reasonably endorse the inconceivability argument without endorsing the conceivability argument, but one cannot reasonably endorse the conceivability argument without also endorsing the inconceivability argument.

Acknowledgments

Thanks to David Chalmers, Peter Epstein, Neil Mehta, Boyd Millar, and Maarten Steenhagen for helpful comments on a presentation of an early draft of this paper. Thanks to Jon Litland for useful discussion. Finally, thanks to two anonymous referees for their exceptionally valuable comments.

References

- Alter, Torin (2016). The structure and dynamics argument against materialism. *Noûs*, 50(4), 794–815.
- Balog, Katalin (2012). In defense of the phenomenal concept strategy. *Philosophy and Phenomenological Research*, 84(1), 1–23.
- Bennett, Karen (2011). By our bootstraps. *Philosophical Perspectives*, 25(1), 27–41.
- Block, Ned, and Robert (1999). Conceptual analysis, dualism, and the explanatory gap. *Philosophical Review*, 108(1), 1–46.
- Chalmers, David (1999). Materialism and the Metaphysics of Modality. *Philosophy and Phenomenological Research*, 59(2), 473–96.
- Chalmers, David (2002a). Consciousness and its Place in Nature. In David Chalmers (Ed.), *Philosophy of Mind: Classical and Contemporary Readings* (247–272). Oxford University Press.
- Chalmers, David (2002b). Does Conceivability Entail Possibility? In T. S. Gendler & J. Hawthorne (Eds.), *Conceivability and Possibility* (145–200). Oxford University Press.
- Chalmers, David (2004). Epistemic Two-Dimensional Semantics. *Philosophical Studies*, 118(1–2), 153–226.
- Chalmers, David (2009a). Ontological Anti-Realism. In David Chalmers, David Manley, and Ryan Wasserman (Eds.), *Metametaphysics: New Essays on the Foundations of Ontology* (77–129). Oxford University Press.
- Chalmers, D. (2009b). The Two-Dimensionalist Argument against Materialism. In Brian McLaughlin, Ansgar Beckermann, and Sven Walter (Eds.), *Oxford Handbook to the Philosophy of Mind* (313–338). Oxford University Press.
- Chalmers, David (2010). *The Character of Consciousness*. Oxford University Press.
- Chalmers, David (2012). *Constructing the World*. Oxford University Press.
- Chalmers, David (2015). Panpsychism and Panprotopsyism. In Torin Alter and Yujin Nagasawa (Eds.), *Consciousness in a physical world: Perspectives on Russellian monism* (246–276). Oxford University Press.
- Dasgupta, Shamik (2014). The Possibility of Physicalism. *Journal of Philosophy*, 111(9–10), 557–592.
- deRosset, Louis (2013). Grounding Explanations. *Philosophers' Imprint*, 13, 1–26.
- Eddington, Arthur (1928). *The Nature of the Physical World*. Cambridge University Press.
- Fine, Kit (2012). Guide to Ground. In F. Correia & B. Schnieder (Eds.), *Metaphysical Grounding: Understanding the Structure of Reality* (37–80). Cambridge University Press.

- Goff, Philip (2017). *Consciousness and Fundamental Reality*. Oxford University Press.
- Hirsch, Eli (2010). *Quantifier Variance and Realism: Essays in Metaontology*. Oxford University Press.
- Jackson, Frank (1982). Epiphenomenal Qualia. *Philosophical Quarterly*, 32(April), 127–136.
- Jackson, Frank (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford University Press.
- Leuenberger, Stephan (2014). Grounding and Necessity. *Inquiry: An Interdisciplinary Journal of Philosophy*, 57(2), 151–174.
- Litland, Jon (2015). Grounding, Explanation, and the Limit of Internality. *Philosophical Review*, 124(4), 481–532.
- Loar, Brian (1990). Phenomenal States. *Philosophical Perspectives*, 4, 81–108.
- Marton, Peter (1998). Zombies versus Materialists: The Battle for Conceivability. *Southwest Philosophy Review*, 14(1), 131–138.
- Papineau, David (2002). *Thinking about Consciousness*. Oxford University Press.
- Pautz, Adam (2014). The Real Trouble with Armchair Arguments against Phenomenal Externalism. In Mark Sprevak & Jesper Kallestrup (Eds.), *New Waves in Philosophy of Mind* (153–181). Palgrave Macmillan.
- Putnam, Hilary (1975). The Meaning of “Meaning.” *Minnesota Studies in the Philosophy of Science*, 7(131), 131–193.
- Rosen, Gideon (2010). Metaphysical Dependence: Grounding and Reduction. In Bob Hale and Aviv Hoffman (Eds.), *Modality: Metaphysics, Logic, and Epistemology* (109–136). Oxford University Press.
- Russell, Bertrand (1927). *The Analysis of Matter*. Kegan Paul.
- Schaffer, J. (forthcoming). Ground Functionalism. *Oxford Studies in Philosophy of Mind*.
- Schaffer, Jonathan (2017). The Ground between the Gaps. *Philosophers’ Imprint*, 17, 1–26.
- Skiles, Alexander (2015). Against Grounding Necessitarianism. *Erkenntnis*, 80(4), 717–751.
- Stoljar, Daniel (2006). *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness*. Oxford University Press.
- Stoljar, D. (2015). Russellian monism or Nagelian monism? In T. Alter & Y. Nagasawa (Eds.), *Consciousness in a physical world: Perspectives on Russellian monism* (pp. 324–345). Oxford University Press.
- Sturgeon, Scott (2000). *Matters of Mind: Consciousness, Reason, and Nature*. Routledge.

Trogon, Kelly (2013). Grounding: Necessary or Contingent. *Pacific Philosophical Quarterly*, 94(4), 465–485.

Tye, Michael (2003). A Theory of Phenomenal Concepts. In Anthony O’Hear (Ed.), *Royal Institute of Philosophy Supplement* (91–105). Cambridge University Press.

Yablo, Stephen (1993). Is Conceivability a Guide to Possibility. *Philosophy and Phenomenological Research*, 53(1), 1–42.

Yablo, Stephen (1999). Concepts and Consciousness. *Philosophy and Phenomenological Research*, 59(2), 455–463.

Yablo, Stephen (2002). Coulda, Woulda, Shoulda. In Tamar Gendler & John Hawthorne (Eds.), *Conceivability and Possibility* (441–492). Oxford University Press.