

Physicalism, the Intelligibility Constraint, and the Myth of Structure and Function*

Gabriel Oak Rabin

Abstract: *The Intelligibility Constraint* on theories of the fundamental states that if the fundamental level is composed purely of Fs, then it must be intelligible how the Fs could generate the remainder of the world’s contents. A prominent line of argument (Alter [2016], Chalmers [1996, 2010b]) claims that physicalism is false precisely because it violates the intelligibility constraint. Physical fundamentalia can’t make consciousness intelligible. The argument has two key premises. First premise: intelligibility is a priori entailment. Second premise: there is no a priori route from the physical to consciousness. I argue that both premises are false. The relevant notion of intelligibility should not be cached out via a priori entailment. Some intelligible-making explanations are not a priori, and some a priori entailments do not yield intelligible connection. The claim that physical fundamentalia cannot a priori entail consciousness is based on the idea that the physical is “structural-functional”, while consciousness is not, and the structural-functional cannot a priori entail the non-structural-functional. The ability of the notion of “structure and function” to bear this dialectical burden is a myth. I argue by exhaustion against a variety of precisifications of “structure and function”.

1 Introduction

Physicalism claims that all the fundamental features of our world are physical.¹ At rock bottom, it’s just quarks, leptons, and bosons, with some

*Acknowledgements.

¹For more on this conception of physicalism, cf. Rabin [2020].

electromagnetic and nuclear forces between them. The rest of the world's contents are grounded in these basic physical building blocks. All the coffees, cars, and koalas arise from quarks, leptons, and bosons, appropriately arranged. Any theory of the fundamental (physicalism included) must satisfy *The Intelligibility Constraint*, which says that if the fundamental level is composed purely of Fs, then it must be intelligible how the Fs could ground the remainder of the world's contents.

Consciousness poses the most significant contemporary challenge to the physicalist picture. Some, most prominently David Chalmers (Chalmers [1996, 2010b]), argue that physicalism fails the intelligibility constraint. It is not intelligible how a world which, at the fundamental level, was purely physical, could at the same time contain consciousness. No matter how you configure the quarks and siblings, you won't get consciousness. More ingredients are required. Thus: there is more to the fundamental grounds of reality than the physical.

We'll call the argument that utilizes the intelligibility constraint against physicalism *The Intelligibility Argument Against Physicalism*.

The Intelligibility Argument Against Physicalism

(I1) If the fundamental is purely physical, then there is an intelligible route from the physical to consciousness.

(I2) There is no intelligible route from the physical to consciousness.

(IC) Therefore: the fundamental is not purely physical; physicalism is false.

(I1) is an instance of the intelligibility constraint, applied to physicalism and supplemented with the fact that the world contains conscious experiences. The focus of this paper is premise (I2), which says that there is no

intelligible route from the physical to consciousness. Why should we believe that this is so? An argument can be gleaned from the work of David Chalmers. It has two basic ideas. First idea: intelligibility is *a priori* entailment. Second idea: there is no *a priori* route from the physical to consciousness. These ideas can be combined into the *A Priori Argument for Unintelligibility*.

The *A Priori* Argument for Unintelligibility

(P1) Intelligibility is *a priori* entailment: There is an intelligible route from the Fs to the Ms if and only if the conditional “If F then M” is knowable *a priori*. (Where F and M are complete descriptions of the Fs and Ms).

(P2) No *a priori* route: The conditional “If P then C” is not knowable *a priori*, where P and C completely describe the physical and consciousness facts.

(C) There is no intelligible route from the physical to consciousness.

This argument is the focus of this paper. Going above and beyond the call of philosophical duty, I will argue that we should reject *both* premises. I reject (P1) because intelligibility should not be cashed out in terms of *a priori* entailment. I reject (P2) because it requires too much. It requires an unbridgeable-in-principle-no-matter-what epistemic gap. We should be skeptical about the existence of such a gap.

I now offer a brief roadmap of the paper. Section 2 (“Against Intelligibility as *A Priori* Entailment”) explores the notion of intelligible grounding and argues that (P1) is false. Section 3 (“From the Physical to Consciousness *A Priori*”) argues against (P2). That section consists of two parts.

First, I lay out some dialectical considerations, including the challenges from idealization and from historical precedent, to the inference from our current inability to see how consciousness might arise from physical processes to a lack of intelligibility in the idealized sense that is relevant to the intelligibility constraint. The reply to these challenges involves a distinction between structural-functional properties, which all physical properties are alleged to be, and non-structural-functional properties, including conscious experience. Sections 3.7-3.10 argue by exhaustion that none of the prominent accounts of the structural-functional can bear the philosophical burden the dualist requires.

2 Against Intelligibility as *A Priori* Entailment

2.1 What is Intelligibility?

The intelligibility constraint insists that the grounding of the non-fundamental in the fundamental be *intelligible*. What is intelligibility? Or, more accurately, what is intelligible *grounding*? (I use ‘ground’ as a label for the generative relationship between fundamentalia and the non-fundamentalia they generate.²) The idea is simple. If P intelligibly grounds Q, a reasonable and intelligent agent who knew that P was the case at the (more) fundamental level, would understand why Q was also the case at the less fundamental level. If P intelligibly grounds Q, there should be no mystery of why Q is the case at a world in which P.

Examples will help illustrate. First example: thermodynamics. The motion of molecules in a closed chamber intelligibly grounds the pressure and

²Two assumptions about how I treat ground are worth mentioning. First, I take the fundamental to consist of those entities that (a) ground everything else and (b) ground everything else (though for a complication cf. [Shumener\[this volume\]](#), Rabin [2020]: 4fn4, 29). Second, I take grounding to be a relation between facts.

temperature of the gas in that chamber. Once you understand how the molecules are moving, their kinetic energy, and the way they regularly strike the walls of the chamber (thereby exerting pressure), it's quite clear why the gas has the temperature and pressure it does. The mean velocity of the molecules does everything you'd expect from temperature, and the mean force exerted on the walls of the chamber by the particles does everything you'd want pressure to do.

Second example: intentionality. Descartes (Cottingham *et al.* [1988]) opted for dualism about the mental precisely because he could not see how the mechanics of mere material objects could ever intelligibly ground *reasoning*, a characteristic feature of intentionality. Reasoning involves transitions between contentful representational states that respect the meaning, or content, of those states. The transition from “today is Thursday” and “Thursday is ice cream day” to “today is ice cream day” respects the content of “today is Thursday” and “Thursday is ice cream day”. The transition from those two contents to “elephants like bananas” does not. Developments since Descartes' time in logic, computing, and in particular the computational theory of mind, have allowed us to see how reasoning could intelligibly emerge from simple mechanical processes. A purely mechanical automated reasoning system can be setup so that it, without in any sense “knowing” or “being aware” of the meaning of its states, transitions between states in a manner that respects the meaning of those states. As Haugeland [1981] puts it, if you set up the syntax correctly, the semantics - the intentional bit - will take care of itself. A related development was the demonstration of how material objects could be appropriately arranged to generate a universal Turing machine, which can, in principle, implement any computational process. The Turing machine, and modern computers, show us how very good at syntax

a “mere” collection of matter can be.

