SPINOZA’S ISOLATIONISM

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

Michael Christopher Rauschenbach

Samuel Newlands, Director

Graduate Program in Philosophy

Notre Dame, Indiana

February 2018
SPINOZA`S ISOLATIONISM

Abstract

by

Michael Christopher Rauschenbach

This project coalesces around a simple but highly counterintuitive claim forcefully emphasized by the 17th century Dutch rationalist, Baruch Spinoza. Following others, I call it the explanatory barrier between the attributes (EB), and I suspect it may be true. My aims here are far more modest than establishing the EB`s truth, however, for fairly strong critical consensus in Spinoza scholarship suggests that Spinoza did not have, and that there really could not really be, any good reason for believing it. That claim is certainly false, in both respects. This project is my attempt to uncover what Spinoza`s reasons might have been and to show that they are compelling.

The explanatory barrier between the attributes is the thesis that there are no causal or explanatory connections between the mental and the physical. Actually, it is broader than this, but this framing is a reasonable place to start. In what follows, I establish the centrality of the explanatory barrier within Spinoza`s overall metaphysics. I argue that Spinoza responds not directly to traditional Cartesian skepticism, but to a related, and more interesting explanatory problem previously unappreciated by his commentators. I show this by giving a new reading of Spinoza`s Short Treatise that reconstructs an
argument for the EB as a response to this problem. Spinoza thinks the explanatory barrier’s truth provides the best explanation for our causal knowledge in ordinary cases. Said differently, the explanatory barrier is the best response to a class of skeptical or explanatory problems that appear intractable if the EB is not true.

The dissertation’s latter half uses this argument to defend what I call Spinozistic Isolationism: Spinozism, plus a compelling justification for the EB. I claim that Isolationism strongly undermines interactionist dualism, accommodates the best arguments for physicalism while retaining key advantages over it, and can be interpreted as either a form of or very nearby rival to Russellian panpsychism. I think the argument given here for the explanatory barrier helps justify Spinoza’s basic metaphysical project. So justified, Isolationism deserves serious consideration as the most compelling route towards resolving longstanding debates in the metaphysics of mind.
This is for my family—mom, dad, and brothers—and friends, but especially for Finola Anne.
CONTENTS

Acknowledgments .............................................................................................................. iv

Introduction: The Isolationist's Claim ................................................................................. 1

Chapter 1: The Explanatory Barrier's Role in Spinoza's Metaphysics .............................. 14

Chapter 2: Spinoza and Skepticism ................................................................................... 71

Chapter 3: A Reading of Spinoza's Short Treatise as an Argument for the EB .......... 142

Chapter 4: Spinoza Confronts the Interactionist Dualist ................................................. 206

Chapter 5: Spinoza and Physicalism ............................................................................... 270

Chapter 6: Spinozistic Isolationism in the Contemporary Debate about Consciousness 331

Works Cited..................................................................................................................... 401
ACKNOWLEDGMENTS

I would like to especially thank my director, Samuel Newlands, for his insightful criticism, constructive guidance, and most importantly for his encouragement and patience throughout this project’s development. Additionally, I thank the rest of my dissertation committee—Jeff Speaks, Leopold Stubenberg, and Brian Cutter—for their advice, their time, and their suggestions for ways to make this project better.

I have been very lucky to have had a number of excellent friends here at Notre Dame, and I thank them for their support, their laughter, and their patience. But most importantly, I thank my family, as well as my best friend and wife, Finola Anne Prendergast, for her brilliant feedback, her lovely laugh, and her sure sense of moral purpose. I have learned, and will continue to learn, so much from her example.
INTRODUCTION: THE ISOLATIONIST’S CLAIM

“Once you eliminate the impossible, whatever remains, no matter how improbable, must be the truth.”

-Sherlock Holmes

This project coalesces around a simple but highly counterintuitive claim whose importance was forcefully emphasized by the 17th century Dutch rationalist, Baruch Spinoza. Following others, I call it the explanatory barrier between the attributes (EB), and I suspect it may be true. My aims here are far more modest than establishing the EB’s truth, however, for fairly strong critical consensus in Spinoza scholarship suggests that Spinoza did not have and that there really could not be any good reason for believing it. That claim is certainly false, in both respects, and this project is my attempt to uncover what Spinoza’s reasons might have been and to show that they are compelling.

The explanatory barrier between the attributes is the counterintuitive thesis that there are no causal or explanatory connections between the mental and the physical.1 Actually, it is broader than this, for reasons that are particular to Spinoza, but he would

1 Sometimes I freely interchange ‘conceptual’ with ‘explanatory’, and this is liable to cause some readers anxiety. Unfortunately, I won’t have the space to justify this practice here, but I hope that one can at least understand the claims being made in cases where I do so.

2 Oppy, “Against Idealism,” characterizes the claim that mentality is late and local as a key tenet of naturalism, but as will become clear, Spinoza and many other contemporary philosophers are unlikely to find this at all persuasive as a binding precondition for being naturalistic.

3 As it turns out, on Spinoza’s view this makes most standard identity claims ambiguous, such that
certainly endorse this framing and it’s a reasonable place to start. Some pedestrian examples, drawn from the realm of what I later discuss under the heading of folk psychology, might help. My stomach was empty, so I thought about food and decided to leave the party. False. Tom is upset because he has not taken his medication today. Not true. I imagined meeting my wife for dinner, desired that this occur at 7pm, and this caused me to call her and make it happen. Nope.

The EB floats free of any underlying ontology, though of course it sits better with some than others. So, for example, the EB entails that no physical events cause or explain any mental events, and vice versa. In property talk: no mental properties are causally or explanatorily relevant to any physical properties, and the other way around. For those drawn to the classical vocabulary of substances: mental and physical substances do not causally interact. No mental substance explains the existence or properties of any physical substance, and of course conversely.

It will turn out that the explanatory barrier, properly interpreted, is substantially less wild-eyed than this brief introduction makes it seem. Spinoza was a genius, not insane, however often those two qualities might seem to coexist in lived experience. In addition to philosophers, poets like Goethe and scientists like Einstein have all found Spinoza’s underlying vision of reality deeply attractive, and by the end of the present work I hope to have added in some small way to the store of reasons this is so. Indeed, as it happens Spinoza wants to combine the EB with a far more common claim, which nonetheless in the present context seems even crazier than the first. Spinoza affirms that the mental is the physical, that mental events are physical events, and that the mental substance is the physical substance, and this despite there being no causal or explanatory
connections between mental and physical properties, events, or substances. All this has
counterintuitive temporal implications worth noting. That the mental is the physical
entails for Spinoza that mentality exists at the beginning of the universe. It is not,
contrary both to commonsense intuition and to scientific orthodoxy (on one
interpretation), a late and local phenomenon. In this case, two wild and wacky claims
combined help to bring Spinoza back to earth. Still, to appreciate the EB’s significance,
one must first revel a bit in its weirdness.

Before briefly cataloguing the contents of the present work, I want to introduce,
briefly defend, and then mostly use without further discussion a set of distinctions critical
for understanding Spinoza’s explanatory barrier between the attributes. Making sense of
the claims that, on one hand, there are no causal or explanatory connections between the
mental and the physical, and, on the other, that the mental is the physical, might seem to
court paradox. Quite related claims pose no such problem though. To take the stock
example: Lois believes Superman can fly, Clark Kent is Superman, but Lois does not
believe that Clark Kent can fly. The explanation: ‘believes’ creates referentially opaque,
intensional (or hyperintensional) contexts. Spinoza’s claim is that both ‘causes’ and
‘explains’ behave similarly, creating referentially opaque contexts in their wake.

2 Oppy, “Against Idealism,” characterizes the claim that mentality is late and local as a key tenet
of naturalism, but as will become clear, Spinoza and many other contemporary philosophers are unlikely to
find this at all persuasive as a binding precondition for being naturalistic.

3 As it turns out, on Spinoza’s view this makes most standard identity claims ambiguous, such that
referential opacity operates in them too. So he endorses, for example, that no mental thing has any physical
effects, that some mental thing m1 is identical to some physical thing p1, that p1 causes physical effect p2,
but that, by the first claim, m1 does not cause p2. See Newlands, “Spinoza’s Modal Metaphysics,” for
much more detail on this odd facet of Spinoza’s view.
More concretely, I want to follow David Nolan in speaking both of extensional, intensional, and hyperintensional positions in sentences, and of phenomena with these characteristics. However, as Nolan mentions, issues here are very tricky, and orthogonal to most of my present concerns. Nonetheless, I need to make Spinoza’s position explicit before mostly ignoring the subtleties in what follows. Nolan defines matters thus:

A position in a sentence is *extensional* if other expressions with the same extension can be substituted into that position without changing the truth value of that sentence. On standard theories, the extension of a name or description is the object designated, the extension of a predicate is the set of (actual) objects to which the predicate applies; the extension of a sentence is its truth value…A position in a sentence is *intensional*, in my sense, if it is not extensional and expressions that are necessarily co-extensive are freely substitutable in that position without change in truth value. A position in a sentence is *hyperintensional* provided it is non-extensional and also not even intensional: even substitutions of necessary equivalents is not guaranteed to preserve truth value.4

On Nolan’s characterization, the set of sentences concerning Lois’s beliefs about Clark and Superman are hyperintensional, at least on the standard understanding of the necessity of identity, according to which necessarily, Clark Kent=Superman. Though ‘Clark Kent’ and ‘Superman’ necessarily refer to the same object, that dashing reporter *cum* superhero, they are not freely substitutable into “Lois believes…can fly” such that truth is preserved.

A flat-footed extension of Nolan’s definitions makes *believing* as a phenomenon, rather than as an operator that creates positions in sentences, a creator of hyperintensional

---

4 Nolan, “Hyperintensional Metaphysics,” 151
As mentioned, Spinoza thinks that causing and explaining are similarly non-extensional. To borrow an example from Charles Jarrett, that explanation is referentially opaque and non-extensional just means:

You can explain (partly) why event d occurred without explaining why event e occurred, even when d and e are one event. An example (derived from Nagel) will illustrate this. Suppose that George is kicked by a certain horse and that we explain this by saying that George foolishly pulled the horse’s tail while standing behind it. The horse in question, however, was at that time the largest mammal in Berkeley, partly because Bimbo (who is normally in Berkeley) was out of town. Saying that Bimbo was out of town does not explain (even partly) George’s being kicked by that horse, but it does explain (partly) his being kicked by the largest mammal in Berkeley. We feel impelled to say, however, that George’s being kicked by that horse is his being kicked by the largest mammal in Berkeley.

Jarrett lays out his account of explanation as referentially opaque in the service of ultimately arguing that causation is too, with the goal of defending Spinoza’s explanatory barrier between the attributes in a very different way than what I pursue here. So he provides a second, causal example. Suppose George hammers a piece of gold into a wedding band at a certain time. Then George causes a wedding band to exist at that time, but does not cause a piece of gold to exist at that time, even though there is just one thing, the gold wedding band, that exists at that time. Causation too, Jarrett thinks, is referentially opaque in the way that explanation and belief are uncontroversially taken to be.

---

5 Actually, I will be flat-footed in what follows in an additional way, since I will also mostly collapse Nolan’s distinction between intensional and hyperintensional phenomena, contrasting both as a unified group with extensional phenomena.

6 Jarrett, “Materialism,” 458-459

7 In addition to Jarrett, “Materialism,” see also Jarrett, “Spinoza’s Denial of Mind-Body Interaction and the Explanation of Human Action,” and for dissent from the view that causation is referentially opaque in this way, see Davidson, “Spinoza’s Causal Theory of the Affects,” and “Mental Events.”
Spinoza’s explanatory barrier between the attributes gains substantially in plausibility if one finds the preceding examples compelling, but my project, for reasons of space, must ignore the particulars of this debate in the philosophy of language and of modality. This represents a significant difference between Jarrett’s strategy for motivating the EB and my own. Jarrett argues from the plausibility of such examples to the possibility of causation’s being a referentially opaque phenomenon. He then uses this possibility to sketch Spinoza’s distinctive metaphysics and to motivate it relative to rival positions, classifying Spinozism as the best way of being a materialist. I proceed in a global, almost certainly riskier, but I think more interesting way.

I attack the central issues from two separate directions, a strategy reflected in the organization of the chapters here. My central thesis is that Spinoza introduces the explanatory barrier to solve an explanatory problem, which bears substantive affinities with well-known skeptical problems like Kripke’s rule-following paradox or Goodman’s grue problem, that interactionist views, those that reject the EB, cannot solve.8 These two lines of attack converge on chapter three, which serves as a hinge joint between them. This all works as follows.

The first part of the project grounds discussion of the explanatory barrier in Spinoza’s philosophy and the surrounding scholarship. In chapter one, I argue that all of Spinoza’s distinctive metaphysical commitments—Substance Monism, Mind-Body Parallelism and ultimately Identity—and his solutions to central natural objections rely on his having some defense of the EB. I further argue that none of Spinoza’s interpreters

---

8 Though the problem bears affinities with these well-known skeptical problems, I do not discuss Spinoza’s problem with detailed reference to these contemporary issues. Again, this is a question of priorities. I make clear in what follows where mine lie.
have provided this defense, and that many of them despair of providing some reasonable defense in principle. I conclude that this lacuna poses a serious threat to the philosophical interest of Spinozism.

In chapter two, I argue that Spinoza does not have a successful argument against the skeptic as defined in most of the literature on Spinoza and skepticism. These issues might initially seem tangential, but they are not. My central claim is that the EB solves an explanatory problem, structurally similar to many well-known skeptical problems, which rival positions cannot solve. But skeptical problems come in many varieties, and the problem I discuss in chapter three is distinct from others in the literature, either on Spinoza or skepticism. Still, the goal of chapter two is to moderate the enthusiasm of those who argue that Spinoza already provides a compelling argument against broadly Cartesian skepticism. If such skepticism were defeated we might seem to be on quite stable epistemic ground already, perhaps minimizing the interest of those issues most relevant to Spinoza’s EB. But chapter two has an additional purpose, for some of the most promising (though ultimately unsuccessful) ways of interpreting Spinoza’s argument against the Cartesian skepticism are structurally similar to the argument for the EB I present in chapter three. This is useful, since otherwise my argument for the EB on Spinoza’s behalf, which is pretty different from anything found in the Spinoza literature, might seem to come from nowhere. Novelty is a virtue only when appropriately grounded in the more familiar, and the discussion of Spinoza’s response to skepticism provides that foundation.

9 Though, again, see the beginning of chapter two for some clarification on what sort of skeptic I have in mind.
Chapter three is the heart of the project. There, I accomplish two tasks. First, I provide a new reading of key sections of Spinoza’s *Short Treatise*, a work that precedes the *Ethics* for which he is justifiably more famous, but which offers essential insights into Spinoza’s own philosophical development. This new interpretation demonstrates how Spinoza grapples directly with the explanatory problem afflicting rival, interactionist positions—those that lack the EB applied to either events, substances, or properties—that I take as motivational for the introduction of the EB that occurs only later, in the *Ethics*. Second, I lay out the explanatory problem in more neutral terms and provide what I call, following Ameriks, a regressive transcendental argument for the explanatory barrier between the attributes on Spinoza’s behalf.

The argument’s regressive or modest character is crucial. Unlike those interpreters I discuss in chapter two, I read Spinoza’s defense of the EB as a challenge to other non-skeptics, like interactionist dualists, physicalists, etc. Their task is to explain how their explicit anti-skeptical commitments are compatible with other aspects of their position and thereby to justify their allegiance to them. Spinoza’s insight is that they are quite hard pressed to do so. This strategy allows me to define and defend a position I call

---

10 Cashing out the EB in terms of either events, properties, or substances gives a hint as to how the EB will be relevant to the dominant view in the metaphysics of mind, physicalism, since most physicalists will still be working with a distinction between mental states and physical states, or mental properties and physical properties, or mental descriptions and physical descriptions, etc. The EB precludes crossover between any of these, and thus serves as a challenge to the physicalist as well, though it’s a challenge of a slightly different sort than that it presents to the interactionist dualist, as chapters 4-5 make clear.

11 See Ameriks, “Kant’s Transcendental Deduction as a Regressive Argument,” for the canonical interpretation of Kant that guides my choice of nomenclature. As I make clear in chapter three, the arguments are not directly analogous, but the comparison is nonetheless fruitful, I think.

12 I take it that a modest, regressive transcendental argument is one that, whatever its other characteristics, proceeds as if skepticism is false, by assuming that we have at least some empirical knowledge. Various ways of fleshing out the distinctive attributes of such arguments are to be found in Ameriks, “Kant’s Transcendental Deduction as a Regressive Argument,” Brueckner, “Modest Transcendental Arguments,” and Stern, “Transcendental Arguments: A Plea for Modesty.”
Spinozistic Isolationism, which conjoins classic Spinozism with a ‘transcendental’ defense of the EB. Since some such defense is required to make viable adherence to Spinoza’s central metaphysical claims, as I argue in chapter one, Spinozistic Isolationism is the viable form of Spinoza’s philosophy. However, this is not self-congratulatory hyperbole, since, as I make clear in chapter three, I think there are good reasons for thinking that Spinoza might himself have been motivated by considerations very similar to those I make explicit in this project.

Chapters four through six turn from the historical to the contemporary and seek to use Spinozistic Isolationism as a fruitful contributor to ongoing debates in the philosophy of mind. In chapter four, I argue that interactionist dualism is particularly ill-suited to respond to the challenge posed by the explanatory problem that motivates the EB’s introduction. In the process, my discussion of Spinoza makes contact both with work on the so-called ‘causal pairing problem’, as well as with work on folk psychological explanations and their relationship to metaphysics and the philosophy of mind. Chapter five shows the form that Spinozistic Isolationism’s engagement with contemporary physicalism takes. There, I argue for two main conclusions. First, I argue that Spinoza can accept the most plausible version of the main (compelling) arguments for physicalism, and thus that physicalists can only rule out Spinoza’s position by definitional fiat. Second, I argue that Spinozistic Isolationism has several significant advantages over physicalism, which has faced a variety of objections concerning, among others, its very definition, its ability to avoid collapsing into a wholesale eliminativism about mental states, and its naturalistic account of intentional content. Spinozistic Isolationism avoids these problems in an elegant way, which I survey in chapter five.
Finally, chapter six injects Spinozistic Isolationism into contemporary debates about consciousness. Specifically, I argue that Spinozistic Isolationism can be fruitfully interpreted either as a form of, or as a very nearby rival to, ‘Panpsychist Russellian Monism’ as it is frequently discussed today. I then claim that it provides more principled motivation for several major arguments in the literature, including David Chalmers’ Hegelian argument for Russellian monism, Galen Strawson’s no-brute emergence argument for panpsychism, and the arguments of Goff, Nagasawa, and Wager for a version of Russellian Monism they call Priority Cosmopsychism. This requires, however, that I discuss whether Panpsychist Russellian Monism must be formulated, as many do following Chalmers, Alter, and Nagasawa, in terms of grounding relations between the phenomenal and the physical (which would violate the EB), or whether this is an oddity of contemporary formulations inessential to the view. I argue that it is the latter. Lastly, I suggest that Spinozistic Isolationism can help solve the combination problem(s) for panpsychism.

To forestall confusion until Spinozistic Isolationism is more precisely described in chapter 6, I want to try to situate the view relative to possible rivals using contemporary terminology. This, it turns out, is not that easy to do. Spinozistic Isolationism is a token identity theory of mental states or events and physical states or events. From the Ethics’

---

13 See Consciousness in the Physical World: Perspectives on Russellian Monism for a masterful collection of much of the most interesting recent work on this family of positions.

14 The relevant citations, discussed in more detail in chapter six, are Chalmers, “Panpsychism and Panprotopsychism”; Strawson, Consciousness and its Place in Nature; Goff, “Cosmopsychism, Micropsychism, and the Grounding Relation”; and Nagasawa and Wager, “Panpsychism and Priority Cosmopsychism.”

15 As mentioned in Chapter 5, Spinoza’s nominalism substantially complicates the statements of his position that immediately follow. I mostly ignore these complications, though occasionally I will point
2p7: the extended mode and the idea of that mode are one and the same thing. Its key innovation is the explanatory barrier, which precludes any causal, conceptual, or explanatory relations between any thinking things, properties, states, or events, conceived as such, and any extended things, properties, states, or events, conceived as such, where truths about how things are conceived are objective and independent of particular acts of thinking. Consequently, Isolationists reject any cross-attribute (i.e. physical-mental or vice versa) grounding claims, since grounding claims express dependence relations that are either themselves explanatory or entail explanations across attributes. From another angle, we can start with what Isolationism is not. Isolationism is not Davidson’s anomalous monism, for reasons Davidson acknowledges. Most importantly, Isolationism requires that causal contexts be opaque, whereas Davidson explicitly affirms only explanatory, but not causal, opacity.\footnote{Davidson, “Spinoza’s Causal Theory of the Affects,” fn. 13, pp. 110-111; these issues are discussed at greater (though perhaps not enough) length in the Introduction.}

Isolationism is also not epiphenomenalist property dualism, since epiphenomenalists characteristically affirm that mental events are caused by physical events, which themselves are causally inefficacious in the physical realm.\footnote{Robinson, “Epiphenomenalism,” Intro} By contrast, Isolationism broaches no asymmetry between the mental and physical. They are equally fundamental and equally causally efficacious, as the Isolationist’s commitment to token identity theory makes clear. Finally, Isolationism is also not, or not just, a garden-variety
form of parallelist property dualism or pre-established harmony, for four reasons.\textsuperscript{18} Leibniz, the most famous defender of parallelism, seems to think that bodies result, or are grounded in, some features of monadic minds and their ideas or perceptions, even if this grounding or resulting relation remains somewhat unclear in Leibniz’s work.\textsuperscript{19} This, of course, violates the explanatory barrier as understood it here. Second, Spinoza is a property pluralist—there are infinitely many attributes, rather than the typical parallelist two. Third, both Spinoza’s reasons for affirming infinitely many attributes, and his reasons for denying causal and explanatory relations among any two sets of attributes, are wholly general, independent of particular features of Thought or Extension. They are thus not held hostage to particular Leibnizian arguments about the impossibility of either inter-monadic causation or of explaining perception in mechanical terms (as in Leibniz’s mill). Fourth and finally, parallelist pre-established harmony is typically taken as a rival to, rather than a version of, token identity theory, and since Spinoza certainly affirms the latter, he cannot be \textit{just} a garden-variety proponent of the former.

The explanatory barrier, which claims that there are no causal or explanatory connections between the mental and physical, seems highly counterintuitive. But I ultimately conclude that Spinozistic Isolationism, a position grounded in Spinoza’s own texts and ways of thinking, offers the best alternative of the most frequently discussed

\textsuperscript{18} I also strongly endorse Strawson’s annoyance, voiced in \textit{Consciousness and its Place in Nature}, 242, fn. 120, with contemporary philosophers’ attempts to map views that either explicitly accept a non-traditional subject/property distinction, as does Strawson, or reject realist views of properties in favor of nominalism, both of which Spinoza arguably does. So if one can only understand Spinoza’s view within a property dualist framework, then I judge this preferable to not grappling with his views at all. It’s just important to recognize then that probably Spinoza’s view of properties and their relations to subjects is either strictly nominalist or much closer to Strawson’s view that there is no real distinction between subjects and their properties in many cases (certainly as regards finite modes of substance and their properties, and also potentially as regards the relationship between substance and its attributes).

\textsuperscript{19} Adams, \textit{Leibniz: Determinist, Theist, Idealist}, Ch. 9
options in contemporary metaphysics of mind. It offers compelling arguments against, and substantial advantages over, interactionist dualism, physicalism, and other forms of panpsychism. I do not pretend to have shown any of this to a level Sherlock Holmes would find acceptable. But I hope to have at least prompted the thought that it is these contemporary positions—various dualisms, physicalisms, and other panpsychisms—that form the set of impossible theses worth eliminating in favor of the improbable truth that Spinoza just might have seen.
1.1 Can the Explanatory Barrier Do What It Must?

Spinoza introduces the barrier between the attributes (EB) to solve an *explanatory* problem animating the opening propositions of the *Ethics*. Before developing this claim, I want to demonstrate the importance of the EB in Spinoza’s overall system, and to argue that even Spinoza’s most sympathetic interpreters offer no substantive defense of it. I then emphasize that ubiquitous reliance on an undefended (and, according to some, indefensible!) principle seriously undermines the various Spinozistic projects his recent interpreters have presented. And, to the extent that these presentations are accurate, Spinozism itself is seriously undermined as well. These are the tasks of this first chapter.

The EB plays a pivotal role both in deriving three of the main theses of Parts I-II of the *Ethics*, and in solving two additional interpretive puzzles in Spinoza scholarship. More specifically, the EB is instrumental in deriving Spinoza’s (1) **Substance Monism**, (2) Attribute and Mode Parallelism (hereafter, **Parallelism**), and (3) **Mind-Body Identity**. The explanatory barrier also addresses (4) **Idealism**—puzzles about seemingly idealistic strains in Spinoza’s thought, oftentimes associated with his definition of ‘attribute’ in 1d4. Finally, the barrier helps resolve (5) **Ignorance**—Tschirnhaus’s and later Bennett’s challenge concerning humanity’s inability to epistemically access more
than two of the attributes. I discuss these five problems in turn. Besides being
philosophically interesting as such, each thesis or puzzle resolution requires a well-
motivated defense of the EB. For either the EB or some proposition entailed by it plays a
major role in deriving the above theses. But many interpreters remain skeptical that any
such defense exists. Hopefully, then, a clear presentation of the EB’s central importance
and ubiquity within Spinoza’s system stimulates interest in a Spinozistic defense of its
well foundedness, which I provide in later chapters.

1.2 An Introduction to the EB

Spinoza endorses a wild and complicated ethical and metaphysical system, at the
heart of which is what I will call the explanatory barrier between the attributes (EB).
This is Spinoza’s shocking claim that the fundamental ways of being an independent
substance, including at least being thinking and being extended, are entirely conceptually,
explanatorily, and causally isolated from each other. This radical causal, explanatory and
conceptual isolation obtains despite an extremely tight relationship between these
fundamental ways of being, and the substance whose reality they express. Spinoza
characterizes this in 2p7s, saying “the thinking substance and the extended substance are
one and the same substance, which is now comprehended under this attribute, now under
that.”

______________________________

20 In what follows, I always underline within a quotation to signify my emphasis of some aspect of
the quoted material, whereas italicized portions are emphasized in the original text. Henceforth, quotations
from the Ethics are given by book, proposition or demonstration, etc., and number, as in 1p4, 1p10s, etc, in
the traditional manner.
Spinoza’s outlier status appears most clearly by contrast with more familiar ways of reasoning. Consider, for example, the physically extended world and the mental world. A natural thought is to regard the mental world as produced, in some complicated way, by the workings of the physical world. This seems to hold both locally (i.e. my thoughts, though modes rather than attributes, for Spinoza, are in some sense produced by my brain) but more relevantly globally. That is, Thought seems to have arisen after a suitably long period during which evolution and other physical processes operated within the physical world. In this loose sense, one might think, what it is to be Thinking is produced by features of Extension. But Spinoza explicitly rules this out.

Spinoza’s view of the relationship between the attributes of Thought and Extension is strange from multiple perspectives. A nice way to appreciate the oddity is to state the view temporally. For example, the physicalist thinks that there was a time when there were physically extended things, but no thoughts or thinkers. By contrast, some theists believe that a Divine thinking (and non-extended) being preexists, ontologically and/or temporally, the physical world it creates. Spinoza’s views preclude, and thus alienate, both camps. This odd view, its consequences and Spinoza’s ultimate justification for it, will be the focus of this study. Beginning to appreciate its subtleties requires, however, some background in Spinoza’s metaphysics.

Spinoza’s metaphysics is a three-legged stool of substance, attributes, and modes. Substance is what is in itself and conceived through itself (1d3), and attributes are initially defined as “what the intellect perceives of substance, as constituting its

---

21 See Melamed, “The Building Blocks of Spinoza’s Metaphysics,” Curley, Behind the Geometrical Method, Bennett, A Study of Spinoza’s Ethics, and others, for various overviews of this ontological tripartition.
The modes, by contrast, are “the affections of a substance, or that which is in another through which it is also conceived” (1d5). For Spinoza, there is just a single substance, God or Nature. Unlike Cartesian or classical views, individual human minds and bodies, rocks, trees, and sports cars are all merely modes of God, or determinate expressions of God, under some attribute, or determinable way of being. They are not substances.

But what is the relation between Substance and its attributes? This is significantly more complicated. In 1p4, Spinoza seems to assert an extremely tight relation between substance and attributes, namely, identity! He says: Therefore, there is nothing…except substances, or what is the same (by D4), their attributes, and their affections.” In a letter to Simon De Vries, Spinoza writes: “I understand the same by attribute [as by substance] except that it is called attribute in relation to the intellect, which attributes such and such a definite nature to substance.” In both, Spinoza equates attributes with substance in some way, although the letter to De Vries, as well as the definition in 1d4, he adds a close connection to the intellect. This becomes relevant for us later.

For now, simply keep in mind that Spinoza employs one classical concept of substance as independent in various ways—self-conceived, self-inhering—but subversively applies it only to God. God’s nature or essence is expressed through attributes, which are intimate, maximally determinable ways of being a substance, the

The relevance of the intellect will be discussed in section 1.6.

See Newlands, Reconceiving Spinoza, Ch. 1, for a defense of interpreting attributes and modes within this framework, as well as Hubner, “The Trouble With Feelings, or Spinoza on the Identity of Power and Essence,” for a similar gloss, albeit with a focus on essences and affects.

I cite from Curley’s translation. Keizer, “Spinoza’s Definition of Attribute,” emphasizes this point.
determinate expressions of which are modes. Objects in the world—those rocks, trees, bodies, minds, and sports cars—are “modes by which God’s attributes are expressed in a certain and determinate way” (1p25c).

To better appreciate Spinoza’s use of ‘attributes’ and ‘modes’, consider what particular thoughts or physically extended things share. My thought of pizza and your thought of your dog Skippy are each about something (pizza, Skippy). Thoughts also, in most cases, have unique bearers (i.e. I can’t have your thought of Skippy, even if, should I see him, I can think of him too). The fire hydrant outside and the mountains the next town over are each physically extended, made of similar sorts of particles and subject, we suppose, to similar sorts of physical laws within a larger web of extended things. A fire hydrant is just a particular way of being Extended (a mode, in Spinoza’s terms), as is a mountain. Extension is just the determinable of which fire hydrants and mountains and quasars are all determinants. The same analysis, in determinable and determinant terms, holds for Thought. That is, within this tripartite ontology the brown color of my hair, my mind, and even my entire human organism are all, equally, mere modes of God, certain determinate expressions of his nature as Thinking or Extended.

We are now in a position to meet the explanatory barrier between the attributes (EB), in Spinoza’s words. He describes it in 1p10, as follows: “Each attribute of a substance must be conceived through itself.” His demonstration then reads: “For an attribute is what the intellect perceives concerning substance, as constituting its essence (by D4); so (by D3) it must be conceived through itself, q.e.d.” But what exactly does 1p10 mean, and why does Spinoza believe it? Fully unpacking the content of this claim

25 This greatly simplifies thorny issues about propositions. But Spinoza would certainly accept it.
will be a gradual process, but we might start by considering Spinoza’s demonstration of 1p10. Unfortunately, his demonstration is not confidence inspiring. It employs two definitions to reach its conclusion, arguing first that an attribute is what is conceived of a substance, as constituting its essence, and that therefore, it must be conceived through itself.

This is puzzling, however, for 1d3 defines *substance* as self-inhering, self-conceived, and conceptually self-contained. As Spinoza puts it: “By substance I understand what is in itself and is conceived through itself, that is, that whose concept does not require the concept of another thing, from which it must be formed.” And 1d1 connects self-conception explicitly to self-causation—“by cause of itself I understand that whose essence involves existence, or that whose nature cannot be conceived except as existing.” But why should constraints on self-conception, self-inherence, and self-causation apply directly to the attributes of substance, rather than only to substance itself? One tempting but incomplete answer is that since, as we have already seen, Spinoza sometimes seems to equate substance and its attributes, of course constraints on the former should apply equally to the latter.

This initial answer fails to do justice to the vast array of uses to which Spinoza puts 1p10, few of which perform functions substantially similar to those of 1d3 and 1d5. As mentioned, he uses it in his proof of Substance Monism, Parallelism and Mind-Body Identity, and others. A reading of the EB and its demonstration on which it was a mere restatement of 1d3’s and 1d1’s constraints on self-inherence, self-conception, and self-causation for substances, which just are their attributes, makes this multiplicity of uses puzzling. Furthermore, Spinoza’s paradigmatic attributes, or maximally determinable
basic ways of being, are Thought and Extension, as they were for Descartes. But whereas Descartes sees thinking and extended things as causally interactive, Spinoza demands, as we will see in more detail, that Thought and Extension be internally self-caused.

Thinking modes only cause thinking effects, and extended modes only cause extended effects, despite the identity of some modes of thinking and extension. This surprising consequence of 1p10 does not, at least at first glance, sound merely like a corollary of the close connection between substance and its attributes. Paul Eisenberg emphasizes both the point about the different contents of Spinoza’s ideas of substance and attribute, as well as the difficulty of maintaining a substance-attribute relation of identity, in arguing:

> If (e.g. Thought), as an attribute of God, is supposed to be identical to the essence of substance but also Extension is identical to the essence of substance, then two things which, according to Spinoza himself, are supposed to be irreducibly different from one another (namely, Thought and Extension) turn out to be identical to one and the same thing and, hence, identical to each other…[Furthermore] it is hard (for me) to see why Spinoza gives the emphasis that he does to the notion of substance if in fact he believes it has the same content as does the notion of attribute(s).²⁶

Spinoza’s employment of the explanatory barrier between the attributes, then, seems hardly justified by the demonstration he gives for it. Nor, it seems, can Spinoza’s casual identification of substance and its attributes be taken completely at face value. Both points bear keeping in mind as we try to understand 1p10’s meaning more fully.

A natural place to look for greater understanding is Spinoza’s elaboration of his demonstration in the scholium to 1p10, where he argues:

> From these propositions it is evident that although two attributes may be conceived to be really distinct (i.e. one may be conceived without the aid of the other), we still cannot infer from that that they constitute two different beings, or

---

²⁶ Eisenberg, “On the Attributes and their Alleged Independence of One Another,” 4, 9
two different substances. For it is of the nature of a substance that each of its attributes is conceived through itself, since all the attributes it has have always been in it together, and one could not be produced by another, but each expresses the reality or being of substance.

Spinoza’s reasoning seems to run as follows. Two attributes may be conceived as really distinct, in the Cartesian sense that one is conceivable without the other. But one cannot infer from this that they constitute the essence of two different beings or substances. Why? Spinoza argues, most broadly, that whether attributes and their bearers have a one-one or many-one relationship, each attribute must still be wholly self-conceived. Descartes argues that each substance has only a single attribute, and that it is incoherent to posit multiple attributes per substance.27 That is, Descartes always maintains a one-one relationship between attributes and attribute bearers. Alternatively, Spinoza ascribes infinitely many attributes to substance. Regardless, Spinoza thinks that there are three reasons that attributes must be wholly self-conceived. Ultimately, these three reasons show why Descartes incorrectly insists that one substance cannot have multiple attributes, but this further consequence comes in addition to the independent arguments for attribute conceptual independence Spinoza gives.