This second example of intelligible ground - the grounding of intentionality from the mechanics of material objects - is more controversial than the first. There are many live issues in cognitive science about whether our brains are really computers, what type of computer, and how the ways in which human minds emerge from brain operations compare to the ways the operations on your screen emerge from the mechanics of microchip transistors. But little of that is relevant here. I do not need to rely on the claim that our brains are computers. The computational theory of mind show us how intentionality, and in particular reasoning between contentful states, *could* intelligibly be grounded in material objects.

2.2 Understanding the Constraint

The Intelligibility Constraint is a transparency thesis about ground. It opposes mysterian theses about the fundamental. I will not argue for the intelligibility constraint here. Instead, I take for granted that something like intelligibility should govern metaphysical theorizing about fundamentality. I examine (i) how we should think about intelligible grounding and (ii) how certain contemporary arguments stack up once we do. These assumptions are not unwarranted. Long before the word ‘ground’ became fashionable, some sort of intelligibility requirement has implicitly governed most theorizing about fundamentality and the metaphysics of consciousness. Some investigation of exactly what the intelligibility constraint says is, however, warranted. Partly to convince the reader that believing in the constraint requires less commitment than one might have thought.

The easiest way to generate an intelligibility constraint would be to endorse a local thesis about the intelligibility of ground:

Local Intelligibility Requirement on Ground: If P grounds Q, then

P makes Q intelligible.

Local Intelligibility Requirement on Ground: If P grounds Q, then P makes Q intelligible.

This thesis has considerable plausibility. For many, grounding just is metaphysical explanation (Dasgupta [2015]: 558). If explananda make their explanans intelligible, then local intelligibility follows almost immediately. But the issue of the way in which grounding is a form of “metaphysical explanation”, and whether the relevant sense of explanation entails anything epistemic - such as intelligibility - is unclear (Kovacs [2017]). An intelligibility requirement is explicitly epistemic: Knowing P should allow a reasoner to make sense of the fact that Q. For those whose conception of grounding already builds in an epistemic sense of “metaphysical explanation”, the intelligibility constraint is nearly analytic. But others who think of grounding as more purely metaphysical will not think intelligible connection so automatic.

Doubts about local intelligibility arise from thinking about grounding as a generative metaphysical relation that occurs only relative to appropriate background conditions (Bennett [2017]: 3.3). If P grounds Q only relative to background condition C, then perhaps only P in conjunction with C, and not P alone, make Q intelligible. For these reasons, I resort to the weaker Global Intelligibility Constraint:

Global Intelligibility Requirement on Ground: The fundamental facts make all facts they ground intelligible.

This requirement is weaker than the global thesis. Suppose P grounds but fails to make intelligible Q because of the failure to include background conditions C. The local thesis fails. But the background conditions C are grounded in the fundamental, and if the fundamental makes C intelligible,

then the fundamental facts can succeed in making Q intelligible even when Q's ground P fails to do so.³ The lesson is that even someone who rejects the idea of the transparency and/or intelligibility of ground can still accept a similar requirement on the collection of all ultimate grounds. From here, I interpret the intelligibility constraint as the weaker, global thesis.

I said I would not offer positive argument for the intelligibility constraint. But two options for how to think of the constraint are worth mentioning. First, the *constitutive interpretation* takes the intelligibility constraint to follow from the nature of grounding and/or fundamentality. Second, the *methodological interpretation* takes the intelligibility constraint to be a good methodological principle for theorizing about ground. On this approach, intelligibility might or might follow from the nature of ground. But requiring that the fundamental grounds make their groundeds intelligible is sound metaphysical practice. I will not adjudicate between these two interpretations. I only stress that, in order to believe in the intelligibility constraint, you need not build any heavyweight explanatory requirements into the nature of ground.

Lastly, the Intelligibility Constraint falls far short of the Principle of Sufficient Reason (Dasgupta [2016], Leibniz [1989], Melamed & Lin [Spring 2020 Edition]). The Intelligibility Constraint does not require that the fundamental facts themselves be intelligible. It might be utterly mysterious why the initial conditions of the universe are what they are, or why a certain quark has the charm it does.

³The issues here roughly parallel discussions of local vs. global supervenience and local vs. global necessitation requirements on ground. In the same way that someone who rejects a local supervenience and/or necessitation thesis can still accept a global thesis (Leuenberger [2014] is an example), someone who rejects local intelligibility can still accept the global thesis. For discussion Leuenberger [2014] and Rabin [forthcoming]: fn1 .

2.3 *A Priori* Entailment Does Not Entail Intelligibility

P a priori entails Q iff the material conditional “If P then Q” is knowable *a priori*. Premise (P1) says that P makes Q intelligible if and only if P *a priori* entails Q. Both directions of this biconditional are false. In this section, I argue against the implication from *a priori* entailment to intelligibility.

Mathematical proof offers a plethora of examples of *a priori* entailment without intelligibility. Proofs by *reductio* often allow one to see *that* P entails Q without offering any insight whatsoever into *why* P entails Q, or *why* Q should be the case at all. Consider the following quote by the mathematician Michael Atiyah.⁴

I remember one theorem that I proved, and yet I really couldn't see why it was true. It worried me for years and years... I kept worrying about it, and five or six years later I understood why it had to be true. (Atiyah [1988]: 305)

Most if not all of these mathematical proofs yield *a priori* justification for “If P then Q”. But, as shown in the above quote, not all proofs make it intelligible why P leads to Q. If this is correct, then *a priori* entailment does not guarantee intelligibility.

2.4 Intelligibility Without *A Priori* Entailment

In this section I argue that P can make Q intelligible without *a priori* entailing Q. I believe that the following can happen. You are told that the Fs are ϕ , and then asked “Are the Gs ψ ?” You think about the matter carefully, considering all sides. You see some connections between Fs and Gs, and between ϕ -ing and ψ -ing. But even after days of *a priori* reasoning, having food

⁴Thanks are due here to discussion with Benedicte Veillet.

and water brought to you in your philosophical armchair, you can't see how the Fs being ϕ should decide the matter one way or the other regarding the Gs being ψ . You are then released from the philosophical armchair and the confines of *a priori* reasoning. You interact with the Fs and the Gs, watching them ϕ and ψ , perhaps even running experiments regarding when Fs ϕ and Gs ψ . With your increased understanding, you get it. "Ah ha! If the Fs ϕ , then the Gs must ψ ." Empirical methods further your understanding of Fs and Gs and ϕ -ing and ψ -ing in a way that *a priori* reasoning could not.

Let's examine a more concrete, albeit slightly fanciful, example. You are having difficulty learning to ride a unicycle. I suggest the following strategy, "This time, before you mount the unicycle, slide on this *Tyrannosaurus Rex* costume, strap a guitar over your shoulder, and place this harmonica between your lips." You are shocked at the suggestion. How could that possibly help? Am I trying to prank you? Being the diligent philosopher you are, you attempt to *a priori* reason your way to a conclusion. You know a lot about the physics of balance, how unicycles move, and the rest. But your *a priori* reasoning yields no verdict on the truth or falsity of the proposition "It's easier to learn to ride a unicycle while wearing a dinosaur costume, holding a guitar, and playing the harmonica." In fact, *a priori* reasoning, as well as common sense, leans toward a "false" verdict.