Spinoza’s three reasons can be found in the scholium to 1p10. First, however many attributes a substance has, it has always had that many. This seems to follow from 1p8schol, where Spinoza argues that every substance is necessarily infinite, and that being infinite is an “absolute affirmation of the existence of some nature.” Were a substance to acquire additional attributes, it would have been limited with respect to its expression of that nature before that point. Second, each attribute expresses the being or

27 CSM I/298, AT VIIIB/349-350—I discuss this passage in greater length below.
reality of substance. This just restates 1d4, the definition of attribute, as essence constituting. Finally, no attribute can produce another attribute. This final point is more interesting than the previous two, for it introduces potentially causal language into Spinoza’s stock of considerations. Moreover, it raises the question of how causal interaction between modes of differing attributes would work in light of one attribute’s inability to produce another.

The three reasons given above converge on a single thought. Spinoza thinks that if substances could gain or lose attributes, or that if one attribute were to be produced by another, *substance itself* would somehow violate the constraint on being self-conceived. More generally, self-conception of attributes is necessary to preserve the same property for substance. For consider two distinct substances, a Thinking Substance and an Extended Substance. If something like the relationship between the physical world and thought sketched earlier is correct, then the essence of Thinking Substances arises from or is caused by some features of Extension. But then the concept of the Thinking Substance does require that of another, from which it must be formed. This follows, at least, if one supposes that the causes of a thing play at least some role in the concept of that thing’s being or reality. But then the Thinking Substance, by 1d3, is really no substance at all!

This supposition about the relationship between a thing’s causes and the concept of that thing is not an idle one. Spinoza argues explicitly that the causes of a thing play a role in the concept of that thing’s being or reality. In fact, Spinoza’s claim is much stronger than this—causes just are reasons. This comes out clearly in 1p11dem, where Spinoza argues:
For each thing there must be assigned a cause, or reason, as much for its existence as for its nonexistence. For example, if a triangle exists, there must be a reason or cause why it exists; but if it does not exist, there must also be a reason or cause which prevents it from existing, or which takes its existence away.

Spinoza’s frequent interchanging of ‘reason’ and ‘cause’ here, as in other passages, suggests a close connection between explanation and causation. He even claims, in a later letter to Tschirnhaus (Ep 60): “In order that I may know which out of many ideas of a thing will enable all the properties of an object to be deduced, I follow this one rule, that the idea or definition of the thing should express its efficient cause.” Here, citing a thing’s efficient cause is taken to be sufficient for deducing all the properties of an object!

But Spinoza endorses even more than this. I take it that he defends, at minimum, a view I will call, following Newlands, the Co-extensive Thesis:

\[
\text{Necessarily, for all } x \text{ and } y, \ x \text{ is conceived through } y \iff y \text{ causes } x \iff x \text{ is in } y \iff x \text{ requires } y \text{ to exist.} \tag{28}
\]

Of course, not every interpreter endorses all of these bi-conditionals, and some, including Newlands himself, defend stronger monistic views. On these stronger interpretations, the bi-conditionals hold because the relations of being conceived through, causing, inhering, and ontologically depending are all identical. That is, Spinoza employs only a single relation of ontological dependence. Given interpretive disagreement on

---

28 Newlands, *Reconceiving Spinoza*, Ch. 3

29 Newlands, *Reconceiving Spinoza*, Ch. 3 See also Laerke, “Spinoza’s Cosmological Argument in the Ethics,” and Garrett, “Spinoza’s Conatus Argument.” To give just a hint of the textual evidence for this thesis: consider the following passages. For the relationship between causation and ontological dependence, see 1p8 and 1a3. The relation between causation and explanation appears most clearly in 1p11dem, but also arises in 1p8s2. For the claim that if something causes something else, it explains it, see 1p3d. Finally, Spinoza argues that if x is conceived through y, then y causes x, in 1a4. The converse can be inferred from
this issue, plus the wide literature already devoted to this question, I will not defend a view in detail here. But I will try to note when some feature of my arguments depends on a particular position.

The upshot of the foregoing discussion is this. Given the Co-Extensive Thesis, as well as the three reasons explored previously, Spinoza argues that there are *independent* reasons for thinking that all attributes are conceptually and causally isolated from each other. Denying the conceptual isolation of attributes directly endangers the conceptual independence of substance itself. Moreover, Spinoza argues that this causal and conceptual isolation holds independently of whether he or Descartes is correct in affirming a many-one or one-one relation, respectively, between attributes and their bearers.

But, of course, once Spinoza has 1p10 in hand, as well as a few other key propositions, he can argue for the anti-Cartesian conclusion of 1p14: Except God, no substance can be or be conceived (more about this argument anon). This rules out causal interaction between any substances, and in particular between Cartesian minds and bodies, because multiple substances simply do not exist! Crucially, this does not amount to Spinoza *assuming*, at this juncture, that his explanation is preferable. He has not simply helped himself to the claim that Thought and Extension are attributes of the single substance but must be conceived without the other for independent reasons. Rather, Spinoza *concludes* this, later, on the basis of his argument for substance monism. This

---

how Spinoza uses 1a4, as in 1p25d and 1p18d. I am indebted to Newlands, *Reconceiving Spinoza*, for helpfully providing citations and interpretive guidance on this important point of controversy.
larger argument relies on 1p10, along with several other key claims, to move beyond what Spinoza could licitly assume here.

Of course, one might object in various ways at this point. If one takes seriously the physical world’s relationship to the mind, Spinoza’s considerations concerning Thought and Extension might not at all seem to justify inferring that each attribute is conceived through itself. Rather, they might suggest that Thought is, despite appearances, no attribute at all. Some sort of physicalism would then result. This becomes especially compelling since Spinoza’s demonstration that Thought is an attribute in 2p1 is less than maximally convincing. I will address this physicalist challenge at length in later chapters. Or, one might find 2p2’s (non) demonstration that Extension is an attribute insufficient (since it just claims to be identical to 2p1’s), and infer idealism. In any case, the considerations adduced above do not conclusively show that attributes conceived as really distinct can be thought, without inconsistency, both to be really attributes and also attributes of the same substance. Yet, further reflection on Descartes prompts perhaps the strongest objections to Spinoza’s view. So before considering the EB’s role in deriving the 5 key theses or puzzles with which I began this chapter, I want to elaborate on what some Cartesian objections can and cannot tell us about Spinoza’s overall argument for the EB in the Ethics. I argue that these Cartesian objections are ultimately inconclusive, but that they point to the limits of Spinoza’s explicit argument for the EB, and make more urgent some general, principled defense of it on Spinoza’s behalf.
1.3 Cartesian Objections to Spinoza’s Understanding of Attributes

Descartes, like Spinoza, argues for the strong conceptual independence of Thought and Extension, such that “the concept of the one is not contained in the concept of the other.” Unlike Spinoza, however, Descartes reasons from attribute to substance plurality, citing the impossibility of one subject’s having two different natures:

As for the attributes which so constitute the natures of things, it cannot be said that those which are different, and such that the concept of the one is not contained in the concept of the other, are present together in one and the same subject; for that would be equivalent to saying that one and the same subject has two different natures – a statement that implies a contradiction, at least when it is a question of a simple subject (as in the present case) rather than a composite one.

For Spinoza, by contrast, conceiving of one attribute without another (of Thought as really distinct from Extension, say) does not license Descartes’ inference to the existence of two substances whose attributes are Thought and Extension, respectively. Whether multiple attributes belong to the same substance or not, Spinoza contends that merely conceiving of single attributes could not possibly be sufficient for knowing which alternative is true. Given the independent reasons Spinoza presents for the causal and conceptual isolation of any two attributes, we are left in the dark about the implications of one subject’s having a thinking nature for the prospects of that same subject’s being extended.

Radical conceptual isolation of the attributes lets us infer almost nothing about whether or not they inhere in multiple substances. In other words, if one accepts attributes

30 CSM II 54, AT VII 78, AT VIIIa 28-29.
31 CSM I/298, AT VIIIB/349-350
as constituting the essence of substances, then two conclusions are consistent with two attributes being really distinct. That is, either the attributes’ conceptual independence necessitates their belonging to two distinct substances. Or, they belong to the same substance, but must be conceived without the other on independent grounds. One explanation is not obviously preferable at this stage.

So 1p10 arises from the conjunction of several lines of thought. First, Spinoza appeals to the restriction on substances as self-conceived, self-inhering, and self-caused. Second, he notes that attributes express the being of substance in a fundamental way. The first two claims make plausible similar restrictions for both substances and their attributes. And, since Descartes too argues for the need to conceive attributes independent of all others, it does seem that the Cartesian inference from attribute to substance plurality proceeds too hastily. Absent additional arguments then, Descartes and Spinoza seem to share justification for **Attribute Conceptual Independence**. But Spinoza, perhaps unlike Descartes, endorses at least co-extension among relations of causation, inherence, and conceptual dependence, as previously mentioned. So conceptual isolation of substances, given the co-extension of the various dependence relations, similarly grounds **Attribute Causal Independence**. If Spinoza can then motivate, as opposed to merely forestalling objections to, the possibility of a single substance having more than one fundamental way of being (**Multiple Attributes Per Substance**), his justification for the EB will be significantly stronger.

Given the importance of these three claims, we might explore Cartesian objections that would, if successful, undermine Spinoza’s understanding attributes in this way. Preliminarily, one might wonder whether Descartes should accuse Spinoza of
fundamentally misunderstanding his notion of real distinction. Recall the beginning of Spinoza’s 1p10: “From these propositions it is evident that although two attributes may be conceived to be really distinct (i.e. one may be conceived without the aid of the other), we still cannot infer from that that they constitute two different beings, or two different substances.” Whereas Spinoza here argues that attributes can be really distinct from each other, Descartes thinks that this is possible only for substances. So Spinoza’s claim, to Descartes, makes little sense on its face.

In Descartes’ Fourth Replies to the Meditations, he states: “[for] establishing a real distinction it is sufficient that two things can be understood as ‘complete’ and that each one can be understood apart from the other.”32 And only substances, for Descartes, can be complete: “[By] a ‘complete thing’ I simply mean a substance endowed with the forms or attributes which enable me to recognize that it is a substance.”33 With respect merely to terminology, then, Descartes is certainly capable of accusing Spinoza of misusing his notion of real distinction—these quotations show that quite clearly. The philosophical point remains substantially more difficult to evaluate, however. We have already seen that one, too simple way for Spinoza to establish that attributes are complete, in a way that is sufficient for being really distinct, is to identify them with substance. This is only questionably coherent, however, for reasons stressed earlier.34

---

32 AT 7:221, CSM 2:156

33 AT 7:222, CSM 2:156. In this paragraph, I exactly mirror the presentation of this relationship as given in Cunning, "Descartes' Modal Metaphysics."

34 See the earlier discussion of the quotations from Eisenberg, “On the Attributes and Their Alleged Independence of One Another.”
Still, Spinoza might be able to give an argument for each attribute’s being understood as ‘complete,’ in some requisite way. Then, since Descartes also accepts that Thought and Extension can be understood apart from each other, Spinoza will have at least shown that two attributes are candidates for being really distinct, pace Descartes’ insistence that they are not. As yet, however, we have not seen any hint of this direct argument for the completeness of attributes in Spinoza. Nor, though, has Descartes offered any principled reason for their wholesale exclusion from the realm of ‘complete things,’ beyond the terminological—“by a complete thing, I mean a substance.” So this preliminary objection strikes me as inconclusive.

However, Descartes does not merely introduce the terminology of ‘real distinction’—he puts it to work. And it might seem that Descartes’ argument for the real distinction of mind and body in the Sixth Meditation poses a more direct challenge to Spinoza. For suppose that Descartes successfully argues, on the basis of the real distinction between mind and body, for the existence of at least two substances. If one further supposes that the explanatory barrier plus some other uncontroversial premises entails Substance Monism, then denying Monism entails the denial of the EB as well. Descartes first states this argument for mind-body distinctness in the following way:

[O]n the one hand I have a clear and distinct idea of myself, in so far as I am simply a thinking, non-extended thing [that is, a mind], and on the other hand I have a distinct idea of body, in so far as this is simply an extended, non-thinking thing. And accordingly, it is certain that I am really distinct from my body, and can exist without it.”35

35 AT VII 78: CSM II 54
Later, Descartes appears to give a second argument for the real distinction between mind and body, couched this time in terms of whether each can be divided.\textsuperscript{36} I will not spend time considering either argument here—others have evaluated both at length.\textsuperscript{37} Of course, many have found Descartes’ arguments here fallacious, or at least puzzling. Perhaps a better response is simply that there are much shorter routes to substance pluralism—here is a table and there is a chair!—that do not engage directly with Spinoza’s understanding of attributes, but would serve to undercut it. So I am not sure that there is a distinctively Cartesian threat here.\textsuperscript{38}

Finally, one might wonder how Spinoza justifies the possible truth of the **Multiple Attributes Per Substance** thesis. While the consequences of this Spinozistic reasoning will be discussed more in the next section, Spinoza’s basic strategy is to argue that the more real a substance is, the more attributes must be attributed to it. This greater reality is then tied to greater power, so that God, as the most real substance, must have the most attributes and the most power, including the most power to exist. Spinoza offers virtually no defense of the first step, his assertion in 1p9 that “the more reality of being each thing has, the more attributes belong to it.” Two cryptic remarks supposedly justify the connection between greater being and greater attributes. The first occurs in Letter 9 to

\textsuperscript{36} AT VII 86-87; CSM II 59

\textsuperscript{37} For some small survey of the momentous secondary literature on Cartesian real distinctions and the mind-body problem, as well as modal questions associated with this style of argument, see: Frankfurt, *Demons, Dreamers, and Madmen*; Broughton, *Descartes’ Method of Doubt*; Bennett, “Descartes’ Theory of Modality”; Rozemond, *Descartes’s Dualism*; and Wilson, *Descartes*.

\textsuperscript{38} I do think, though, that exploring some of the presuppositions of Descartes’ real distinction argument, specifically as regards his model of divine creation and his tacit assumption of the substantial distinctness of God and creatures, serves to illuminate key theological differences between Spinoza and Descartes.
Simon DeVries, while the second is a single reference to 1d4, the definition of attribute, in the demonstration of 1p9 within the *Ethics*. But there are numerous problems here. For one, 1d4 is notoriously problematic in its own right.\(^39\) Further, the connection between the definition of an attribute and 1p9 remains opaque. In Letter 9, Spinoza offers two arguments. He says:

> I advanced two proofs, the first of which is as follows: It is clear beyond all doubt that every entity is conceived by us under some attribute, and the more reality or being an entity has, the more attributes are to be attributed to it. Hence an absolutely infinite entity must be defined…and so on. A second proof—and this proof I take to be decisive—states that the more attributes I attribute to any entity, the more existence I am bound to attribute to it; that is, the more I conceive it as truly existent. The exact contrary would be the case if I had imagined a chimera or something of that sort (Ep9).

Spinoza offers these considerations within a larger discussion of two types of definition. Within this conversation, they start to make a bit of sense, but only a bit. Earlier in the letter, Spinoza argues that the DeVries reading group has fallen into confusion because they have failed to distinguish between two types of definition. The first “serves to explicate a thing whose essence alone is in question and the subject of doubt, and [the second] is the definition which is put forward simply for examination” (Ep9). Now, if Spinoza takes himself to be explicating a thing *whose essence alone* is in question—God—then we can see him reasoning as follows. Spinoza tells us “that to the essence of any thing belongs that which, being given, the thing is necessarily posited and which, being taken away, the thing is necessarily [NS: also] taken away; or that without which the thing can neither be nor be conceived, and which can neither be nor be

\(^{39}\) See, for example, Newlands, “Thinking, Conceiving, and Idealism in Spinoza,” and Keizer, “Spinoza’s Definition of Attribute: An Interpretation.”
conceived without the thing” (2d2). God is defined as “a being absolutely infinite, that is, a substance consisting of an infinity of attributes, of which each one expresses an eternal and infinite essence” (1d6). So then God can neither be nor be conceived without infinitely many attributes, which constitute his essence. But now Spinoza’s contribution from Letter 9 becomes relevant. For his first “proof” there notes that since we conceive everything under some attribute, it seems that we can define an absolutely infinite entity, whose essence would consist in all possible attributes. Spinoza calls this substance God.

The second “proof” then works in tandem with this first claim. Spinoza asks us to recognize that an essence is that which, once given, necessarily posits the existence of the being whose essence it is, and which can neither be nor be conceived without that being. Consider some idea, of a horse, say. Spinoza might suppose his opponents will agree that the idea of the horse can be conceived qua thinking (i.e. as a psychological state) or extended mode (i.e. as a brain state), under the attributes of Thought or Extension.

Now notice the symmetry in Spinoza’s definition of essence. Not only can the entity neither be nor be conceived without its essence, but the essence can neither be nor be conceived without the entity whose essence it is. But Thought and Extension both exist—think of the idea of the horse! Spinoza would like us to infer that the most real being would have both attributes, of Thought and Extension, along with any others there might be. For if what belongs to a thing’s essence necessarily posits that thing’s existence, having any and all attributes contribute to your existence, a la Spinoza’s God, seems to give the maximal essence, the most real being.

In the next section, I briefly examine Descartes’ reaction to this line of reasoning, in the context of Spinoza’s argument for Substance Monism. Here, I anticipate those
conclusions by noting that having multiple attributes only contributes to substance’s maximal reality if such a possibility is antecedently assumed to be coherent. So while I will argue that Descartes’ objection is inconclusive, I by no means take myself to have provided any compelling Spinozistic defense of Multiple Attributes Per Substance either.

The preceding discussion strongly suggests that Descartes lacks a conclusive objection to 1p10, interpreted as requiring both conceptual and causal isolation between the attributes of Thought and Extension. However, I also hope to have emphasized the general limits of Spinoza’s argument for the EB. At minimum, given only the texts examined thus far, Spinoza must first motivate Attribute Conceptual Independence. Though this is common ground between him and Descartes, many contemporary philosophers have raised objections to even this point.40 He must then illuminate the connection between Attribute Conceptual Independence and Attribute Causal Independence. Recall that earlier, I merely provided some textual evidence for attributing the view to Spinoza, because of how he thinks about the interaction among various forms of ontological independence. I did little to motivate the view on independent grounds.41 Finally, some argument for the possibility of Multiple Attributes Per Substance seems to be required.42

40 See, for some examples of the claim, often cashed out in contemporary terms as the inconceivability of zombies (and thus a conceptual connection between the physical and mental of some kind): Della Rocca, “Spinoza and the Metaphysics of Skepticism,” Dennett, “The Unimagined Preposterousness of Zombies,” Kirk, “The Inconceivability of Zombies,” and Zombies and Consciousness, as well as many others.

41 See, however, Newlands, “Another Kind of Spinozistic Monism,” as well as Reconceiving Spinoza, for a more concerted defense of this connection.

42 In a more recent paper, Della Rocca, “Explaining Explanation,” has given a PSR-inspired defense of the claim that a substance requires more than 1 attribute, and that, if there are multiple attributes, they must all be attributes of the same one substance. While he has defended the latter claim in “Spinoza’s Substance Monism,” as well as in Spinoza, the former claim is new. Della Rocca imagines a scenario in
All this takes place against the backdrop with which we began: the sheer bizarreness of Spinoza’s views. Their oddity, at least, seems uncontroversial. For Spinoza asserts that the Thinking and Extended Substance, or that particular minds and bodies, are both one and the same thing, as well as completely causally isolated from each other! Spinoza might rightly emphasize the difficulty of pointing to any specific inconsistency in his system while respecting the EB. After all, how could one identify an inconsistency between Thought and Extension without appealing to cross-attribute relationships that the EB explicitly precludes? But this is hardly sufficient for demonstrating that no such inconsistency exists.

One might try to bring in Spinoza’s commitment to the Principle of Sufficient Reason (PSR) here, to justify the connection between Attribute Conceptual Independence and Attribute Causal Independence. After all, what could explain why modes of totally conceptually isolated attributes could nonetheless causally interact? Alternatively, why think relations (causation and conceptual containment) that are so similar in at least some circumstances could have radically different conditions of application? But this just

which there is only a substance S, and a single attribute, A. In this case, what explains why S is distinct from A? Merely saying that S is a substance and A an attribute is insufficient, given their overall similarities (i.e. both self-conceived, self-caused, etc.). So this is another Identification-friendly context, in which it seems that S=A. But then we have a substance that, according to Della Rocca, is “primitively conceived,” conceived not in virtue of any particular feature of it, but just brutally, primitively. This, Della Rocca thinks, would be unacceptable to Spinoza, and so, he thinks, substance cannot have only a single attribute. However, Della Rocca thinks, his argument “provides for a multiplicity of attributes, [but] it does not seem to provide for an infinity of attributes—something Spinoza apparently endorses” (33). But I do not think this can be right. For if Della Rocca’s argument works, then a situation in which there are any finite number of attributes—2, 5, 36, etc.—seems amenable to being ruled out by applications of the PSR to the question of whether substance, a being absolutely infinite, is limited to some finite (and arbitrarily-seeming) number of ways of being expressed. So if Della Rocca’s argument on Spinoza’s behalf gets the Multiple Attributes per Substance thesis up and running, I think it establishes the stronger claim of attribute infinitude that Spinoza endorses. This would be a serious result, since Della Rocca clearly employs both the PSR and the EB in his defense of Spinoza here, thereby making all the clearer the necessity of some defense of both principles.
saddles Spinoza with two problems, for he must now justify his commitment to the PSR and rule out the existence of some non-arbitrary competing explanation of the phenomena. The foregoing does suggest that Spinoza may have some ad hominem reasons for pushing Descartes to accept his profoundly anti-Cartesian conclusions. It also suggests that Cartesian objections surrounding these claims are inconclusive. But the explanatory barrier between the attributes still lacks a general, principled defense. This absence becomes even more frustrating in light of the EB’s central role in deriving nearly every major Spinozistic thesis. I now turn to showing how these arguments, from the EB to one of the five central claims with which I opened this chapter, are supposed to work.

1.4 Substance Monism and the Explanatory Barrier

First, I discuss what I take to be the strongest reconstruction of Spinoza’s argument in the Ethics for substance monism. **Substance Monism** claims: necessarily, there exists exactly one unique substance, and that substance is God, a being “absolutely infinite, consisting of an infinity of attributes, of which each one expresses an eternal and infinite essence” (1d6). In discussing Michael Della Rocca’s argument for Substance Monism, I aim to show how the EB works as an ineliminable part of Spinoza’s argument for his best-known claim.43

Before delving into the EB’s specific justificatory role, it is useful to have a skeletal version of the argument for Substance Monism in mind. The first crucial claim is 1p5, where Spinoza argues that no two substances share a single attribute. In 1p7,

43 This argument is developed in Della Rocca, “Spinoza’s Substance Monism,” as well as in Della Rocca, *Spinoza.*
Spinoza then claims that it pertains to the nature of a substance to exist. He then conjoins 1p7 with 1p11, the claim that God, a substance of infinitely many (i.e. at the very least, all possible) attributes, necessarily exists. But if God as substance possibly exists as having all the attributes there are, and if attribute sharing is ruled out by 1p5, one can infer 1p14, that besides God no other substance can be or be conceived. No other attributes are left for other possible substances—Spinoza’s God is an attribute glutton!

One can raise the following question about this basic line of thought.44 Suppose it pertains to the nature of substance to exist, from 1p7, and accept 1p5’s preclusion of attribute sharing. Then if any other substance is possible (such as purely thinking or purely extended substances, ts1 and es1, respectively), God’s existence is impossible. For 1d3 defines God as the being with all possible attributes, and 1p5 precludes attribute overlap among substances. So ts1’s mere possibility, in conjunction with 1p7, guarantees its existence, which in turn renders the nature of God impossible because internally contradictory.

In an earlier paper in response to this question, Don Garrett frames the question, following Spinoza, in terms of which possible substances have the most power to exist.45 Garrett argues that Spinoza first equates having more attributes with greater reality, and greater reality with greater power to exist. Spinoza then argues, in 1p11dem: “So, if what now necessarily exists are only finite beings, then finite beings are more powerful than an absolutely infinite Being. But this, as is known through itself, is absurd.” Della Rocca’s

44 In contemporary discussions, Garrett, “Spinoza’s Ontological Argument,” bears responsibility for highlighting this challenge. Della Rocca, “Spinoza’s Substance Monism” responds, on Spinoza’s behalf, to Garrett’s framing of the issues.

45 See, for example, Garrett, “Spinoza’s Ontological Argument,” esp. 218.
response to this framing is simply this: “[Spinoza’s] opponent takes seriously the possibility that ts1 [thinking substance 1] has more power to exist than does God. In doing this, the opponent is taking seriously the possibility that some difference between ts1 and God gives ts1 more power to exist than God.” Della Rocca helpfully fills out the imagined objection by noting that, for someone like Descartes, both ts1 and es1 have far more power to exist than Spinoza’s God. Descartes endorses this relative power claim simply because he finds the notion of a substance with multiple attributes incoherent, as we have already seen in section 1.2. The substances Ts1 and es1, by contrast, face no such obstacle.

This consideration alone suggests that both ts1 and es1 might exist, and indeed even comfortably coexist, since Substance Monism has not yet been established. Notably, this would be a slightly different line of thought than the Cartesian worries of the previous section. There I argued, with Spinoza, that Descartes could not simply infer immediately from a plurality of attributes to a plurality of substances. I also discussed Spinoza’s own arguments for **Multiple Attributes Per Substance**, and found that they seemed to require supplementation, to avoid merely assuming this possibility. Thus, both the arguments for Spinoza’s view, and the Cartesian objections to it, remained inconclusive there.

Here, however, given Spinoza’s other commitments, Descartes can object by relying on the even more plausible, minimal claim that rocks, say, are purely extended substances. Despite its remaining difficult, in the most general cases, to see how Descartes can positively conceive of any given object as either wholly extended or

46 Della Rocca, “Spinoza’s Substance Monism,” 25
wholly thinking, the case for rocks in particular being wholly extended seems intuitively strong. If rocks lack mental activity, this casts doubt, not on 1p10 directly, but on the coherence of Spinoza’s definition of God in 1d6. Or, to press the challenge elsewhere, it endangers his argument for God’s necessary existence in 1p11. For Descartes can argue that rocks’ purely extended existence, combined with Spinoza’s commitment to the no-attribute sharing thesis in 1p5, serves as the reason why God’s nature is internally contradictory. So defined, if rocks are purely extended substances, Spinoza’s God, as having infinitely many attributes, does not exist. Either the definition of God, or the claim of his existence, then has to go.

Here is another way to see the challenge. The Cartesian objector might comfortably grant that Extension itself constitutes an attribute of Substance, as in a wholly extended rock. But then, given 1p5’s prohibition on attribute sharing, the option to attribute even more reality to the wholly extended rock, by making it Extended and Thinking, is a vacuous one. Such an attribution represents a genuine impossibility. So, in conclusion, it looks as if Spinoza does very little to motivate the idea that any substance could have multiple essences, or more carefully, multiple basic ways of being “that the intellect perceives of a substance, as constituting its essence.”

Perhaps Spinoza does more than nothing. For instance, Newlands, Reconceiving Spinoza, Ch. 4 has recently argued that Spinoza has a conceptualist account of essences, according to which “a thing’s essence is the containing of states of the world in the concept of a thing’s categorical basis.” A thing’s essence is an explanatory power between a thing’s categorical basis, or those properties that give rise to and account for its power (Ch. 4). But, as Newlands himself makes clear, this conceptualist account of essences is shot through with reliance on the EB. For instance, he reasons: “Now consider a particular attribute, such as Thought. Does Thought express, contain, or explain all of God’s effects? It would seem not, according to 1p10 and the explanatory barrier between attributes. The attribute of thought explains only God’s thinking effects. Likewise, the attribute of extension explains only God’s extended effects. Given that essences are explanatory connections, differences among the explananda entail the non-identity of the corresponding essences. Hence, the thinking substance has a different essence than the extended substance. And insofar as God is both a thinking and extended substance, it follows that God has multiple essences” (Ch. 4). This makes it an interesting (indeed, more compelling, variant) on the Della Rocca strategy to
played by these correlations between power and reality, and between reality and multiple attributes, it looks like Spinoza merely begs the question against the Cartesian position. Certainly, if Spinoza’s argument really must rely on God’s greater power to exist, relative to other possible substances, additional arguments would seem to be required.

Before moving on to Della Rocca’s preferred defense of Spinoza, which does not so directly rely on considerations concerning God’s greater power to exist, I do want to make one final point here, this time in Spinoza’s favor. I do not think that Spinoza’s argument works to rule out the option of there being one substance for each attribute (Thought, Extension, Attribute Z, etc.). But I think it is far more effective against someone who tries to combine this one substance per attribute picture with a commitment to God’s existence. Traditional theology conceives of God as maximally real and maximally existent. But as Spinoza’s objector would have it, each attribute constitutes the essence of a substance, but it is not the case that they all constitute the essence of the same substance, on pains of incoherence. Still, how can a picture with a Thinking Substance, an Extended Substance, an Attribute Z substance, etc., plus the claim that no two substances can share an attribute, make room for God? In what does God’s essence consist, if not in any of the attributes that seem to constitute the essences of everything that we can conceive?

So if a commitment to God’s existence is presupposed as common ground between Spinoza and his objector, then I think his argument regains some of its force. For then the traditional theologian’s commitment to God’s maximal reality tells against an a
priori commitment to the view that substances cannot have multiple attributes. Of course, this commitment to God’s existence need not be shared by someone simply interested in Spinoza’s proof of Substance Monism on independent grounds. But if one maintains the view that substances cannot have multiple attributes, and confronts the no-attribute sharing thesis, then one seems pushed towards the view according to which there exists exactly one Thinking Substance, one Extended Substance, etc. Yet this seems incompatible both with one’s belief in God’s existence, as traditionally construed, as well as with the strong pre-theoretical tendency to consider oneself, as well as other human beings, as substances in their own right.

Fortunately, however, Della Rocca provides a far more compelling reconstruction of Spinoza’s argument for Substance Monism, one firmly grounded in the EB. He begins by assuming that no two substances share a single attribute (1p5), which he takes himself to have argued for, on good Spinozistic grounds.48 Della Rocca then argues that every possible attribute is instantiated by some substance, though of course at this stage of the argument these substances are not presumed to be identical. His reason is simple. Attributes, like substances, are self-conceived (1p10d), so since self-conception considerations in 1p7 generate the claim that every possible substance exists, the corresponding claim for every possible attribute, on similar grounds, appears to follow too. And, of course, each attribute must be an attribute of some thing—free-floating attributes are only marginally coherent. Della Rocca concludes: “We can draw the

48 Della Rocca, “Spinoza’s Substance Monism,” 17-22. Della Rocca argues basically that any attribute X shared between two substances would be insufficiently conceptually independent, in violation of 1p10, and would thus be conceived through some other attribute, Y or Z, endangering X’s status as an attribute.
conclusion that if x is an attribute then x is possessed by one substance and by no other. Thus we can speak of the x-substance and, similarly, of the thinking substance and the extended substance.49

With every possible attribute instantiated by a unique substance, the stage is set to begin asking whether any of these unique substances are identical to any of the others. Della Rocca argues, relying on a version of the principle of the identity of indiscernibles, that in fact each unique substance—say, the extended substance or the thinking substance—is identical to each of the others. That is, there is one unique substance, God, that possesses all the attributes, and Substance Monism is true. The trick is to begin asking, for any purportedly non-identical pair of unique substances X and Y, what difference explains their nonidentity. This trick gains its power from the fact that each substance seems to share many attribute-neutral properties with every other. Della Rocca defines these attribute-neutral properties as “those properties such that saying that something has one of these properties does not entail anything about which particular attribute or attributes it has or does not have.”50 For example, each substance is self-conceived, in itself, infinite, etc.

This creates what might be called an Identification-friendly Context, which arises regularly in the early parts of the Ethics. In Identification-friendly contexts, a reason or set of reasons creates pressure to identify purportedly distinct entities, be they substances

49 Della Rocca, “Spinoza’s Substance Monism,” 27

50 Della Rocca, “Spinoza’s Substance Monism,” 27; Newlands, “Thinking, Conceiving, Idealism,” 46 appears to treat attribute-neutral properties as ways of talking about how things are structured in the most general way. For one attribute-neutral relation, that of being ‘conceived as,’ Newlands interprets this as roughly synonymous “with ‘insofar as it is structured thusly,’ where ‘thusly’ names an attribute context.”
or modes, and no sufficiently compelling reason for maintaining nonidentity emerges.\textsuperscript{51}

In fact, often the prima facie best reasons for maintaining such a nonidentity are ruled out precisely because of the explanatory barrier between the attributes. So the rational pressure to identify, based ultimately on Spinoza’s commitment to the PSR and EB, comes to dominate.

In this Identification-friendly context, a survey of purported reasons for non-identification yields no promising candidates. For example, one cannot argue that it is in virtue of the fact that ts\textsubscript{1}, say, has Thought that it lacks Extension. This line of reasoning violates the EB, since Della Rocca thinks that the EB entails (3): Each attribute of a substance, independently of any other attribute of that substance, is sufficient for conceiving of that substance.\textsuperscript{52} To see why, recall that Spinoza’s PSR allows us to ask for the cause or reason of each thing’s existence and nonexistence (1p11dem). So consider whether it is possible for Spinoza’s substance with both Thought and Extension to exist. Whatever reason one gives for its nonexistence must respect the radical independence of the attributes. For a substance, considered as extended, must have its actual properties, as well as those it lacks, explained in either extended or attribute-neutral terms. Appealing to a substance’s possession of Thought, therefore, to explain why it lacks Extension, impermissibly violates the EB.

The final attempt Della Rocca considers for blocking identification consists in arguing: “It is part of the thing’s nature or essence to possess or lack that property.

\textsuperscript{51} Newlands, “Another Kind of Spinozistic Monism,” 488, defines Spinoza’s inference as follows: Necessarily, for all x and y, the existence of some explanatory grounds for the identity of x and y, and the lack of any non-primitive explanatory grounds for the non-identity of x and y entail the identity of x and y.

\textsuperscript{52} Della Rocca, “Spinoza’s Substance Monism,” 21.
Perhaps, then, it is part of the nature or essence of the thinking substance that it lacks extension."\(^5^3\) This position, however, seems to be precluded by the open question that the EB poses to its defenders. Meghan Sullivan suggests that facts about essences or natures might be ‘question closers.’ That is, “once an essence fact has been identified, it is meant to strike us as inane to ask for further explanation of that fact.”\(^5^4\) But, given that no facts about a particular attribute can explain those of another, Spinoza’s open question is: what other possible reason could there be for assuming that a substance essentially lacks a given attribute? Absent some answer, and given some shared attribute-neutral properties, Spinoza therefore demands the identity of the thinking and extended substance. Finally, since the two attributes chosen were in no way special, one can run this pairwise argument for every set of two attributes. This culminates in the assertion that one unique substance has all the attributes, given the no-attribute sharing and every-attribute-instantiated theses we have thus far been assuming.

Because Della Rocca frequently employs this style of open question objection, and because I think there is something illicit about it, at least as a completely general strategy, three important caveats are in order here concerning Spinoza’s argument. First, one should recognize the inherent limitations of such a strategy, or really any argument that relies substantially on the PSR. The limitation is simply this: there will always be the problem of motivating the claim that one has surveyed all possible explanations and found them insufficient. Additionally, brute facts have a dandelion-like propensity for popping up in strange places in even the most hardened rationalist’s philosophical

\(^{5^3}\) Della Rocca, “Spinoza’s Substance Monism,” 29

\(^{5^4}\) Sullivan, “The Irrelevance of Essence,” 505
system. So it is always open to the objector to posit just the putative brute fact at issue as fundamental. Indeed, it seems plausible that a Cartesian objector might wish to do so here. Unless the rationalist can ensure that her system is principled all the way down, the prospects for which are tenuous at best, the objector can rightfully ask why a brute fact is precluded from entering at precisely this moment.