You take my advice despite your inclings that I'm pranking you. Lo and behold, disguised as a musical extinct apex predator, you accomplish your best unicycle ride yet: 40 meters! (At the end you crash hard on your face due to your miniature t-rex arms' inability to catch your fall). Despite the painful lump growing on your forehead, you beam a smile. *It is easier to ride a unicycle while dressed as a musical tyrannosaurus! What a surprise!*

Only once you actually ride the unicycle does it become intelligible how

such a counter-intuitive method could make learning to ride a unicycle easier. *A priori* reasoning can't provide this understanding, partly because one can't actually ride a unicycle, or put on a dinosaur costume, while in the *a priori* armchair.

The basic idea is that sometimes empirical *a posteriori* experience of the world provides understanding of, and makes intelligible, phenomena that *a priori* reasoning cannot. Thus, intelligibility does not require *a priori* entailment.

3 From the Physical to Consciousness *A Priori*: Debunking the Myth of Structure and Function

3.1 Intelligibility as *A Priori* Entailment

Premise (P2) says that the material conditional "If P then C" is not knowable *a priori*, where P and C completely describe the physical and consciousness facts, respectively. In this section, I argue against this premise. I will not *provide* an *a priori* route from the truths of microphysics to the truths about conscious experience. Instead, I will argue that, given the role that (P2) needs to play, we have good reasons to think it false.

For the *a priori* entailment account of intelligibility to even get off the ground, it must be the case that the normal mundane cases of intelligible grounding admit *a priori* entailment. There is no mystery of how a chamber that, at the more fundamental level, contained particles bouncing around could, at the less fundamental level, contain an ideal gas that exerted pressure p. The former intelligibly grounds the latter. But do the particle-facts *a priori* entail the pressure-facts? If not, the *a priori* entailment account of intelligible ground is a non-starter. While a "yes" answer is by no means a

surety, it's not implausible.⁵ Here, I wish to spot the proponent of *a priori* scrutability the claim that normal cases of intelligible grounding are also cases of *a priori* entailment. In fact, we'll assume that there is *a priori* entailment from the fundamental to all the non-consciousness facts. This fits with the dualist programme: the metaphysical issues raised by consciousness are supposed to be special and specific to conscious experience. If they aren't, and analogous issues arise for, e.g., the *a priori* entailment of facts about koalas, or neurons, then the dualist's arguments prove too much by leading to dualism about koalas and neurons as well.

The type of *a priori* entailment relevant to (P2) does not involve what you or I could figure out given armchair enough and time. Nor what the sharpest *a priori* reasoner on the planet now, or even 100,000 years from now, could do. The agents with respect to which we ask, "Could this person *a priori* deduce B from A?" to test for *a priori* entailment must be idealized in various ways. To start, they should have unlimited computational power and an ability to simultaneously consider an unlimited number of facts (e.g. the location of every fundamental particle in the universe). They must also be ideally rational, not suffer any cognitive deficiencies, and grasp all possible concepts. When we ask, "Does A *a priori* entail B?", we should envision the capacities of a God or a computationally enhanced alien genius, rather than any human being. This ideal sense of *a priori* entailment is what is at stake in premise (P2).

3.2 The Challenge From Idealization

When we recognize that the sense of the *a priori* relevant to the *A Priori* Argument for Unintelligibility is heavily idealized, a challenge to premise

⁵Deniers of a general thesis of *a priori* entailment of all truths from the fundamental level include Block & Stalnaker [1999], Ladyman & Ross [2007], Proponents include Chalmers [1996, 2012], Chalmers & Jackson [2001], Jackson [1998].

(P2) arises. The challenge, in its simplest and most naive form, is simply this: How are we, limited creatures that we are, to know what is and is not *a priori*, in this heavily idealized sense? Computationally unlimited gods can solve in seconds problems we'd take centuries to decipher.

The challenge, in this simple form, is not terribly difficult to defeat. It's clear that some types of truths simply don't *a priori* entail other types of truths. No matter how much information I provide about the distribution of plants in my yard, you'll never be able reason *a priori* to the identity of the best-selling rapper of the 2000s. The problem is not one of computational power, non-ideal rationality, or failure to grasp missing concepts. Truths about yard plants simply don't have the right character to *a priori* entail truths about album sales. Bottom line: even as limited creatures, we can sometimes determine what *a priori* entails what in the idealized sense.

A more nuanced form of the challenge proves more substantive. First, notice that we possess good evidence for physicalism. I won't go into the details of that evidence here.⁶ If we did not have such evidence, philosophers wouldn't spend so much ink defending the cause, nor would physicalism be the default view in contemporary philosophy of mind. Even if we did not possess good evidence for physicalism, the context of the current dialectic forces us to act as if we did. Remember, the problem of intelligible connection from the physical is supposed to be particular to consciousness. There is no analogous hard problem of koalas. We are operating under the assumptions that (i) intelligibility is *a priori* entailment and (ii) the non-conscious features of the world are *a priori* entailed by the physical. This entails that physicalism is doing extremely well, all things considered. Physical fundamentalia

⁶Two prominent arguments for physicalism include the argument from causal closure (Montero [2003], Papineau [2001]) and the argument from methodological naturalism (Stoljar [2009]: section 17).

intelligibly ground 99% of the world's contents. This lends support, albeit indirectly, to the claim that the physical intelligibly grounds the remaining 1%: consciousness. Thus we have indirect evidence that, in the idealized sense, the physical *a priori* entails consciousness. Call this *the inductive argument* for physicalism.

The challenge from idealization, in its more substantive form, is this. Both as a matter of fact and as a result of the current dialectic surrounding the hard problem's challenge to physicalism, we have very good evidence that physicalism is the correct theory of the complete nature of the world's fundamental building blocks. Instead of responding to the current non-ideal explanatory gap between the physical and consciousness by abandoning physicalism, instead reject the inference from the current explanatory gap to an idealized *a priori* gap. This physicalist move is perfectly compatible with acknowledging that we do have some access to what is and is not *a priori* in the ideal sense.

On the other side, we also have evidence to the contrary conclusion that the physical does not *a priori* entail (in the idealized sense) consciousness. This evidence comes from the famous explanatory gap in explaining consciousness as a physical process. We, non-ideal beings that we are, can't see how the physical could ground consciousness. And it's sometimes hard to see how idealization could help.

We have evidence both for and against an ideal *a priori* entailment from the physical to consciousness. Which is stronger? Our evidence for physicalism, or our evidence that the *a priori* gap between the physical and consciousness will remain in the limit case? One's response will depend on how strong one thinks the evidence for physicalism really is, and how good our access to what *a priori* entails what in the idealized sense at the limit

scenario in which we have complete information about the fundamental level. Here rational parties can reasonably disagree. This is part of the reason there has been so much discussion in the philosophical literature on exactly this issue.

Those who think that the *a priori* gap will remain in the limit case opt for dualism or non-physicalist monism. Those who think that the evidence for physicalism is stronger will choose to keep the faith, and hope for the emergence of an *a priori* route from the physical to consciousness, thus rendering the physical grounding of consciousness intelligible. (A third option is to reject the intelligibility constraint on theories of the fundamental. I won't consider this move.)