Second, one might think that essences are simply not the sort of thing of which one can coherently have more than one. Then, if attributes are defined as “what the intellect perceives of a substance, as constituting its essence,” some fact about essences, rather than any illicit appeal across attributes, would guarantee that each substance has only one. Then, given the uniqueness of essences, Spinoza’s proof in 2p1 that Thought is an attribute of God would show that God is solely thinking. Really appreciating something’s essence as Thinking would show that thinking swallows up, as it were, everything it is to be that thing. No appeal to anything about Extension would then be required to show that the unique thinking and unique extended substance are distinct. At best, Spinoza would have delimited a single, extended substance, as well as a single, different, thinking substance. But the status of putatively distinct thinking and extended substances, such as my mind and my body, would be even less clear, in this one-substance-per-attribute system, than in Spinoza’s preferred monistic one.

My third caveat points to an important feature of the dialectical situation surrounding Garrett and Della Rocca’s respective answers to the question of which substances are possible. Della Rocca takes issue with Garrett’s framing of the question in terms of God’s greater power to exist, relative to purely thinking or purely extended substances. For he argues that Spinoza has simply begged the question, on this reading,
about which substances are more powerful. In order to avoid saddling Spinoza with an argument he finds question begging, Della Rocca attempts to reorient the argument for Substance Monism around the EB, to avoid the issues plaguing Garrett’s proposal.

This reorientation succeeds only if the EB receives independent support. For it becomes clear that, as Della Rocca himself admits, “we can see Spinoza as holding that it *is* possible for a substance to have more than one attribute. This is possible precisely because of the conceptual separation among the attributes.”55 The heart of the matter is that identification can only be favored if the explanatory barrier has some independent justification. This would make its application here, to the question of justifying the possibility of substances’ having multiple attributes, appropriately uncontroversial. Otherwise, Della Rocca has merely relocated the source of the justificatory lack. Reorienting Spinoza’s arguments around the explanatory barrier remains a powerful strategy only if we can make some progress in justifying it.

1.5 Parallelism and Mind-Body Identity

We have seen how Spinoza’s argument for Substance Monism depends in perhaps unexpected ways on the explanatory barrier between the attributes. In this section, I want to explore one way of arguing for Parallelism and Mind-Body Identity using the explanatory barrier. Others have discussed arguments connecting the EB with these two claims, but especially with Mind-Body Identity, in detail. Nonetheless, a clear description

55 Della Rocca, “Spinoza’s Substance Monism,” 29
of one explication of this dependence helps to emphasize the radical extent to which the
EB permeates Spinoza’s metaphysical system.  

Take Parallelism first, the canonical statement of which occurs in 2p7, where
Spinoza writes: the order and connection of ideas is the same as the order and connection
of things. This seems to require that for each thing there is one corresponding idea, and
vice versa, which also implies, since ideas are things, that for each idea there is one
corresponding idea of that idea. Moreover, the relations—order and connection—among
ideas exactly mirror the relevant relations among things. Della Rocca draws additional
attention to the representational nature of the parallelism, noting: “An important feature
of parallelism is the fact that an idea represents the item with which it is parallel.” Della
Rocca also offers a helpful, albeit dense, summary of the way that Parallelism depends on
the EB. He says:

First, the explanatory barrier between thought and extension is (at least partly)
responsible for the essence requirement on representation. This requirement,
together with Spinoza’s causal notion of essences, generates the causal
requirement on representation. The causal requirement in turn, together with the
claim that God has unconfused ideas of each thing, generates Spinoza’s
representational parallelism, according to which the order and connection of ideas
in God’s mind matches the order and connection of the objects of those ideas.

---

56 For the connection between the EB and Parallelism, the classical examples are Gueroult,
Spinoza 2 Vols, and Della Rocca, Representation and the Mind-Body Problem. Recent examples include
Melamed, “Spinoza’s Metaphysics of Thought.” Philosophers discussing mind-body identity, often in
connection with the EB, include: Jarrett, “Spinoza’s Denial of Mind-Body Interaction and the Explanation
of Human Action”; Della Rocca, “Causation and Spinoza’s Claim of Identity,” and Representation and the
Mind-Body Problem”; Odegard, “The Body Identical with the Human Mind: A Problem in Spinoza’s
Philosophy”; Koistinen, “Causality, Intensionality, and Identity: Mind Body Interaction in Spinoza”;

57 Melamed, “Spinoza’s Metaphysics of Thought,” has recently offered serious arguments against
the view that Spinoza maintains a single doctrine of Parallelism. I discuss his views in section 1.7.

58 Della Rocca, Representation and the Mind-Body Problem, 18

59 Della Rocca, Representation and the Mind-Body Problem, 101
Let me briefly explain the key steps of the dependence of Parallelism on the EB outlined here. What Della Rocca calls the essence requirement on representation is the view: “The content of an unconfused idea of x must be ‘the thing with E’”, where E is the individual thing’s essence. The causal requirement on representation, relatedly, is the view: “In order for a mind to have an unconfused idea of x, that mind must have an unconfused idea of x’s finite cause, c.” The link between these two requirements, broadly speaking, arises from Spinoza’s views on the relationships among definitions, causes, and essences. For example, Spinoza says in TIE 95: “To be called perfect, a definition will have to explain the inmost essence of a thing.” Since he also explains, in Letter 60, “[that] the idea or definition of a thing should express its efficient cause,” the relevant connection seems to be the following. In order to give the definition of a thing, one should identify its essence, and the best way to do so is to locate its (finite) efficient cause.

To see how the EB grounds Parallelism, I follow Della Rocca in speaking of essential and non-essential representation, where essential representation involves representing an item x by grasping either its essence E or some property that x has only because it also has E. Non-essential representation, by contrast, involves picking out x by a set of properties that, while they do serve to uniquely identify x, do not do so by grasping x’s essence E or any properties that x has solely in virtue of having E. Now

60 Della Rocca, Representation and the Mind-Body Problem, 84
61 Della Rocca, Representation and the Mind-Body Problem, 70
62 The discussion of why the thing’s finite, rather than infinite, efficient cause should be identified is not particularly relevant for my purposes here. Della Rocca does discuss, however, some reasons for thinking that the finite reading is preferable.
consider the idea R that represents any arbitrary extended mode x, for instance my baseball hat. Suppose R represents x non-essentially, so that R singles out my baseball hat by properties other than its essence E. One natural thought is that the representation R’s ability to uniquely identify my baseball hat might depend on a cluster of non-essential physical features surrounding it. Perhaps, for example, my hat is the sole ball cap located in a house whose deed is in in my name at the county assessor’s office.

Here is the problem. It now looks like the properties of some idea R depend on some physical facts—the ball cap’s location in a certain house, and the relevant deed’s physical location and causal history (i.e. actually having been signed by me, etc.) But this dependence of the mental item R’s properties on some physical states of affairs violates the EB. For, as is evident from 2p5, Spinoza says that a thinking mode’s properties “express, in a certain way, God’s nature insofar as he is a thinking thing. And so (by 1p10) it involves the concept of no other attribute of God, and consequently (by 1ax4) is the effect of no other attribute than thought.”

Parallelism’s dependence on the EB becomes clearer if one combines the thought that God has unconfused ideas of each thing with the claim that the EB makes cases of non-essential representation, as interpreted above, inadequate. For if God has unconfused ideas of each thing, God’s ideas must satisfy both the essence and causal requirements, such that in having an idea R of my baseball hat x, God has an idea R of x as ‘the thing with E’, where E is given as the complete expression of x’s efficient cause. If the EB precludes God’s non-essential representation as inadequate and if non-essential

63 Here I follow and elaborate on an example Della Rocca, Representation and the Mind-Body Problem, gives on pp. 98-99.
representation is interpreted as dependent on non-essential physical facts, then we are pushed in one of two directions.

First, all non-essential representation might be inadequate. This thought entails parallelism directly, for if the restrictions on essential representation are correct, then for God to represent x, he must represent it as the thing with E, where E gives the complete efficient cause. If God has ideas of each thing, then this secures Parallelism. Or, second, non-essential representation must be interpreted as being of representations of those non-essential physical facts, rather than as dependent on the non-essential physical facts themselves. But this second thought—that every non-essential physical fact has a corresponding representation, which is required to ground non-essential representations of idea R—just entails parallelism between Thought and Extension. Idea R(x) has a corresponding Idea(Idea(R(x))), and this idea of idea R must be explained in terms of representations of the physical objects in x’s causal history. In either case, the order and connection of ideas is the same as the order and connection of things, each of which is represented by its corresponding parallel idea.

Once Spinoza’s system has Parallelism up and running, it can supplement the EB to generate Mind-Body Identity. The basic line of thought is this. Modes have many attribute-neutral properties (Della Rocca sometimes uses ‘extensional’ synonymously). These are properties that a mode has that do not presuppose or entail anything about the attribute to which it belongs. Earlier, attribute-neutral properties were also described as structural, not dependent on a particular attribute description. They also have intensional properties, like being a thought about the Lakers or being a baseball cap mode, that do presuppose a mode’s being of a particular attribute. That is, if a mode R is a thought
about the Lakers, this directly entails that R is thinking. Similarly, being a baseball cap mode x directly entails that x is extended. Della Rocca then gives the following argument, which relies on a principle he calls Leibniz’s Law: A=B iff A and B have all their extensional [i.e. neutral] properties in common.64 He says:

Parallelism helps us to see that mind and body share all their neutral properties. Since all extensional properties are neutral, and since the mind and the body must have some extensional properties, it follows that mind and body share all their extensional properties. By Leibniz’s Law we can, therefore, conclude that mind=body.65

Della Rocca gives several examples of neutral properties, and cites Spinoza frequently using 2p7, where the Parallelism thesis is stated and argued for, to secure the mind and body’s sharing of several sorts of neutral property. His examples include having a certain number of immediate causes and effects, having certain temporal properties, and being a certain sort of complex individual.66 Della Rocca argues that mind and body share all relevant sorts of neutral properties by noting that neutral properties, for Spinoza, are attribute-independent, by definition. Thus, they focus only on general or structural features that modes have. But recall that for Spinoza, x’s essence is closely tied to x’s finite cause. So these structural features of modes seem to contribute to the explanation of that mode’s place in larger causal and explanatory networks, the isomorphism of which Parallelism guarantees across attributes. So Spinoza can be seen as thinking that every sort of neutral property will be related to the order and connection of modes in some way, or will be otherwise irrelevant to questions of identity. Thus,

64 Della Rocca, *Representation and the Mind-Body Problem*, 131
65 Della Rocca, *Representation and the Mind-Body Problem*, 137
66 Della Rocca, *Representation and the Mind-Body Problem*, 133-140
parallelism guarantees the sharing of all neutral properties between parallel modes, each of whose existences are guaranteed by the Parallelism thesis itself.67

Each step of this compressed argument, of course, requires careful defense, which Della Rocca spends a great deal of time trying to provide. And obviously, there are other ways to state Spinoza’s route to mind-body identity. Notably, however, most of the descriptions of Spinoza’s line of argument do rely on drawing some distinction similar to that which Della Rocca draws between intensional and extensional properties. Koinstinen, for example, focuses on Spinoza’s notion of causation in intensional and extensional terms, but others hone in on explanation, or some other notion.68 Here, I am not so concerned to defend Spinoza’s argument. Consequently, I am not at this point interested in exploring ways that one or another step might be resisted. Rather, the point is simply this: the EB is crucial for generating Parallelism, and this thesis plays the starring role in many reconstructions of Spinoza’s argument for Mind-Body Identity.

The centrality of the EB’s importance in this reconstructions mirrors Spinoza’s own strategy for motivating the identity claim in 2p7s. There, he appears to invoke 2p5 and 2p6, which mention 1p10’s explanatory barrier explicitly. So the EB should be seen as lying at the heart of several of Spinoza’s most important theses, including Substance Monism, Parallelism, and Mind-Body Identity. Or so the history of scholarship on Spinoza, buttressed by substantial textual references from the Ethics and other works,

67 Della Rocca carefully argues for the plausibility of this general line of thought, and also notes some properties that might not contribute to order and connection but are also universally instantiated and thus, one might think, thereby irrelevant to questions of identity, since noting that an object has them does not serve to single it out in any way (i.e. being one of many). Again, these subtleties are mostly irrelevant to my purposes at this point.

would suggest. Della Rocca has played a special role in making this pervasive
dependence on the EB more evident. So his insistence that the EB is both inadequately
defended and absolutely essential to Spinoza’s metaphysical system is worrisome.\textsuperscript{69} At least, anyone interested in the viability of Spinoza’s project should be searching for some adequate defense of EB compatible with his other commitments. Before attempting to do just this, however, I want to hammer home the importance of the EB by describing the role it plays in illuminating several lesser-known but fascinating puzzles in the \textit{Ethics}.

\textbf{1.6 Blocking Idealism with the Explanatory Barrier}

Jonathan Bennett’s book, \textit{A Study of Spinoza’s Ethics}, modeled a comprehensive reading of Spinoza’s greatest work. It also generated or elaborated in greater detail upon several riddles or lacunae that even his earliest readers have felt threaten the coherence of Spinoza’s overall metaphysical system. The following two sections outline some recent responses to two of Bennett’s outstanding questions, highlighting in the process the surprising utility of the EB in combatting diverse problems.

Samuel Newlands has recently argued that the EB plays the important role of preemptively defending Spinoza from charges of idealism. Newlands outline of this defense requires further reflection on the connection between concepts, conceiving, and conceptual relations, on one hand, and ideas, thinking, and mental relations, on the other. Newlands responds to Bennett’s first charge, of “a lopsidedness in Spinoza’s system which he does not mention, could not explain, and should not have tolerated” by first

\textsuperscript{69} The relevant expressions of pessimism about the explanatory barrier’s defense occur in Della Rocca, \textit{Representation and the Mind-Body Problem} and “Spinoza’s Metaphysics of Skepticism.”
expanding the scope of the charge. In the process, he catalogues many of the various ways that Spinoza seems to explicate the most crucial tenets of his system in conceptual terms, which are read naturally as mental ones, associated with the attribute of Thought.\textsuperscript{70}

The central worry is that the attribute of Thought somehow plays a more fundamental role in Spinoza’s metaphysics than that played by the other attributes. Thought’s pivotal, undergirding role appears to endanger at least four theses that nearly all commentators agree Spinoza defends. Newlands labels these views attribute plenitude, attribute independence, attribute parallelism, and attribute parity.\textsuperscript{71} Briefly, these views amount to the claims that (1) Spinoza affirms the existence of more than one attribute and (2) insists on the conceptual separation and self-containment of each attribute and its modes from every other (the EB, in my terms). He also (3) believes in an isomorphism between (at least) the modes of Thought and those of every other attribute (what I have called Parallelism). Finally, Spinoza (4) rejects the notion that any attribute is more fundamental, more attribute-like, than any other. Giving up this last thesis of attribute parity would amount to a sort of Spinozistic idealism. For the fundamentality of Thought in God seems incompatible with the affirmation of Extension as a genuine essence of substance, a genuine attribute, independent of and not ultimately reducible to Thought.

As Newlands reconstructs Bennett’s challenge, it has two key steps. First, Spinoza’s privileging of conceptual relations entails a privileging of the mental, and second, Spinoza’s privileging of the mental leads to inconsistency with the four theses

\textsuperscript{70} Bennett, \textit{A Study of Spinoza’s Ethics}, 62

\textsuperscript{71} Newlands, “Thinking, Conceiving, Idealism,” 34
described above.72 Denying the first step, which Newlands calls mentalism—every conceptual relation is a mental relation—is his central way of meeting Bennett’s challenge. Rather than recapitulating this argument in its entirety, I just focus on the role that the EB plays in securing Newlands’ desired conclusion.

First, the EB helps generate an argument for the thesis of attribute parity. As Newlands writes:

In his opening definitions, Spinoza claims that instances of causation and being-in are co-extensive with instances of conceptual dependence. Although Spinoza uses numerous synonyms for forms of dependence, each is ultimately equivalent to causation, being in, or being conceived through. Hence, if one attribute, A1, depends in any way on another, A2, then A1 is conceived through A2. However, the conceptual dependence of one attribute on another is explicitly ruled out by the attribute independence doctrine. So the conceptual isolation of attributes plus the co-extension of conceptual dependence with any other form of dependence entails attribute parity.

Now, Newlands defends a view much stronger that the mere co-extension of causation, being-in, and being conceived through, which he calls Conceptual Dependence Monism. According to CDM, there is only a single, conceptual, dependence relation in Spinoza, which is identical with causation, inherence, and conceptual containment.73 But for our purposes, in the context of worries about the connection between the denial of attribute parity and idealism, this stronger thesis need not be maintained. Indeed, even the Co-Extensive Thesis is unnecessary to obtain the desired conclusion.

Rather, we can generate an argument for attribute parity with much humbler premises. First, Spinoza attempts in 1d4 to offer a definition of attribute, as “what the

72 Newlands, “Thinking, Conceiving, Idealism,” 34

73 See, for example, Newlands, “Another Kind of Spinozistic Monism,” for a detailed defense of CDM.
intellect perceives of a substance, as constituting its essence.” This definition, as we
noted earlier with reference to Letter 60, must express the essence of what an attribute is
by expressing its cause. But, as Newlands himself notes, one might naturally read
Spinoza as asserting, from 1d4 plus the causal requirement from Letter 60, (A) the
intellect to be the cause of the various attributes, including Extension, and its modes. On
the other hand, it seems clear that Spinoza should be read in 2p6 as inferring the opposite.
Namely, that (B) because “the modes of each attribute involve the concept of their own
attribute, but not of another one; [they, by Ia4] have God for their cause only insofar as
he is considered under the attribute of which they are modes.” More straightforwardly,
Spinoza here simply denies that the intellect is the cause of Extension or its modes.

But then it seems that we do not even require the general co-extension of
causation, being-in, and being conceived through to get an objection to Spinoza’s
position going.74 The following way of stating the objection makes precise the conflict
between the two statements, (A) and (B), given above. Rather than general co-extension
of Spinozistic dependence relations, it suffices merely to note that Spinoza holds that if x
is conceived through an attribute A, then x has God, considered under A, as its cause.
This is equivalent to saying that if it is not the case that x has God, considered as A, as its
cause, then it is not the case that x is conceived through A.

74 Newlands discusses the definition of attribute in 1d4 at both the very beginning, and the end, of
“Thinking, Conceiving, Idealism.” However, he does not explicitly make the point about the independence
of his argument even from the more moderate claim of co-extension among causation, conception, and
inherence, and indeed at points seems to imply the opposite—namely, that co-extension is required. And
this is a good thing, for while some (i.e. Garrett, “Spinoza’s Conatus Argument,”) have defended the co-
extension of the three dependence notions, others (i.e. Lærke, “Spinoza and the Cosmological Argument
According to Letter 12,”) have explicitly denied it.
If this claim is combined with the causal requirement on definitions from Letter 60, Spinoza’s explicit avowal of Extension as an attribute of God in 2p2, and the definition of attribute in 1d4, then we have the following argument.

1. Extension is defined as what the intellect perceives of God, as constituting its essence (1d4, 2p2).
2. The causal requirement on definitions states that a definition must express the cause of what is defined.
3. The intellect, in its perception of God, is the cause of God considered as Extended (1,2).
4. God is the cause of the essence of any extended mode x (1p25).
5. The objects of ideas (i.e. extended things) follow solely from God considered as Extended (2p6cor).
6. If x causes y, and y causes z, then x, at least to some extent, causes z.
7. God considered as Extended is the cause of any extended mode x (4,5).
8. The intellect, in its perception of God, is the cause of extended mode x, at least to some extent (3,6,7).
9. The “modes of each attribute have God as their cause only insofar as he is considered under the attribute of which they are modes” (2p6cor, invoking EB).
10. Contradiction of 8 and 9.

The contradiction arises here, on my view, only upon the presumption that x’s being conceived through God considered as A implies that x is caused by God considered as A (to get 1-3), which is a weaker presumption than either co-extension of causation,
conception, and inherence, or of Newlands’ preferred Conceptual Dependence Monism. The logical weakness of this assumption, of course, I take to be a strength of the reading I am offering here, which is otherwise quite close to that offered by Newlands.

The upshot here is that (A) must be rejected. 1d4 should not be read as suggesting that the intellect is, in any sense, the cause of the other attributes. Rather, 1d4 is a failed definition by its own lights, and dropping its apparent causal implications restores attribute parity.\(^75\) So, given the internal problems with 1d4 independent of the EB, as well as the EB’s centrality to the rest of Spinoza’s system, we can see how EB should be seen as successfully motivating attribute parity. For denying attribute parity necessitates the rejection of EB, which compromises a whole host of Spinoza’s other commitments.

Newlands also uses the EB to show why mentalism must be rejected in order to save the proof, in 2p2, that Extension is an attribute of God. The reason is straightforward. Newlands first cites Spinoza’s argument in 2p2, which tells us only “[that] the demonstration of this [i.e. Extension is an attribute of God] proceeds in the same way as that of the proceeding Proposition.” Newlands reconstructs 2p2 to parallel 2p1 as follows: “Therefore (by 1d5) there belongs to God an attribute whose concept all singular bodies involve, and through which they are also conceived.”\(^76\) But, as Newlands rightly notes:

If the fact that a substance has the attribute of Thought cannot be used to explain why a substance lacks the attribute of extension, surely the fact that a substance has the attribute of thought cannot be used to explain why a substance has the attribute of extension. Yet if conceptual relations just are relations of Thought [i.e.

\(^75\) Newlands, “Thinking, Conceiving, and Idealism,” makes the point about 1d4 as a failed definition at the end of this paper.

\(^76\) Newlands, “Thinking, Conceiving, and Idealism,” 39
on the assumption of mentalism], then by appealing to conceptual relations between bodies to prove that God is extended, Spinoza will have thereby appealed to Thought to prove that God is extended.

Newlands point here can be made more explicit by recalling my earlier discussion of Descartes in section 1.3 and of Della Rocca’s reconstruction of the argument for Substance Monism in 1.4. There, having the attribute of Thought was seen as insufficient for inferring that a substance lacked the attribute of extension. Such a guarantee would, in using the presence of one attribute to explain the absence of another, run afoul of the EB. Here, Newlands mirrors this point by arguing that conceptual relations cannot be mental relations. For the existence of such conceptual/mental relations between bodies, as in the reconstructed 2p2 above, could not explain why Extension is an attribute of God, again on pain of violation of the EB.

Newlands’s reflections on the relevance of the EB to the denial of mentalism, the demonstration of attribute parity, and the proof that Extension is an attribute of God serve to further emphasize the central thrust of this opening chapter. That is, it becomes still clearer that any hope of preserving the overall, non-Idealistic and non-mentalistic coherence of Spinoza’s system also requires a principled defense of the EB. So, if we worry, with Bennett, that one naturally understands a privileging of the conceptual as a privileging of the mental, and a privileging of the mental entails idealism, we get the following. Absent countervailing arguments against this natural tendency, Spinoza’s

77 This is so even if one reads Newlands as arguing merely that the explanatory barrier is merely sufficient for alleviating worries about idealism. To simplify the logical relationship between various theses a bit: if (EB→¬Idealism), then, by contraposition, (Idealism→¬EB). But, as the foregoing was meant to suggest, (¬EB→¬(Substance Monism, Parallelism, Mind-Body Identity). At least, Spinoza fails to suggest other ways these theses might be justified within his system.
privileging of the conceptual is a privileging of the mental. This privileging of the mental is a brand of Idealism. Idealism entails the negation of the explanatory barrier between the attributes. Certainly, Spinoza gives no indication of thinking that these theses can be justified independently of his commitment to the explanatory barrier. And without the EB, Spinoza cannot derive the theses of Substance Monism, Parallelism, and Mind-Body Identity to which he is obviously committed. This generates the alleged incoherence in Spinoza’s system. Of course, one might try to provide the requested countervailing arguments without at all relying on the EB (or on any thesis that requires the EB for its derivation). Alternatively, one might suggest that this privileging of the mental is not yet full-blown idealism. But given how Spinoza actually argues, as well as the preceding sections, I take all this as good evidence for the central theoretical importance of justifying the EB.

1.7 Knowledge of Other Attributes

I conclude this chapter by considering Yitzak Melamed’s promising response to a puzzle about the lack of human knowledge of attributes other than Thought and Extension. Spinoza’s contemporary Tschirnhaus first expressed puzzlement at this restriction, and Jonathan Bennett has forcefully renewed the challenge. The worry is this. Spinoza argues that every extended mode $em_1$ is identical to a thinking mode $tm_1$ and a mode of attribute $z$, $zm_1$. So why can human minds know nothing about the infinitely many other attributes, besides Thought and Extension, whose existence Spinoza affirms in 1p14? There, God is shown to be a substance “consisting of infinitely many attributes, each one of which expresses an eternal an infinite essence (1d6). From each of these
infinitely many attributes follow infinitely many modes, and we have already considered
Spinoza’s claim that these modal orders parallel, and are identical with, each other. But
while the thinking mode tm₁ that is my mind represents my body, em₁, and nothing else
(2p13), it is also identical with zm₁. However, it appears not to represent zm₁ at all! Or,
more carefully, if tm₁ represents zm₁, Spinoza claims that I remain, in principle, wholly
unaware of it.

Spinoza’s correspondent Tschirnhaus asked for an explanation of this apparent
inconsistency, writing:

Will you please let me have a proof of your assertion that the soul cannot perceive
any more attributes of God than Extension and Thought? Although I can
understand this quite clearly, yet I think that the contrary may be deduced from
the Scholium or Proposition 7, part II of the Ethics.”78

Spinoza’s response to Tschirnhaus in Letter 66, which Melamed expands upon
and defends, runs as follows:

However, in reply to your objection, I say that although each thing is expressed in
infinite modes in the infinite intellect of God, the infinite ideas in which it is
expressed cannot constitute one and the same mind of a singular thing, but an
infinity of minds. For each of these infinite ideas has no connection with the
others, as I have explained in that same Scholium to Proposition 7, Part II of the
Ethics, and as is evident from Proposition 10, part I. If you give a little attention
to these, you will see that no difficulty remains, etc.

Bennett notes, in calling attention to what he perceives as the inadequacy of
Spinoza’s response: “In the two late letters he tries to avoid this conclusion [if it exists,
humans must know the third attribute] by a move which is so abrupt, ad hoc, and
unexplained that we cannot even be sure whether it is a retraction of the metaphysics or

78 From Tschirnhaus’s letter to Spinoza, as quoted in Melamed, “Spinoza’s Metaphysics of
Thought,” 656.
of the epistemology.” By contrast, Melamed is significantly more optimistic. He defends Spinoza by first arguing for a related thesis: that Spinoza’s famed doctrine of Parallelism, as stated in 2p7 and 2p7s, is really two distinct doctrines, each of which is textually and philosophically independent.

More interestingly for our purposes, Melamed’s defense of Spinoza’s response crucially relies on the explanatory barrier between the attributes. Rather than rehearsing his whole defense here, I will first state his solution to Tschirnhaus’s challenge. And while I do not find his arguments for the existence of two independent parallelism doctrines in Spinoza convincing, I think they arise from salutary motivations. His solution, I think, is best interpreted as a claim about the relative fundamentality of two aspects or more precise readings of Spinoza’s unified Parallelism doctrine. Melamed advances this relative fundamentality claim, perhaps, to forestall worries that Parallelism among modes of all attributes arises solely because of an interesting feature of Thought. That is, Melamed wants to use the EB to answer Tschirnhaus’s challenge, while precluding Spinoza’s being any sort of Thought-centric, reductive idealist—as Newlands did in the previous section. Seeing the explanatory barrier thus operating again to preclude Spinozistic idealism and to resolve a longstanding puzzle in Spinoza’s metaphysics of ideas offers another occasion to ask about its defense.

Melamed’s answers Tschirnhaus’s question extremely elegantly on Spinoza’s behalf. He argues:

Spinoza’s invoking of E1p10 in the context of our impossibility of having

---

79 Bennett, “A Study of Spinoza’s Ethics,” 78
80 Melamed, “Spinoza’s Metaphysics of Thought,”
knowledge or ideas of other attributes seems to imply that not only modes of different attributes are separate from each other, but also ideas of modes of different attributes—though belonging to the same attribute—are conceptually separate from each other.”

As modes of different attributes can be conceptually and causally isolated, as the EB states, so too can ideas or representations of modes of different attributes be causally and conceptually isolated. More carefully, my idea or representation of extension is my mind, and what it represents is my body. But there is another aspect of the same idea, which is identical to my mind, which represents a mode of the third attribute, rather than my body. A version of the explanatory barrier operates within Thought, among aspects of ideas that represent modes of different attributes, and precludes their causal or conceptual interaction. This limits human knowledge of the other attributes. For Spinoza claims that the “object of the idea constituting the human mind is the body, or a certain mode of extension which actually exists, and nothing else” (2p13). But, as in the inter-attributes mode case, this causal and conceptual isolation is not enough to guarantee nonidentity. Indeed, after one grants the parallelism of the various idea-aspects, it becomes difficult to resist their identification, for the same reasons as those at play in Della Rocca’s Leibniz’s Law argument for Mind-Body Identity.

To fully appreciate Melamed’s reply, we require a basic understanding of his theory of idea aspects. Melamed first distinguishes between modes under specific attributes, and modes under every attribute, calling the former ‘Modes of an Attribute’ and the latter ‘Modes of God’. This distinction neatly parallels Della Rocca’s notion of extensional/neutral and intensional properties. When we are speaking of a Mode of an

81 Melamed, “Spinoza’s Metaphysics of Thought,” 656
Attribute, the properties we ascribe are intensional in the way defined earlier, whereas when we are speaking of a Mode of God, the properties ascribed are neutral or extensional. These latter properties are those used, in conjunction with the EB and Della Rocca’s version of Leibniz’s law, to establish mode identity. In Melamed’s reading of Spinoza, then, one and the same thing is expressed in infinitely many aspects, which are “attribute-specific units…of the same thing under all attributes.”

A closely related distinction arises similarly within the attribute of Thought, which has a special richness for Spinoza, on Melamed’s reading. This occurs because, as Melamed emphasizes, 2p7 states: the order and connection of ideas is the same as the order and connection of things. That is, the primary function of thinking modes is to mirror the total ordering of all things, including those relations of causal and conceptual isolation between each mode’s attribute-specific properties. So, Melamed thinks:

Each idea has infinitely many aspects, so that each idea-aspect represents uniquely a mode of God under a particular attribute (i.e. a mode of an attribute). These idea-aspects relate to the infinitely faceted idea in the same way the infinitely many modes of attributes relate to a mode of God, and the infinitely many attributes relate to God. In all three cases we have a res (Substance, Mode of God, the infinitely faceted idea) having infinitely many aspects (each aspect tied to a particular attribute, and each aspect being causally and conceptually isolated from the other.

Melamed’s reading has significant advantages. For one, it allows him to precisely state how exactly the order and connection of ideas is the same as that of things. It also shows how Spinoza can justifiably appeal to the EB in his reply to Tschirnhaus to guarantee that the same causal and conceptual isolation that arises between modes of

82 Melamed, “Spinoza’s Metaphysics of Thought,” 646
83 Melamed, “Spinoza’s Metaphysics of Thought,” 654
different attributes must arise within Thought. Again, Thought’s primary job is to mirror the order and connection of things, as 2p7 states. So it is no wonder that, like Modes of God having multiple aspects (multiple modes of an attribute for the one mode of God), so too ideas have multiple distinct representations of the infinitely many modes of other attributes. These include Extension (bodies), Thought (ideas of ideas), and all the others (those unknown to human minds).

Melamed develops this reading, as previously mentioned, as a consequence of his broader claim that Spinoza’s Parallelism really splits into two, philosophically independent doctrines. He labels these Ideas-Things Parallelism and Inter-Attributes Parallelism. Inter-Attributes Parallelism claims: “The causal order of things in each attribute corresponds to the causal order of things in each of the other attributes. In addition, parallel items are just different aspects of one and the same thing.”

Ideas-Things Parallelism, by contrast, is a more specific claim, which states: “The causal order of ideas corresponds to the causal order of things. In addition, ideas represent the things that parallel them.” He offers at least four sophisticated arguments for this bifurcation claim, all of which, however, I find ultimately ineffective.

84 Melamed, “Spinoza’s Metaphysics of Thought,” 651
85 Melamed, “Spinoza’s Metaphysics of Thought,” 651
86 Melamed first gives examples where one parallelism might hold but the other might fail to do so. These examples, however, do not represent genuine possibilities, for Melamed later grants (673) that (a) Spinoza endorses both of them and also that (b) both theses are clearly necessary truths, so that (c) in the broadest sense neither could have been otherwise. Second, he discusses differences in how Spinoza cites 2p7 and 2p7s, but since he does not regard this as conclusive (642), neither do I. Third, he gives an interesting argument that Ideas-Things Parallelism cannot always involve identity, for some ideas cannot be identical to their objects (i.e. Idea Dei, the idea of God). This is so because, “According to Spinoza, the idea of God is an infinite mode of Thought. The object of this idea is God. Since God is a substance and the idea of God merely a mode… it seems that there cannot be a relation of identity between the two” (641). But this seems wrong. For 2p3 states that in God “there is necessarily an idea, both of his essence of and of everything that necessarily follows from his essence.” And 2p4dem suggests that an infinite intellect
I think Melamed erroneously postulates a bifurcated Parallelism doctrine because he does not recognize the full scope of the representational requirement of what he calls Ideas-Things Parallelism. He overlooks a crucial feature of the EB in its relation to Thought. Namely, the combination of the EB and of the representational nature of thought requires that all relations, including attribute-neutral relations of paralleling, be mirrored in Thought. So Thought represents all identity and non-identity relations, as well as any relations of isolation between modes of various attributes. This holds regardless of whether one has already established any parallelism thesis—the paralleling relations exist quite apart from any justifications for a particular philosophical claim.

This allows the most promising objection to Melamed—that Inter-Attributes Parallelism is derivable from Ideas-Things Parallelism—to succeed. 87

I think Melamed erroneously postulates a bifurcated Parallelism doctrine because he does not recognize the full scope of the representational requirement of what he calls Ideas-Things Parallelism. He overlooks a crucial feature of the EB in its relation to Thought. Namely, the combination of the EB and of the representational nature of thought requires that all relations, including attribute-neutral relations of paralleling, be mirrored in Thought. So Thought represents all identity and non-identity relations, as well as any relations of isolation between modes of various attributes. This holds regardless of whether one has already established any parallelism thesis—the paralleling relations exist quite apart from any justifications for a particular philosophical claim.

This allows the most promising objection to Melamed—that Inter-Attributes Parallelism is derivable from Ideas-Things Parallelism—to succeed. 87

87 Ideas-Things Parallelism states that the “order and connection of ideas is the same as the order and connection of things (2p7). This initially appears to state a claim about two domains (ideas and things), as opposed to the infinitely many domains characterized in Inter-Attributes Parallelism (the infinite domains of modes of each attribute). But appearances are misleading. The identity here holds between the whole order and connection of ideas, on one hand, and the order and connection of all things, including ideas plus all non-Thinking modal orders, on the other. So the domain of things is a collection of infinitely many other domains. Furthermore, these orders and connections include both intra-attribute relationships, as well as attribute-neutral ones. For instance, these orders among things include having a certain number of immediate causes and effects, having certain temporal properties, and, crucially, paralleling or being identical to certain other modes!

This last aspect is key. Ideas-Things Parallelism requires that, if a certain extended mode em1 is identical to zm1 (as Inter-Attributes Parallelism holds), then this identity relation between em1 and zm1 must be present in Thought. And, as we saw in section 1.5, the identity relation obtains whenever all attribute-neutral properties are shared between ‘two’ modes of different attributes. But representing the identity between em1 and zm1 occurs within Spinoza’s larger substance-mode ontology. This significantly constrains the candidates for doing the representing.