Two lines of thought further strengthen the physicalist line. First, if one relaxes the assumption that intelligibility is *a priori* entailment (which I think one should), we don't need an *a priori* link to make the physical grounding of consciousness intelligible. This makes defense of physicalism much easier. Second, looking at the history of science and of explanatory gaps sheds light on the current debate and favors the more conservative physicalist option.

3.3 The Challenge From Historical Precedent

The dualist believes that we should respond to the current explanatory gap between the physical and consciousness by revising our theory of what is fundamental. It's not all physical stuff, as science would lead us to believe. However, we have faced explanatory gaps before. In the 18th and 19th centuries, the materialist conception of the fundamental (a predecessor to physicalism) faced the challenge of *life*. At the time, no one could see how material components could generate the phenomena characteristic of life. This led many prominent thinkers to endorse vitalism, a form of dualism about life (rather than about the mental). The vitalists claimed that "there are additional

fundamental forces in nature that are on a par with those Newton ascribed to all matter” (Bechtel & Richardson [1998]: 1).

Descartes (Cottingham *et al.* [1988]) similarly rejected materialism, famously opting for a type of substance dualism, because he thought that material substances could not account for the intentional process of reasoning. In both cases, no one could see how a given theory of the fundamental (materialism) could intelligibly ground a troublesome target phenomena (life in the first case, reasoning in the second). Applying something like the intelligibility constraint to an explanatory gap, these thinkers rejected the materialist theory of the fundamental.

Unfortunately, these dualists turned out to be wrong. Material fundamentalia are capable of generating both life and reasoning. And they needn't violate the intelligibility constraint in order to do so. The respective explanatory gaps were bridged by further developments and the lure of dualism faded.

19th century chemists created urea, an organic compound, from an inorganic substance (ammonium cyanate). This called into question the division between the inorganic and the organic on which vitalism relied. The organic had been alleged to require and imbue the vital force. But scientists had managed to bridge the gap between the inorganic and the organic in a lab without any living thing or vital force playing a special role. Other advancements, including the manufacture of fatty acids and better understanding of cell development, sealed vitalism's fate.

The mystery of intentionality remained long after Descartes' concerns about reasoning. (I take the problem of accounting for reasoning in a material world to be a sub-problem of accounting for intentionality in a material world.) Not until 20th century advancements in logic and computing did the way in which intentionality and reasoning could arise from the non-

intentional become intelligible. Much remains to be discovered, and arguments remain about how, why, and whether cognition is computation. But the basic idea that a computer, such as a Turing machine, can be constructed from simple and purely physical ingredients and simultaneously generate the massively complicated behavior and correspondences that are characteristic of intentionality, is not under debate. Thus, the computational approach to mind shows how intentionality might be grounded in an underlying level that is not itself intentional.

The developments that shed light on how the target phenomena in these two cases (life and reasoning) could be intelligibly grounded in physical fundamentalia were of quite different natures. In the case of life, the advancement was mostly empirical. Laboratory work yielded better understanding of the ingredients involved in living and non-living substances and how those two related to each other. In the case of intentionality and reasoning, the advancement was more conceptual than empirical.

3.4 Taking up the Challenges

Contemporary dualists are well aware of the challenge from historical precedent. They maintain that the case of consciousness is importantly different from the cases of life and intentionality. Chalmers [2003] writes that “in explaining life... the only phenomena that present themselves as needing explanation are phenomena of adaptation, growth, metabolism, reproduction, and so on, and there is nothing else that even calls out for explanation. But... the case of consciousness is different and possibly unique, precisely because there is something else, phenomenal experience, that calls out for explanation.” (109)

I agree with Chalmers that the case of consciousness is different than the case of life. Consciousness experience is neither adaptation nor reproduction.

But the case of thermodynamics is different than the cases of consciousness and of life. Pressure and temperature are neither conscious experience nor reproduction. This response to the challenge from historical precedent requires that consciousness lie on one side, while life, thermodynamics, intentionality, and the remainder of the non-fundamental world all lie on the other side. What makes the problem of consciousness special and unique, and different from every other explanatory gap ever faced in the history of science?

Chalmers' basic answer, on which others (Alter [2016]) offer variations, is that physics, life, thermodynamics, intentionality, coffee, and all the rest are "structural-functional". Consciousness is not. Structural-functional phenomena can never make intelligible, or *a priori* entail, non-structural-functional phenomena. Thus, physical fundamentalia can never make intelligible, or *a priori* entail, consciousness. Or so the story goes. (I have serious doubts).

3.5 Clarifying the Desiderata

Premise (P2) of the *A Priori* Argument for Unintelligibility says that the conditional "If P then C" is not knowable *a priori*, where P and C completely describe the physical and consciousness facts, respectively.

In the remainder of this paper, my plan will be to explore a variety of accounts of what "structural-functional" amounts to. What exactly is the "structural-functional", such that (i) physics, life, and the most of the world is structural-functional, while consciousness is not, and (ii) structural-functional truths can never *a priori* entail non-structural-functional truths? My plan will be to argue that there is no account of structure and function that can simultaneously meet both desiderata.

Clarification of the desiderata is necessary. In order for an account of "structural-functional" to vindicate premise (P2), it must be the case that (a) non-conscious properties are completely structural functional, (b) at least

some conscious properties are not, and (c) it is plausible that structural-functional properties are incapable of intelligibly generating non-structural functional properties. The basic idea of the argument is that some features of consciousness (perhaps their phenomenal essence, e.g. the “painfulness”) are beyond the reach of intelligible grounding from physical properties. If physical properties are even partly not structural-functional, then the claim that physical properties can’t intelligibly ground consciousness because of the structural-functional / non-structural-functional divide does not stick. The non-structural-functional aspects of physical properties could do the grounding. Consciousness must have at least some non-structural-functional features: the features that are out of reach. Consciousness can have some structural-functional features. Relations of similarity and difference between color experiences, as represented in a color wheel, might provide an example of a structural-functional feature of consciousness. If some non-fundamental non-conscious features of reality are even partly not structural-functional, then the charge that the structural-functional cannot intelligibly ground the non-structural-functional does not stick, because we are operating under the assumption that all the non-conscious features of reality (the coffees, cars, and koalas) can be intelligibly grounded in the physical. Thus, all the non-conscious features of reality must be structural-functional.

3.6 Accounts of “Structure and Function”

Each of sections 3.7-3.10 argues that one of the following four accounts of what it is to be “structural-functional” does not succeed.

- extrinsic
- dispositional
- not absolutely intrinsic

- spatio-temporal-causal

3.7 First Account: Structural-functional as Extrinsic

Philosophy has long acknowledged a division between intrinsic and extrinsic properties. The intrinsic properties of a thing are those properties the thing has in virtue of how the thing itself is. Extrinsic properties are properties held at least partly in virtue of how matters other than the thing itself are. Examples illustrate the distinction. Consider a house. The house's intrinsic properties include having 9 rooms, one staircase, a mass of 60,000kg, two bathrooms, a peaked roof, and a paucity of cabinets in the kitchen. Extrinsic properties include being located in Vermont, being owned by the Vanderbergs, and being coveted by all the neighbors. The *duplication test* provides a useful heuristic for determining intrinsicity. x 's having of property F passes the duplication test if any duplicate of x also has F . Always passing the duplication test is good evidence that a property is intrinsic.⁷ There are complicated issues about what exactly the intrinsic/extrinsic divide comes to, and whether there is really one distinction at play here (Cf. Marshall & Weatherson [2014]). But the rough and ready distinction should suffice for our purposes, mostly because the distinction seems hopeless for the task at hand.