So if em1=tm1 and em1=zm1, some Thinking mode or modes represents both relations, by Ideas-Things. Further, since em1 parallels tm1, tm1 represents em1, also by Ideas-Things. But representing em1 requires representing its relations, both to other extended modes and to modes of attribute z, like zm1, to which it is identical. However, by 2p13, tm1, considered as my mind, cannot represent the em1-zm1 relation, since the human mind is ignorant of attributes other than Thought and Extension. Melamed’s
Rather than focusing on his unsuccessful argument for the existence of two Parallelism doctrines, then, we might more productively see Melamed as arguing for the \textit{relative fundamentality} of Inter-Attributes Parallelism to Ideas-Things Parallelism. He sometimes seems to argue precisely in this way:

The identity of modes of different attributes was arguably proven in E1p16, and it is only recalled in 2p7s. We may push further and ask what are Spinoza’s reasons for accepting the infinitely faceted structure of modes in E1p16, and the answer seems to be given by the only text invoked in E1p16d: the definition of God (E1d6). Indeed the duplicated infinity is already asserted in this crucial definition.\footnote{Melamed, “Spinoza’s Metaphysics of Thought,” 650}

Melamed’s thought seems to be that 1d6’s definition of God grounds the proof for God’s existence in tandem with other independent theses. This conclusion of 1p14 gives us 1p16, where the identity of modes of different attributes is asserted, and it is this direct consequence of Spinoza’s definition of God that gives us the Inter-Attributes Parallelism.

\begin{quote}
88 Melamed, “Spinoza’s Metaphysics of Thought,” 650
\end{quote}
Ideas-Things Parallelism is downstream of Inter-Attributes Parallelism, which is a more fundamental consequence of God’s nature. But if this is Melamed’s point, then he has put it quite carelessly. For the fact that Ideas-Things follows from Inter-Attributes by no means removes the possibility that Inter-Attributes also follows from Ideas-Things; that is, the two theses might be inter-derivable! If Inter-Attributes Parallelism is ultimately just a consequence of 1d6 and the fact that Spinoza has, on unrelated grounds, proven 1p14, then it of course follows that, assuming 1p14 and 1d6 and 2p7 (Ideas-Things), one can derive 2p7s (Inter-Attributes). After all, Melamed, quoted above, has just said that 2p7s is derivable from 1p14 and 1d6 alone!

Moreover, it is not obvious that Inter-Attributes Parallelism is more fundamental than Ideas-Things Parallelism. Spinoza explicitly claims that 2p7 follows from 1Ax4, which states that knowledge of an effect depends on knowledge of its cause. This is closely related to 1Ax3, which states, “From a given cause an effect necessarily follows; and, on the other hand, if no definite cause by granted, it is impossible that an effect can follow.” The relations among 1Ax3, 1Ax4, and 2p7 might be expressed in this way. First, 1Ax3 states a version of the principle of sufficient reason. 89 1Ax4 states that knowledge of these effects depends on knowledge of these causes, or that identifying the sufficient reason for some effect requires knowing its cause. Finally, 2p7 makes clear that knowledge, in the sense of clear and distinct ideas existing at least in God’s mind (as all-knowing), always mirrors the existing causal order of modes under any attribute, that is, of all things.

89 See Garrett, “Spinoza’s ‘Ontological’ Argument,” 202, for the same claim.
So we have a set of tight interrelationships among Spinoza’s idea of God, His Intellect, and what it represents. This ties in to Spinoza’s commitment to the PSR and the EB. These play crucial roles in securing the existence of the God whose properties—as Melamed notes concerning Inter-Attributes Parallelism in 1p16 and as I have just noted concerning 1Ax4 and 2p7—generate the relevant parallelisms. One option is that Inter-Attributes Parallelism should be read as a direct consequence of some facts about God’s infinite nature, and Ideas-Things as a more distant consequence thereof. A second is that God’s existence and a fact about knowledge stated in 1Ax4 can generate Ideas-Things Parallelism, from which Inter-Attributes Parallelism can be derived.

Choosing between the two options seems immaterial. Each parallelism doctrine is closely connected both to the other, and to the PSR and EB, which are the most basic principles of Spinoza’s system. It is no wonder that questions of relative fundamentality, at this deep level, are difficult to adjudicate. But we should still carefully distinguishing between difficult questions of relative fundamentality, on the one hand, and the question of whether Inter-Attributes Parallelism is derivable from Ideas-Things Parallelism, on the other. This latter question, for both the simpler and more nuanced reasons given here, deserves an unqualified ‘yes’.

Finally, we might speculate why Melamed takes such pains to argue that Inter-Attributes Parallelism is not derivable from Ideas-Things. This should help reiterate the overwhelmingly beneficial aspects of Melamed’s elegant defense of Spinoza’s response to Tschirnhaus’s challenge concerning ignorance of other attributes. My suggestion is that Melamed, like Newlands in the previous section, wants primarily to combat any reading of Spinoza that pushes him towards idealism. Asserting the primacy or even co-
fundamentality of Ideas-Things parallelism, one might think, makes the pervasive inter-attribute mode identity depend too heavily on an interesting property of thinking modes. Namely, it results from their order and connection being the same as the order and connection of things. If this is part of Melamed’s motivation, then it is a salutary part. For, as Newlands and Melamed both convincingly argue, an Idealist Spinoza is an inconsistent one. Melamed puts it this way:

The answer [to why Spinoza cannot be a Reductive Idealist] seems to be quite simple. Spinoza could not embrace idealism because of his commitment to the conceptual separation of the attributes. Had any of the other attributes been reduced to Thought, this attribute would not be “conceived through itself” (E1p10), and hence it would cease to be an attribute. The conceptual barrier between the attributes blocks the possibility of reducing any of the attributes to another attribute.90

I wholeheartedly concur with this straightforward and powerful statement of the EB’s role in blocking the inference to any sort of reductive idealism. I simply reject that a multiplicity of parallelism doctrines, or a strict emphasis on one aspect of Parallelism at the expense of another, is required to block the inference. Most importantly, I also see the now-familiar challenge in Melamed’s invocation of the EB for precluding idealism: to find some principled motivation for the EB.

This concludes our survey of the various and occasionally surprising roles that philosophers have found for the explanatory barrier between the attributes in Spinoza’s thought. After introducing the explanatory barrier, we have seen that Spinoza’s proof for Substance Monism, his argument for Parallelism, and his affirmation of Mind-Body Identity all rely at crucial stages on the EB. We have further considered how the

90 Melamed, “Spinoza’s Metaphysics of Thought,” 676
explanatory barrier between the attributes might be invoked to resolve longstanding
riddles in Spinoza scholarship concerning idealism and ignorance of other attributes.
Before turning to my novel Spinozistic defense of the EB, which I state as a preemptive
challenge to a sort of explanatory skepticism, we must examine, in the next chapter,
whether Spinoza responds successfully to skepticism in completely different ways. If so,
this would strongly suggest that the EB is not introduced primarily for this purpose,
contrary to the arguments I present in chapter three.
CHAPTER 2:
SPINOZA AND SKEPTICISM

2.1 Introduction

The mature Spinoza, author of the *Ethics*, apparently cares little for epistemology. He has been called many things—atheist, God-drunk pantheist, and mystic—but most interpreters agree that questions about knowledge are not at the heart of Spinoza’s philosophical projects.\(^9\) Indeed, even in his earlier *Treatise on the Emendation of the Intellect*, which is much more explicitly concerned with the truth of and justification for our ideas, Spinoza conceives of his task in decidedly non-epistemological terms:

> “Meanwhile man conceives a human nature much stronger and more enduring than his own, and at the same time sees that nothing prevents his acquiring such a nature…This, then, is the end I aim at: to acquire such a nature, and to strive that many acquire it with me. That is, it is part of my happiness to take pains that many others may understand as I understand, so that their intellect and desire agree entirely with my intellect and desire. To do this it is necessary, first to understand as much of Nature as suffices for acquiring such a nature; next, to form a society of the kind that is desirable, so that as many as possible may attain it as easily and surely as possible” (*TIE* 13-14, G II/8-9).

In this passage, self-improvement followed by social transformation animates Spinoza’s search for knowledge. He aims to strengthen human nature in himself, to make

\(^9\) For the original attribution of God-intoxication, see Carlyle "Novalis."; Nadler, *Spinoza: A Life*, as well as “Spinoza the atheist” discusses the propriety of labeling Spinoza as either a pantheist or atheist, and comes down on the side of atheism. Perler, “Spinoza on Skepticism,” whom I’ll discuss at greater length later, notes the attractiveness of considering Spinoza as a mystic specifically with respect to epistemological questions, before ultimately rejecting this view.
it more permanent, and to create the sort of social arrangement in which the study of nature, in a manner like that of modern scientific communities, unifies the participants within it. He idealizes the pursuit of pragmatic goals within a collective society devoted to discovering truths about the natural world. The solitary thinker of Descartes’ *Meditations*, consumed by skeptical doubts, seems wholly absent from his vision.

Spinoza’s sharp methodological differences with Descartes have not, of course, escaped critical notice. Nevertheless, philosophers wanting to engage with Spinoza’s views on questions concerning skepticism have mostly focused their attention on this early *Treatise* or on Spinoza’s commentary on Descartes. Nearly (but not quite) all of the examples I discuss in this chapter follow this pattern. While certainly understandable, I think this tendency to focus on Cartesian questions is mistaken, and defending this claim is the first task of this chapter. I begin here for three reasons. First, if Spinoza is seen primarily as engaging with radical Cartesian skepticism, then his arguments are quite bad—they have hardly any anti-skeptical force at all. More specifically, I argue that almost all extant interpretations of Spinoza’s arguments of this sort fall prey to the same problem. Each argument faces, in a sense to be defined more precisely below, the dilemma of being either redundant or impotent.92

Second, this tendency causes philosophers interested in Spinoza’s philosophy to miss his best insights. Interpreters have been distracted with epistemological problems quite foreign to Spinoza’s true concerns, and which arise only with great difficulty within

92 I borrow these terms from Craig, “Nozick and the Skeptic: the Thumbnail Version,” who speaks of a similar dilemma for Nozick’s anti-skeptical strategy. Stern, “Transcendental Arguments: A Plea for Modesty,” also employs a similar conception of redundancy for ambitious transcendental arguments, which share many of the features I’ll be concerned with here.
his mature metaphysical system. They have missed the innovative justification for
isolation between the attributes that Spinoza first explores in the *Short Treatise on God, Man, and His Well-Being*, and implements in the *Ethics*.

This difference between my interpretation and others bears emphasizing. I contend that Spinoza is uninterested in directly engaging the radical Cartesian skeptic on his own terms. If we must read him this way, then his arguments against the skeptic fail. Thankfully, this way of reading Spinoza is neither necessary nor most fruitful. Read otherwise, his arguments regain their force. I argue in chapter three that Spinoza aims to point out irremediable explanatory problems *within other anti-skeptic’s* (whether *Descartes’s, or some non-Spinozistic naturalist system*) favored solutions to skeptical problems. Defending this anti-skepticism, for Descartes, involves mustering the full resources of his mature metaphysics. Yet, Spinoza argues that the Cartesian or non-Spinozistic naturalistic metaphysical systems used to defeat the skeptic are explanatorily inadequate. They run the risk of creating as many problems as they solve. *Given* this explanatory inadequacy, Spinoza thinks, their proponents should abandon these metaphysical systems in favor of his own. His system, armed with the explanatory barrier, avoids these tensions.

I have just spoken of Spinoza’s identification of ‘irremediable explanatory problems’ within competing metaphysical systems. But this is not the only perspicacious way of talking about the claims I’m interested in. Indeed, in one of the central passages where Spinoza first worries explicitly about the problems besetting Cartesian interactionism, he states the issue as a sort of skeptical challenge.93 This is no accident.

---

93 *KV* I/97; I focus on this passage at length in the next chapter.
Serious explanatory problems can easily morph into skeptical worries as follows. Consider any system or theory S, whether S is Newtonian physics, Cartesian interactionism, or a statistical model for forecasting American elections. Suppose S cannot explain a large swath of the data, which appears similar to data it purports tocompetently organize. If it is the only system relied upon for judgments about the data in question, some doubts naturally emerge. Why should we be so confident in S’s other claims or predictions, given that it cannot explain such closely related phenomena? Suppose we have relied on S’s claims in the past. Were we justified in doing so, given S’s inability to capture anything like the full picture? Were S’s past successes only happy accidents, consequences of the data we chose to consider? What if the system is not just incomplete, but is producing false results? How can we be reasonably confident in S’s accuracy, given its radical incompleteness, provided there is no additional system capable of explaining the recalcitrant data?

Nonetheless, the force of Spinoza’s charge is most apparent if his challenge is stated in terms of explanation failure. For ultimately, Spinoza is not a skeptic, nor does he think anyone else should be. He does believe, however, that Descartes launched his anti-skeptical project from unstable interactionist foundations. Spinoza thinks that only a system buttressed with the explanatory barrier between the attributes meets the explanatory demands he thinks need meeting. But this causes problems not only for dualists, but also for physicalists who accept distinct mental properties as well, as we will see in later chapters.

This leads to my final reason for beginning with diagnosing the failure of previous interpreters’ reconstructions of Spinoza’s anti-skeptical arguments. In chapter
one, I argued that nearly every important philosophical thesis that Spinoza holds arises, in some intimate sense, from the explanatory barrier between the attributes. Later, I show how Spinoza introduces the EB to help justify the collective explanatory project of those interested in acquiring the aforementioned “stronger and more enduring human nature [whose] intellect and desire agrees entirely” with those of others. If, by contrast, Spinoza actually seeks to respond to skepticism in the way that this chapter’s commentators have thought, then the EB is wholly unnecessary for solving at least that problem. But then the EB remains in the state I left it in the previous chapter—Spinoza lacks both a reasonable defense of its truth and a plausible motivation for introducing it. Instead, I want to clear the way for the EB’s introduction and use in unifying Spinoza’s system by showing precisely how existing discussions of Spinoza and skepticism fall short. My reading focuses our attention on the explanatory adequacy of competing anti-skeptical systems, rather than on the anti-skeptical arguments themselves. It gives the explanatory barrier between the attributes a role and a defense, thereby providing an exciting alternative to the dominant interpretive trend.

This chapter has two main parts. In the first (sections 2.2-2.5), I engage a number of interpreters who see Spinoza as directly responding to a version of radical Cartesian skepticism. In 2.2, I try to identify the sort of skeptic they have in mind, and I suggest one commonsense way of responding to this skeptic. In 2.3, I describe a general problem for combatting Cartesian skepticism of this kind in some philosophical manner, as opposed to commonsense fashion, stated as a dilemma. Direct responses to Cartesian skepticism must avoid the twin vices of redundancy or impotence—I argue that most extant interpretations of Spinoza do not. In 2.4, I briefly mention a general line of defense for
the most controversial premise in my argument for the dilemma, with reference to the epistemology of disagreement. Then, I discuss a first group of interpretations in 2.5 that identifies some local feature of Spinoza’s epistemology or philosophy of mind, and argues that this local feature helps him successfully combat the Cartesian skeptic. I show that this first group of interpretations faces the dilemma outlined in 2.3.

In the chapter’s second part (2.6-2.7), I consider several somewhat more sophisticated attempts to develop a distinctively Spinozistic anti-skepticism. This second group of interpretations focuses more on global, fundamental principles within Spinoza’s metaphysics. They wield these fundamental principles in arguments designed to showcase Spinoza’s preventative strategy against the skeptic, which attempts to silence skeptical arguments before they start. For accepting the principles appears to presuppose our having a great deal of knowledge about the world and its rational structure. Spinoza can then devote his remaining energy to showing the theory-ladenness of the skeptic’s premises. These preventative strategies bear resemblance to my own reading of Spinoza, but I nevertheless argue, in 2.6, that they remain unsuccessful. They also fall prey to the dilemma of redundancy or impotence. Mostly, this results from their failure to appreciate that Spinoza’s real target is not the skeptic, but the anti-skeptic who embraces some alternative to Spinozism. If my arguments succeed, they will have made clear why Spinoza’s strategy requires reinterpretation. They thus clear the ground for a shift in focus in which different texts, as well as the explanatory barrier between the attributes, assume a starring role.
2.2 Defining the Skeptic and a Commonsense Response

Given Spinoza’s astute commentary on Descartes’ philosophy, it is no wonder that most of his interpreters have seen Spinoza as intent, on those few occasions where he focuses primarily on epistemology, to combat a brand of radical Cartesian skepticism. However, commentators differ in how seriously they try to connect Spinoza’s discussion to Descartes’ actual views. Thus, for my purposes it will be helpful to have a general Cartesian skeptic in mind, regardless of his resemblance to any historical position. By a Cartesian skeptic, then, I mean (a) someone who doubts the truth of even his most clear and distinct ideas, (b) within the context of a broadly Cartesian theory of mind. These doubts arise at least primarily (c) due to skepticism about the origins of his cognitive faculties, or, more minimally, about the origins of the beliefs arising from them.94 Though all of (a)-(c) are not stated explicitly in every discussion I consider below, all cooperate to create the picture of the skeptic that Spinoza, it is claimed, has primarily in mind. Cartesian skepticism is true when a skeptical scenario obtains that makes the Cartesian skeptic’s doubts justified.

So, for instance, Willis Doney argues: “Spinoza tacitly accepts this assumption of Descartes’; namely, that ignorance of one’s origin [c] is a necessary as well as sufficient condition of doubt—or, as Spinoza specifies in the Improvement of the Understanding, of genuine doubt—about clear and distinct perceptions [a].”95 For both Descartes and Spinoza, ignorance of one’s origin—that one might be the creation of a deceiving

---

94 In what follows, all discussions of ‘doubt’ and ‘certainty’ concern epistemically normative versions of these attitudes or states. As far as I can tell, the fact that someone is doubtful or feels certain, if this fact is taken as independent of their reasons for doing so, is irrelevant to epistemology.

95 Doney, “Spinoza on Philosophical Skepticism,” 619
demon—provides a legitimate reason for doubting the truth of one’s clearest ideas. Doney later connects Spinoza’s alternative to Descartes’ own anti-skepticism to perceived shortcomings in the Cartesian theory of judgment [b].