In order for intrinsicity to successfully provide a precisification of “structural-functional” that vindicates premise (P2) of the *A Priori* Argument for Unintelligibility, it must be the case that all physical properties are extrinsic and some conscious properties are not. But this is clearly false. The paradigms of intrinsic properties discussed above, such as having a mass of 60,000kg, are

⁷Problems arise when we consider necessarily instantiated properties (e.g. “being such that there are infinitely many prime numbers”), which are nevertheless intrinsic, as well as upon deeper investigation of what it takes for two things to be “duplicates” in the relevant sense.

clearly physical. Other physical properties, such as being located in Vermont or within 1 kilometer of a disco, are extrinsic. Physical properties can be either intrinsic or extrinsic.

Furthermore, on this account it's false that structural-functional properties are incapable of intelligibly generating non-structural functional properties. Extrinsic properties can *a priori* entail intrinsic properties. If I have the extrinsic property of being ten meters to the left of something cube-shaped, this *a priori* entails that something ten meters to my right has the intrinsic property of being cube-shaped.

In sum, on the account of “structural-functional” according to which structural-functional properties are extrinsic, and non-structural-functional properties intrinsic, it is false that the physical is exclusively structural-functional and also false that the structural-functional is incapable of intelligibly generating the non-structural.

3.8 Second Account: Structural-functional as Dispositional

Dispositional properties are properties whose nature is exhausted by how the bearer of that property is disposed to behave in certain situations. Fragility is a paradigm of a dispositional property. Whether a thing is fragile is determined by whether that thing is inclined (“disposed”) to break when struck. In contrast, categorical properties are characterized by how the thing itself is, rather than how the thing interacts with other entities or behaves in certain situations. (For more on the distinction cf. Choi & Fara [2016]: section 2). The dispositional/categorical distinction cross-cuts the extrinsic/intrinsic distinction. The fragility of a martini glass is simultaneously dispositional and intrinsic, because the glass's fragility is entirely due to the structural and material makeup of the glass itself.

Application of the dispositional/categorical distinction to the metaphysics

of consciousness is popular among Russellian monists and panpsychists of various stripes (cf. Chalmers [1996]: chapter 4, Chalmers [2003], Montero [2010], Rosenberg [2004], Seager & Allen-Hermanson [2015]: section 4.3, Strawson [1997, 2006], Stoljar [2001], and the essays in Alter & Nagasawa [2015] and Jaskolla [forthcoming].)

The basic idea in these views is that the physical world is purely dispositional. All it is to be an electron, or a boson, or have a certain mass, is to be disposed to react to other physical stuff, such as a proton or other masses, in certain ways. Consciousness is different. A ticklish experience cannot be exhausted by the ways in which someone experiencing the tickle is disposed to behave.

I am skeptical about this final step. *Contra* the claim that ticklish experiences are not wholly dispositional, the nature of a ticklish experience does seem to be exhausted by how it is disposed to make the subject of that experience feel. Similarly, the nature of a ticklish experience might be exhausted by how it is disposed to combine with other “atomic” experiences to yield a complete experiential field, i.e. the overall conscious state of the subject.

Another problem with the dispositional account of the structural-functional is that many paradigm physical properties are categorical. Ellis [2010] writes that “shape, size, orientation, speed, handedness, direction” and “angular separation” are all categorical. But, clearly, all these properties are physical. So on this account it’s false that all physical properties are structural-functional.

Furthermore, even if fundamental physical properties (being negatively charged, having a mass of one gram) are dispositional, it is implausible that all the non-conscious properties are similarly dispositional. An atom-by-atom duplicate of a koala generated at random from the void will have all the same

dispositions as a koala. But it's plausible that the atom-by-atom duplicate will not be a koala, because it does not have the correct causal origins or bear the appropriate relations to the koala species that inhabits Australia. Thus, the property of being a koala is not purely dispositional.

The dispositional account account of the structural-functional is such a non-starter that it's tempting to think that I must have mischaracterized the view. But the proponents of using the dispositional/categorical distinction to divide the structural-functional from the rest don't tell us much beyond simple examples like fragility. They point to Russell [1927]'s *The Analysis of Matter* and literature in the philosophy of science. But it's clear that, in those discussions, physical properties lie on both sides of the dispositional/categorical divide. Fans of pan-psychism often point to the idea, much discussed in the philosophy of science, that "dispositional properties must have categorical bases" to support their pan-psychist view that categorical mental properties underpin basic physical properties, like charge and mass, which are alleged to be dispositional. First, "dispositional properties must have categorical bases" is far from a truism of contemporary philosophy of science. But, more importantly, the philosophers and metaphysicians of science simply don't mean what the pan-psychists mean when they used the words 'categorical' and 'dispositional' (cf. the essays in Marmodoro [2010]).

3.9 Third Account: Structural-functional as Absolutely Intrinsic

Another account of the structural-functional can be gleaned from the work of Alter & Nagasawa [2012] and Pereboom [2011, 2014]. We start with Pereboom's distinction between comparatively and absolutely intrinsic properties. An *absolutely intrinsic* property is a property that "has no extrinsic aspects or components" (Pereboom [2014]: 50). In contrast, a *comparatively intrinsic* property does have extrinsic aspects or components. On this account,

a structural functional property can be either intrinsic or extrinsic, but it can't be absolutely intrinsic. The general picture we get is that some features of consciousness are absolutely intrinsic, while physical properties are either extrinsic or comparatively intrinsic, and such properties are incapable of *a priori* entailing absolutely intrinsic properties.

What does it take for an intrinsic property to have “extrinsic aspects or components”? For conjunctive properties (“determined mother”), one of the conjuncts might be extrinsic and the other intrinsic. This would yield a property with an extrinsic component or aspect. But such a conjunctive property will be, invariably, extrinsic, not intrinsic. Stoljar [2015]: 331 proposes to understand the distinction using the notion of parthood.⁸ The idea is that an intrinsic property F of x might be due to the properties G_1, G_2, G_3, \dots of x 's proper parts. In such a case, if any of G_1, G_2, G_3 are extrinsic, then F is comparatively intrinsic. Comparatively intrinsic properties have a “lower grade” of intrinsicity, because they are derived from extrinsic properties. If all of G_1, G_2, G_3, \dots are intrinsic, then F is absolutely intrinsic.

I have several concerns about this understanding of absolute intrinsicity. First, what if an entity has no proper parts? The distinction seems not to apply. Perhaps all its intrinsic properties are *de facto* absolutely intrinsic.⁹ Or *de facto* comparatively intrinsic?

The “absolutely intrinsic” gambit only works if some properties associated with consciousness are absolutely intrinsic, while no non-conscious properties are. The best candidates are purely experiential properties, such as experiencing a ticklish sensation. This property seems like a decent candidate for absolute intrinsicity. But I'm unsure how to evaluate the claim.

⁸My presentation differs from Stoljar's. I use the idea of a property being “due to”, “derived from”, or “grounded in” another property. Stoljar uses the notion of necessity, where certain properties or patterns of properties necessitate others.

⁹This is the option Stoljar chooses.

The property seems to have some extrinsic aspects or components. The relevant sensation is similar to the property of experiencing a sensation of fingers drawing across one's bare skin. Though perhaps that does not count as an "aspect" or "component" of the property. If so, I'd like to hear a lot more about what does and does not count as an "aspect" or "component" of a property.

Furthermore, on some theories of consciousness, the property of experiencing a ticklish sensation won't count as intrinsic at all, let alone absolutely intrinsic. Consider a panpsychist position with a holistic bent, which claims that which experiences an agent has are partly determined by the "phenomenal field" surrounding that agent. Whether I experience blue or violet is partly determined by the panpsychist goings-on in my immediate surroundings. On this type of view, my experience of violet is partly due to factors outside me, which are not parts. "Experiencing violet" is not an intrinsic property at all, let alone absolutely intrinsic.

Perhaps the suggested panpsychist view is far-fetched. But it's clearly an option. Part of the problem here is that it is sometimes difficult to tell whether a property is intrinsic or extrinsic, and doubly so whether its intrinsicity is absolute or comparative. The problem is particularly acute in the case of consciousness. We know so very little about the factors that lead to conscious experiences that it is virtually impossible to adjudicate the question of whether my property of having a ticklish sensation is due to the extrinsic or the intrinsic properties of my parts. Pereboom and company act like it is obvious, or at least relatively easy to determine from the armchair via introspection, that phenomenal experiential properties are absolutely intrinsic. In so far as I understand the notion, I find the issue difficult to adjudicate. Furthermore, if Pereboom and company are right, then we can

rule out, from the armchair via introspection, the holistic panpsychist view I discussed above. But that can't be correct.

In sum, the notion of absolutely intrinsicity cannot bear the burden that its proponents place on it in performing the vital task of distinguishing the structural-functional from the non-structural-functional.

3.10 Fourth Account: Structural-functional as Spatio-temporal-causal

Chalmers [2010a] offers the following account of the structural-functional, which he calls 'structural-dynamic':

A structural-dynamic description is one that is equivalent to a Ramsey sentence whose O-terms are limited to spatiotemporal expressions, nomic expressions, and mathematical and logical expressions. (210n18)

The notion of a Ramsey sentence comes originally from the philosopher and mathematician Frank Ramsey. Carnap [1958] and Lewis [1970] developed the idea further. Ramsification is a method for simultaneously defining one or more expressions in terms of others. The terms to be defined are dubbed "theoretical", while the terms doing the defining are called "observational", and tend to count as already understood. The method originates in the philosophy of science. It was designed for defining theoretical scientific expressions in terms of an allegedly observational vocabulary. However, the method is general enough that neither set of expressions need be in any important sense tied to theory or observation. The basic idea of Ramsification involves taking a sentence and replacing occurrences of the theoretical terms with (possibly higher-order) variables, and then quantifying over those variables. Consider the theoretical sentence 'All electrons attract pro-

tons, repel other electrons, and leave wave-like patterns when unobserved during a double-slit experiment.’ We “ramsify” the sentence by replacing all occurrences of ‘electron’ and ‘proton’ with second-order variables E and P. These terms count as “theoretical” in this context, while the remainder are “observational”. The resulting Ramsey sentence is ‘ $\exists E \exists P \forall x \forall y \forall z$ If Ex, then (i) if Py, then x attracts y, (ii) if $x \neq z$ and Ez, then x repels z, and (iii) x leaves wave-like patterns when unobserved during a double-slit experiment.’ The sentence says that there are two properties, E and P, such that E behaves like the electron-property and P behaves like the proton-property.

Chalmers’ idea is that any direct mention of electrons, or protons, or rocks, all of which are structural-functional, can be “Ramsified out” and replaced with quantifiers and a statement containing only a limited vocabulary of privileged expressions. In the example above, ‘attract’, ‘repel’, and ‘acts like a wave’ were privileged and did not get replaced by quantifiers. One might wonder whether those expressions could also be replaced. Chalmers believes one can repeat the Ramsification process until the only remaining vocabulary is spatio-temporal expressions, mathematical and logical expressions, the term ‘cause’, and nomic expressions such as ‘it is a law that’.

The account of the structural-functional on offer is that x is structural-functional if and only if it can be given a Ramsey-style definition that uses only the vocabulary of spatio-temporal, causal, nomic, and mathematical expressions. Let’s evaluate this account. We should remember that, the dialectic here requires that, outside of conscious experience, our entire world is structural-functional. All the koalas, cars, and coffees, as well as the electrons, quarks, and nuclear forces, are structural-functional. That means all descriptions of the world in terms of koalas, cars, coffees, and all the rest can be replaced by Ramsey-style descriptions containing spatio-temporal

expressions, the phrases '... causes ...' and 'it is a law that', along with some maths.

I must confess that I find the suggestion that almost all expressions can be reductively defined in this way preposterous. In the background, I hold a general philosophical view that few terms, if any, can be given definitions, in the sense of necessary and sufficient conditions. I believe that the impossibility of definition holds even when we give ourselves recourse to the full and robust vocabulary contained in say, the Oxford English Dictionary. Chalmers claims that *almost every term* (or at least all the structural-functional terms, which is a lot) can be given a definition in the *tremendously restricted* vocabulary of the spatio-temporal-causal-nomic plus maths. This claim boggles the mind. Philosophers have failed for centuries to come up with a definition of 'knowledge' using such terms as 'justification' and 'belief'. I find it hard to believe that they could succeed if they used the words 'cause' and 'located in the spatio-temporal manifold at location x, y, z, w' instead of 'justification' and 'belief'.

We see now that the spatio-temporal-causal account of the structural-functional is based upon an implausible and unargued for brand of strong reductionism about almost everything. We should reject it.

A defensive point on behalf of the proponent of the spatio-temporal-causal account deserves to be made. For the question of whether the structural-functional could *a priori* entail truths about consciousness, the proponent of the structural-functional as spatio-temporal-causal does not require something as strong as definition. Definition requires two-way entailment between *definiens* and *defiendum*. But *a priori* entailment is a one-way relation. For current purposes, all that the proponent of the structural-functional as spatio-temporal-causal requires is (a) one-way *a priori* entailment between

spatio-temporal-causal and all the non-consciousness truths and (b) the impossibility of entailment from the spatio-temporal-causal to (certain types of) facts about consciousness. The giving of definitions for ‘coffee’, ‘koala’, and ‘quasar’ in terms of the spatio-temporal-causal is one way to achieve (a), but that method is supererogatory.

This retreat position only makes matters worse. First, it remains implausible that such an impoverished description could *a priori* entail a complete story of the world (minus consciousness). Second and more importantly, the account on offer, which says that all the non-consciousness truths are *a priori* entailed by a quantified spatio-temporal-causal Ramsey sentence while the consciousness truths are not, is circular. Once we abandon the definitional approach, structural-functional just means “a priori entailed by a quantified spatio-temporal-causal Ramsey sentence” and non-structural-functional means “not so entailed”. We were looking for an account of the structural-functional distinction that could answer the question, “Why does the Ramsey sentence a priori entail the truths about coffees, koalas, and quasars, but not the consciousness truths?” We are told, “Because the former are structural-functional, whereas the latter are not.” We follow up, “Why does the Ramsey sentence a priori entail all the structural-functional truths, but not the others?” The view under consideration is that structural-functional just means “a priori entailed by such a Ramsey sentence” and non-structural-functional means “not so entailed”. Without a doubt, this conception of the structural-functional succeeds in delivering the desired difference in entailment status. But it does so by *fiat*. The word ‘structural-functional’ is doing no actual work in the reasoning. It’s merely a stand-in for the phrase ‘a priori entailed by the Ramsey sentence’. To call the truths about coffees, koalas, and quasars “structural-functional” and the conscious-

ness truths “non-structural-functional” is just to assert that the former are a priori entailed and the latter are not. But this is exactly the conclusion that the structural-functional distinction was supposed to establish! This conception of “structure and function” does nothing to explain why there is a difference between the two. It may, however, work as an effective placebo or smokescreen on those who are not paying close attention and/or are already predisposed to believe the conclusion being argued for.

The proponent of the structural-functional as spatio-temporal-causal has a difficult sell. They ask us to be tremendously optimistic about the prospects for *a priori* entailment from their restricted base (“it entails almost everything there is!”) but not too optimistic (“but not consciousness!”). This is a very fine line to walk. Why should our optimism, or our beliefs about *a priori* entailment, fall exactly where the fan of the intelligibility argument requires? This final point, about the difficulty of the task that those who defend the intelligibility argument face, extends well beyond the particular view that the structural-functional is the spatio-temporal-nomic, and even beyond the view that intelligibility is *a priori* entailment. The opponent of physicalism has a very fine line to walk, regardless of how she spells out the notion of intelligibility.

4 Keep the Faith, O Ye Physicalists

Fundamentality physicalism says that all the fundamental building blocks of our world are physical. Any theory of the fundamental must satisfy the intelligibility constraint. Physicalism does well. For almost all the contents of our world, from coffee to cars to koalas to quasars, there is no mystery of how a world that, at the fundamental level, was purely physical, could, at the same time, contain coffees, cars, koalas, or quasars. There do seem to be

genuine issues regarding how a world that was fundamentally physical could contain conscious experience. We can't see how physical materials could metaphysically ground consciousness.

If we accept the intelligibility constraint, then a lack of intelligible route between the physical and consciousness entails the falsity of fundamentality physicalism. I am inclined to accept the intelligibility constraint, and this corollary of it. But this does not decide the issue of physicalism itself. The question now becomes: in light of the present gap, what should we think about whether there is an intelligible route, in the futuristic ideal case, from the physical to consciousness?

As discussed in section 3.3, humankind has faced similar gaps in the past. Our predecessors could not see how material fundamentalia could ground life or intentionality. Some responded by rejecting materialism and opted for forms of dualism or primitivism. But they turned out to be wrong. Further developments bridged the gaps and allowed us to see how intentionality could arise from materially implemented computational processes, as well as how organic materials and life might be grounded in inorganic materials that contained no vital force. These cases encourage us to be conservative when revising our theory of the fundamental. We can make mistakes about what, in the ideal case and the long run, intelligibly grounds what.

Some think that the historical precedents have little bearing on the problem of consciousness. They allege that the problem of consciousness is importantly different than previous explanatory gaps. According to this line of thought, we can tell, either from examining our physical and phenomenal concepts or from peering into the nature of the physical and the phenomenal, that the lack of intelligible connection between the two will always remain, no matter how much we learn. I am doubtful that we can see clearly enough

into either our concepts or into the nature of the things themselves to justify such confidence.

One particular implementation of this strategy is the structure and function gambit, which holds that physical fundamentalia and the all the worldly contents they ground are structural-functional, while consciousness is not. I have argued at length (sections 3.6 - 3.10) that this strategy does not work. Once you place pressure on the notion of “structure and function”, the promise it holds for justifying an inference to an unbridgeable gap between the physical and consciousness crumbles.

Some contemporary philosophy takes the attitude that the metaphysical issues surrounding the hard problem are pretty much figured out. The major landmarks in logical space have been mapped; no empirical breakthrough could alter the terrain. All that’s left for future generations is to fill in the details. I think this line of thought mistaken, especially considering how young inquiry into these topics is.

We remain in the early days in the study of consciousness and of the brain. On the empirical side, despite exciting advances and significant recent progress, we remain grossly ignorant of how the brain actually works, and how various mental phenomena are grounded in it. We don’t even know what the neural substrates of conscious experience are! On the conceptual side, we haven’t been grappling directly with the hard problem for long. Until relatively recently, the problem of phenomenal consciousness was rarely separated from other questions in the study of the mind. Much high-quality thought has gone into the topic. But it’s still early days, and there remains room for significant conceptual breakthroughs.¹⁰

¹⁰My optimism about the potential for a conceptual and/or empirical breakthrough that could “close” the explanatory gap is, to some degree at least, shared by Stoljar [2006] and Sturgeon [2000].

I am optimistic about a conceptual breakthrough that would make intelligible how physical processes could ground consciousness, in a manner similar to that by which the computational theory of mind made intelligible how intentionality could be grounded in a world that was fundamentally non-intentional. This conceptual breakthrough might be aided by empirical developments. Mapping the correlations between the varieties and dynamics of conscious experience and the operations of the brain will point us towards the promising routes for intelligibly connecting the two.

Making the generation of consciousness from the physical intelligible might require some tweaking of both our physical and phenomenal concepts. But a change in those concepts, or in how we view the problem, does not entail that we have changed or abandoned the subject. Instead, we will have come to a better understanding of the target phenomenon. I believe that such a conceptual shift happened with the computational theory of mind. One could consider the type of intentionality that computational processes intelligibly ground “light-weight” compared to the intentionality that Descartes wanted. (Of course, there are interpretative issues here). Some contemporary philosophers who reject the computational approach probably continue to operate with a more heavy-weight conception of intentionality. “That’s all just mechanical processes,” they exclaim. “Where’s the *understanding* and the *meaning*?” But part of what the computational theory teaches us is that purely physical syntactic engines can operate in a semantically respectable way without a need for any extra element, or homunculus, doing the understanding or bestowing meaning. In the same vein, I am suspicious that our concept of phenomenally conscious experience is too heavy-weight. The intelligible grounding of consciousness in the physical might require some lightening up. But I don’t think that such a lightening entails a change of

subject or abandoning the project of explaining conscious experience.

As a theory of the fundamental, physicalism does shockingly well. For almost every phenomenon our world contains, from cars to coffees to koalas to computers, it is intelligible how that phenomenon could be generated from a small supply of physical fundamentalia: quarks, leptons, the nuclear forces, and their like. That is an impressive performance by physicalism. It should not be forgotten when evaluating physicalism's prospects. Consciousness provides a notable exception. In the face of this obstacle, some wish to abandon physicalism. I urge physicalists to keep the faith. Abandoning the *a priori* entailment model of intelligible grounding opens up space to resist the temptation toward dualism offered by the hard problem of consciousness. Humankind's investigation of consciousness is in its earliest stages. There is much that we do not understand. My bet is that, eventually, after we have learned much more about consciousness, partly through *a posteriori* investigation, the generation of consciousness from the physical will become intelligible. It would be high folly to opt for dualism and declare the mystery of consciousness impenetrable when we are only beginning to peel back the layers of consciousness's onion.

References

- The Stanford Encyclopedia of Philosophy*.
- Alter, Torin. 2016. The Structure and Dynamics Argument Against Materialism. *Nous*, **50**, 794–815.
- Alter, Torin, & Nagasawa, Yujin. 2012. What Is Russellian Monism? *Journal of Consciousness Studies*, **9**, 67–95.
- Alter, Torin, & Nagasawa, Yujin (eds). 2015. *Consciousness in the Physical World: Perspectives on Russellian Monism*. OUP USA.

- Atiyah, Michael. 1988. *Collected Works*. The Clarendon Press Oxford University Press.
- Bechtel, William, & Richardson, Robert C. 1998. Vitalism. *In*: Craig, E. (ed), *Routledge Encyclopedia of Philosophy*. Taylor and Francis.
- Bennett, Karen. 2017. *Making Things Up*. Oxford University Press.
- Block, Ned, & Stalnaker, Robert. 1999. Conceptual Analysis, Dualism, and the Explanatory Gap. *Philosophical Review*, **108**, 1–46.
- Carnap, Rudolf. 1958. Beobachtungssprache und theoretische sprache. *Dialectica*, **12**, 236–248.
- Chalmers, David. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- Chalmers, David. 2003. Consciousness and Its Place in Nature. *Pages 102–142 of*: Stephen Stich, Ted Warfield (ed), *The Blackwell Guide to Philosophy of Mind*. Malden, MA: Blackwell.
- Chalmers, David. 2010a. *The Character of Consciousness*. Oxford University Press.
- Chalmers, David. 2010b. The Two-Dimensional Argument Against Materialism. *In*: Chalmers [2010a].
- Chalmers, David. 2012. *Constructing the World*. Oxford: Oxford University Press.
- Chalmers, David, & Jackson, Frank. 2001. Conceptual Analysis and Reductive Explanation. *Philosophical Review*, **110**, 315–61.
- Choi, Sungho, & Fara, Michael. 2016. Dispositions. *In*: Zalta, Edward N. (ed), *The Stanford Encyclopedia of Philosophy*, spring 2016 edn. Metaphysics Research Lab, Stanford University.
- Cottingham, John, Stoothoff, Robert, & Murdoch, Dugald (eds). 1988. *Descartes: Selected Philosophy Writings Volume I*.
- Dasgupta, Shamik. 2015. The Possibility of Physicalism. *The Journal of Philosophy*, **111**, 557–92.
- Dasgupta, Shamik. 2016. Metaphysical Rationalism. *Noûs*, **50**(2).

- Ellis, Brian. 2010. Causal Powers and Categorical Properties. *In: Marmodoro [2010]*.
- Haugeland, John. 1981. Semantic Engines: An Introduction to Mind Design. *In: Haugeland, J. (ed), Mind Design*. MIT Press.
- Jackson, Frank. 1998. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press.
- Jaskolla, G. Bruntrup L. forthcoming. *Panpsychism*. Oxford University Press.
- Kovacs, David. 2017. Grounding and the Argument From Explanatoriness. *Philosophical Studies*, **174**, 2927–2952.
- Ladyman, James, & Ross, Don. 2007. *Everything Must Go: Metaphysics Naturalized*. Oxford.
- Leibniz, G. W. 1989. *Philosophical Essays*. Indianapolis: Hackett. Translated by Roger Ariew and Daniel Garber.
- Leuenberger, Stephan. 2014. Grounding and Necessity. *Inquiry : An Interdisciplinary Journal of Philosophy*, **57**(2), 151–174.
- Lewis, David. 1970. How to Define Theoretical Terms. *Journal of Philosophy*, **67**, 427–46.
- Marmodoro, Anna (ed). 2010. *The Metaphysics of Powers: Their Grounding and Their Manifestations*. Routledge.
- Marshall, Dan, & Weatherson, Brian. 2014. Intrinsic vs. Extrinsic Properties. *In: SEP [n.d.]*.
- Melamed, Yitzhak, & Lin, Martin. Spring 2020 Edition. Principle of Sufficient Reason. *In: Zalta, Edward N. (ed), Stanford Encyclopedia of Philosophy*.
- Montero, Barbara. 2003. Varieties of Causal Closure. *Pages 173–187 of: Walter, Sven, & Heckmann, Heinz-Dieter (eds), Physicalism and Mental Causation*. Imprint Academic.
- Montero, Barbara. 2010. A Russellian Response to the Structural Argument Against Physicalism. *Journal of Consciousness Studies*, **17**(3-4), 3–4.

- Papineau, David. 2001. The Rise of Physicalism. *In*: Gillett, Carl, & Loewer, Barry M. (eds), *Physicalism and its Discontents*. Cambridge University Press.
- Pereboom, Derk. 2011. *Consciousness and the Prospects of Physicalism*. Oxford University Press.
- Pereboom, Derk. 2014. Russellian Monism and Absolutely Intrinsic Properties. *Pages 40–69 of*: Kriegel, Uriah (ed), *Current Controversies in Philosophy of Mind*.
- Rabin, Gabriel Oak. 2020. Fundamentality Physicalism. *Inquiry*.
- Rabin, Gabriel Oak. forthcoming. A Short Argument From Modal Rationalism to Fundamental Scrutability. *Thought*.
- Rosenberg, Gregg H. 2004. *A Place for Consciousness: Probing the Deep Structure of the Natural World*. Oxford: Oxford University Press.
- Russell, Bertrand. 1927. *The Analysis of Matter*. London: Routledge.
- Seager, William, & Allen-Hermanson, Sean. 2015. Panpsychism. *In*: Zalta, Edward N. (ed), *The Stanford Encyclopedia of Philosophy*, fall 2015 edn. Metaphysics Research Lab, Stanford University.
- Stoljar, Daniel. 2001. Two Conceptions of the Physical. *Philosophy and Phenomenological Research*, **LXIII**, 253–281.
- Stoljar, Daniel. 2006. *Ignorance and Imagination*. Oxford University Press.
- Stoljar, Daniel. 2009. Physicalism. *In*: SEP [n.d.]. Stanford Encyclopedia of Philosophy.
- Stoljar, Daniel. 2015. Russellian Monism and Nagelian Monism. *Pages 324–345 of*: Alter, Torin, & Nagasawa, Yujin (eds), *Consciousness In The Material World*.
- Strawson, Galen. 1997. The Self. *Journal of Consciousness Studies*, **4**, 405–28.
- Strawson, Galen. 2006. Realistic Monism - Why Physicalism Entails Panpsychism. *Journal of Consciousness Studies*, **13**(10-11), 3–31.
- Sturgeon, Scott. 2000. *Matters of Mind: Consciousness, Reason and Nature*. Routledge.