Della Rocca, in two papers on the subject of Spinoza and radical skepticism, characterizes the skeptic as “one who doubts whether we have any epistemic purchase on reality at all. This doubt emerges as a doubt even about ideas that, above all other ideas, we take to be true. These are the ideas that are representationally in order [a].” He too notes the connection between Descartes’s bifurcation of will and intellect [b], on one hand, and Spinoza’s rejection of this bifurcation as a consequence of the principles he uses to respond to skepticism, on the other.

Martha Bolton states the position she considers somewhat differently, initially, as “we are unsure whether we have the general ability to discern the truth, and as a consequence, we cannot be certain in any particular case that what we take to be true is really so.” But she too invokes Descartes’ notion of a deceitful god [c], stating: “Although some things are quite evident (clear and distinct) [a] to us…this uncertainty about our cognitive faculties [c] gives a genuine reason to doubt about everything else.” So (a)-(c) seem to capture key features constituting the skeptical opponent against whom Spinoza’s interpreters see him arguing.

---

96 Doney, “Spinoza on Philosophical Skepticism,” 622
97 Della Rocca, “Spinoza and the Metaphysics of Skepticism,” 859
98 Della Rocca, “Spinoza and the Metaphysics of Skepticism,” 855
99 Bolton, “Spinoza on Cartesian Doubt,” 379
100 Bolton, “Spinoza on Cartesian Doubt,” 379
So described, this Cartesian skeptic bears strong resemblance to the external world, brain-in-a-vat (BIV) skeptic who occupies an outsized place in the recent epistemology literature. But there are many sorts of clear and distinct perceptions, so classical Cartesian intuitions shouldn’t focus one’s attention unduly on the problem of the existence of the external world, to the exclusion of related issues. One might clearly and distinctly perceive some logical or mathematical relations, or that one is appropriately applying a given concept. So, minimally, Cartesian skepticism as defined here includes what Williamson calls skepticism about reason as well as skepticism about judgment. Modal intuitions, including descriptions of what is possible, or which possible worlds are nearby worlds, can also be clear and distinct. Cartesian skepticism, as I’ve defined it, threatens them too. This breadth is no surprise—Descartes himself draws similar conclusions in the First Meditation, saying, “may I not similarly go wrong when I add two and three or count the sides of a square, or in some simpler matter, if that is even imaginable?” Given such BIV scenarios, or skepticism about reason, judgment, and/or modality, one might wonder how one could possibly respond to skepticism characterized so radically. It’s not at all easy to see how one might remove doubts about whether our

---

101 For a random sampling of many, see Putnam, *Reason, Truth, and History*, 1-21; Ayers, “What are we to say to the Cartesian skeptic”; Black, “A Moorean Response to Brain in the Vat Skepticism”; and Nozick, *Philosophical Explanations*.

102 Williamson, “Philosophical ‘Intuitions’ and Skepticism about Judgment”

103 This is Craig’s main point in “Nozick on Skepticism: The Thumbnail Version.” Intuitions about which possible worlds are nearby possible worlds are just as vulnerable to skeptical challenge, and so natural evaluations of which worlds are near and distant will typically end up presupposing that the actual world, or some world much like the actual world, is not a skeptical world from the outset.

104 CSM II, 14
thoughts have any justificatory purchase on a world external to us or independent of our representations.105

The most natural, commonsense response to Cartesian skepticism is simply to assume that we have some basic knowledge, and thus that we do not in fact doubt, and are not justified in doubting, the truth of all of our most clear and distinct perceptions. Typically, the class of knowledge claims actually assumed by any particular ordinary person will be implausibly large, but that is no real objection. Minimally, a commonsense response can credit us with knowledge of basic mathematical and logical truths, as well as plenty of basic empirical knowledge of the external world. Namely, commonsense assumes that such a world exists, and that it is populated by objects that have at least some of the properties we typically take them to have.

This commonsense response to Cartesian skepticism has received, and continues to receive, philosophical polishing.106 But sometimes Spinoza seems to endorse a

---

105 This is related to the point Williamson presses about psychologizing evidence. If we are only ever allowed to use as evidence claims to the effect that an agent justifiably believes, or seems to perceive, that X is F, rather than using as evidence the claim that X is F (because the agent knows it), it seems impossible to defeat determined skepticism. Williamson counsels us to avoid trying to defeat it, except perhaps to “save a soul” (“Philosophical ‘Intuitions’ and Skepticism about Judgment,” 124)—this chapter is an extended attempt to show why we can’t help but heed his advice.

106 Moorean, as well as Neo-Moorean (i.e. Sosa, Black, Pryor), responses to Cartesian skepticism exemplify this tendency to add philosophical polish to the commonsense view. Moore’s original strategy was to take the negation of the skeptic’s conclusion, to accept the skeptic’s conditional premise (i.e. if I know I have hands, I know I am not in a Cartesian skeptical scenario), and to argue that the negated conclusion was far more plausible than the skeptic’s initial premise. This combines to create the anti-skeptical argument. Some Neo-Moorean responses can be found, for example, in Black, “The Moorean Response to Skepticism,” Sosa, “How to Defeat Opposition to Moore,” Pryor “What’s Wrong with Moore’s Argument.” Sosa provides a canonical neo-Moorean explanation for the plausibility of skepticism by suggesting that philosophers have wrongly assumed Sensitivity: S’s belief is sensitive iff were it not so that p, S would not believe that p. Instead, Sosa argues, Safety is the more relevant principle: S would believe that p only if it were so that p. Since, he observes, subjunctive conditionals do not contrapose (but one might naturally assume that they do), which would give Sensitivity from Safety, we have both a response to the skeptic and a diagnosis of the skeptical argument’s plausibility. While this latter explanation for our confusion is perhaps philosophically interesting, I doubt that either Moore or his later
reasonably unpolished, commonsense indifference to skeptical problems. In his *Treatise on the Emendation of the Intellect*, Spinoza says: “A true idea (for we have a true idea) is something different than its object” (*TIE* II/14, 13-14). After assuming that we have true ideas, he later claims that truth suffices for certainty: “And from this, again, it is clear that, for the certainty of the truth, no other sign is needed than having a true idea” (*TIE* II/15, 9-11). Lest one think this blithe indifference confined to Spinoza’s early work, in 2p43 of the *Ethics* he basically repeats the first quote, saying: “He who has a true idea…cannot doubt the truth of the thing.”

In another passage, Spinoza confronts the objection that any truth-seeking method will require some justification, which itself requires further justification, leading to an infinite regress that prevents our coming to know the truth. There, Spinoza gives this extended, snarky rebuttal of the cogency of the skeptic’s position, couched in commonsense language:

> Matters stand here as they do with corporeal tools, where someone might argue in the same way. For to forge iron a hammer is needed; and to have a hammer, it must be made; for this another hammer, and other tools are needed; and to have these tools, too, other tools will be needed, and so on to infinity; in this way someone might try, in vain, to prove that men have no power of forging iron (*TIE* II/13, 24-29).

Here, Spinoza expresses commonsense incredulity at the skeptical consequences of requiring antecedent justification for the most ordinary facts of life, as prerequisite for getting on with the business of believing and acting.

followers add much to the naïve commonsense response that most ordinary college undergraduates find compelling.
On the basis of quotes like these, Spinoza’s interpreters have accused him of widespread confusion or at least unjustified overconfidence with respect to skeptical problems. Curley and Cook accuse Spinoza of simply “ignoring or dismissing the skeptical challenge.”\(^{107}\) Jonathan Bennett says: “I think that Spinoza is muddled about skepticism.”\(^{108}\) Hubbeling suggests: “Spinoza has not really struggled with doubt as for example Descartes did…His rejection of Skepticism is harsh and by no means does justice to the often sincere and deep-going struggles of the Skeptics.”\(^{109}\)

Even if this interpretive judgment is unduly harsh, there is a more serious problem looming. Philosophers dissatisfied with the commonsense response to Cartesian skepticism face the challenge of giving some philosophical argument against it. But, to put my cards on the table, I am quite pessimistic that a direct philosophical response to the radical Cartesian skeptic can succeed, and am even less confident if the Cartesian skeptic is made less dogmatic.\(^{110}\) Of course, the devil is in the details, and sophisticated and promising responses to various skeptical scenarios have been explored in recent

\(^{107}\) See, for example, Curley and Cook, quoted in Perler, “Spinoza on Skepticism,” 2, online. The relevant citations are Curley, *Behind the Geometrical Method*, 65-67 and Cook, *Spinoza’s Ethics*, 82.

\(^{108}\) Bennett, *A Study of Spinoza’s Ethics*, 176—also quoted in Perler, “Spinoza on Skepticism”

\(^{109}\) Hubbeling, *Spinoza’s Methodology*, 35

\(^{110}\) This makes me no radical, however. Timothy Williamson, in the paper quoted previously, accepts much the same conclusion, provided that one rejects his preferred view—our total evidence is all that one knows. Generally, I think, an inch of daylight is all a determined skeptic needs, and this fact, as Williamson argues at length, is insufficiently appreciated in the literature. Finally, to clarify the point about less dogmatic skepticism: I think that certain forms of Pyrrhonian skepticism, where this skepticism is seen as an ability (see Sextus Empiricus, *Outlines*, I/8), will certainly resist any refutation, direct or indirect, as well as any self-refutation. But of course the Pyrrhonian might give Cartesian skeptical arguments, within the broadly Pyrrhonian framework that eschews beliefs (and thus asserting premises of arguments). These versions are less dogmatic, and therefore less open to challenge (i.e. that the skeptic is asserting or presupposing something like a closure principle, which can be shown to be false) than typical Cartesian skeptical arguments. But, to be clear, this Pyrrhonian alternative, or my confidence in its success, is no part of my view in this chapter, or anywhere else. For more on these issues, see Barnes, “The Beliefs of a Pyrrhonist,” and Burnyeat, “Can the Skeptic Live His Skepticism?”
literature. I cannot do justice to most of these arguments here. I do think, however, that most discussions of Spinoza’s response to Cartesian skepticism can be suitably unified, as facing a dilemma of impotence or redundancy.

2.3 The Dilemma for Philosophical (not Commonsense) Anti-Skepticism

Jim Pryor, in a paper defending G.E. Moore’s argument against skepticism, offers a helpful framework for thinking of the ‘redundancy’ horn of my dilemma. He starts with the nagging feeling most readers of Moore can’t shake: Moore simply begs the question against the skeptic. My notion of a response’s being redundant also begins with this rough intuition about anti-skeptical arguments. But we must make the intuition more precise. For none of Spinoza’s interpreters reconstruct his arguments in straightforwardly question begging ways. That is, no argument is such that one of its premises, or anything logically equivalent to it, appears in its conclusion. Pryor shows that one must work hard to give the question begging charge more robust footing. To do so, I’ll present an

111 Pryor considers five alleged types of epistemic dependence an argument might exhibit. I’ll focus on the final two shortly, but will first note a few things about the earlier possibilities. The first two possible types of dependence apply to all deductively valid arguments, and Pryor dismisses their perniciousness for that reason. Type III dependence might seem more serious, since there “for you to have justification to believe the premise, it’s necessary that you have justification to believe the conclusion.” Pryor thinks Type III dependence, too, impugns perfectly respectable arguments, those “where the connection between premise and conclusion is so obvious that understanding the premise well enough to be justified in believing it requires you to take any justification for the premise to also justify you in believing the conclusion” (“What’s Wrong With Moore’s Argument,” 358-59).

Anthony Brueckner (“Cartesian Skepticism, Content Externalism, and Self-Knowledge,” 169-171) defines ‘epistemic circularity’ in precisely this Type III way, and has a related worry: “this objection can be made against any modus tollens argument whose conditional premise is not in dispute. Surely not all such arguments are epistemically circular.” In a footnote he observes that, unless one is prepared to face the argument for question begging directly, one might have no way of blocking the conclusion that all modus tollens arguments with disputed conditional premises are circular. His claim, like Pryor’s, is dismissive but also more expansive—if such arguments are epistemically circular, so too are all or most philosophical arguments. On his behalf, we might add that the skeptic’s own arguments face this challenge too, if the skeptical premises are asserted in the standard way. Many traditional skeptics thus fare no better than the anti-skeptics on this account of epistemic dependence.
argument for the redundancy of a certain category of response to Cartesian skepticism, into which the anti-skeptical arguments considered in this chapter fall.

In speaking of ‘justification’, I’ll follow Pryor in understanding this as “quality that hypotheses possess for you when they’re epistemically likely for you to be true, and thus epistemically appropriate for you to believe [whether or not you do believe them].” I’ll employ a graded notion of justification, and I’ll suppose, as Pryor does, that the connection between prima facie and all-things-considered justification is the following. To the extent that prima facie justification is not defeated by counter-evidence, or to the extent that undermining counter-evidence is itself defeated, prima facie justification confers some degree of all-things-considered justification. Further, I will not be assuming what most philosophers would regard as an unreasonable standard for knowledge, such as infallibilism. Indeed, I’ll try to respect the intuition many philosophers have, which is that the intractability of Cartesian skeptical problems can be attributed, in part, to tacitly presupposing unreasonably strict knowledge standards. Still, it turns out that, unlike the Moorean cases of direct perception or reliance on one’s memory that Pryor primarily discusses, the sources of justification given by Spinoza’s interpreters are highly questionable. So even though I presuppose nothing close to infallibilism here, this won’t help these interpretations much.

112 Pryor, “What’s Wrong with Moore’s Argument,” 352

113 In “Fallibilism and Concessive Knowledge Attributions,” 127, Stanley defines fallibilism as the view: “someone can know that p, even though their evidence for p is logically consistent with not-p.” The infallibilist, by contrast, requires that if one knows that p, one’s evidence is logically inconsistent with not-p, that is, that one’s evidence entails p.
I now argue that any argument with 1) highly philosophical premises and 2) the denial of Cartesian skepticism as its conclusion is redundant or impotent. In giving this argument, I’ll also be providing three criteria for identifying such arguments, which I’ll then use in working through interpretations of Spinoza’s response to skepticism. The argument runs thus:

1. Extant interpretations of Spinoza’s response to skepticism involve attributing arguments with highly philosophical premises to him.
2. This class of highly philosophical anti-skeptical premises requires conservative rather than liberal treatment.
3. If a hypothesis H requires conservative rather than liberal treatment, it requires antecedent justification J as a precondition for being used to justify some other proposition P.
4. If having justification J for believing the conclusion is among the conditions required to antecedently justify H as an argument premise, then the argument displays illicit epistemic dependence (IED).

In addition to Craig, “Nozick on Skepticism: The Thumbnail Version,” several other philosophers appear to have identified a problem similar to what I discuss here. For instance, Anthony Brueckner, “Modest Transcendental Arguments,” 73, says: “The general pattern, then, is to impugn ambitious Kantian transcendental arguments by alleging that they can achieve their goal only by relying upon some highly controversial supplementary premise. Even worse for the transcendental arguer, the supplementary premise turns out to be one which by itself would be sufficient to generate an anti-skeptical conclusion, thereby rendering superfluous the rest of the transcendental argument. For example, the verificationist can argue from the bare meaningfulness of sentences concerning physical objects to the conclusion that standardly available sense-experience counts as evidence which confirms the truth of such sentences. The idealist can argue that given the obtaining of the right patterns of sense-experience, it simply follows that there are physical objects (since, on idealism, facts about objects just consist in experiential facts). Jonathan Vogel, “Dismissing Skeptical Possibilities,” 242 makes similar remarks, saying that most responses to skepticism are “encumbered by questionable philosophical machinery,” the implicit contrast for which appears to be some sort of commonsense anti-skepticism. Being so encumbered, Vogel seems to suggest, makes the arguments redundant or impotent in roughly the way I argue for above.
5. Arguments that display illicit epistemic dependence are either redundant or impotent.

6. Spinoza’s arguments against skepticism, as given in the extant interpretations considered here, are either redundant or impotent.

I’ll take some care in showing (1), but at first glance it shouldn’t be too surprising. Spinoza builds a systematic metaphysical edifice in the *Ethics*, and to some extent in earlier works, so his response to skepticism unsurprisingly has deep roots in the rest of his thinking. At least, that is, if Spinoza’s interpreters are dissatisfied crediting him with nothing more than the commonsense response previously discussed.

The interpretive consensus that this commonsense response is dissatisfying can help clarify the perhaps worryingly vague notion of ‘highly philosophical premises’ employed here. Now I don’t happen to find this notion problematically vague. I think we mostly have a sense of which sorts of knowledge ordinary people commonly assume themselves to have, the assumption of which suffices to defeat the Cartesian skepticism considered here. And we mostly have a sense of how one’s response to skepticism changes after a bit of exposure to philosophy. Either one begins to find the skeptical position significantly more plausible, or one’s account of why skepticism is false becomes far more complicated, often burdened with new theoretical commitments. Of course, only rarely, if at all, do philosophers ever replace their commonsense commitments with new theoretical ones. In practice, philosophers merely tend to rely on both a new, complicated philosophical explanation (professionally), and their same old commonsense commitments (all the rest of the time).

But, for my purposes, I can say at least a bit more to clarify the notion. A highly
philosophical premise is typically (not always) some \textit{a priori} claim that tries to establish some class of knowledge (i.e. empirical, self-knowledge, a priori logic or mathematics) in order to defeat the skeptic. Arguments with highly philosophical premises try to avoid merely proceeding on the assumption that we already know many things. The main premise typically involves specialized content, and appeals to some general feature of our thought, language, or cognitive apparatus. It’s the sort of premise that will seem sufficiently substantive to satisfactorily attribute to Spinoza without his thereby appearing to just boringly beg the question against the skeptic. Unlike with ordinary knowledge claims—‘I know she has two dogs and four cats’—establishing a highly philosophical premise takes work, as typically there is considerable philosophical disagreement about its truth. This last aspect will be particularly relevant for my purposes.

Premise (2) requires sustained discussion and defense—I’ll return to it below. But I need to first clarify the terminology used in premises (2) and (3), which I again borrow from Pryor. Start by thinking of the set of conditions M whose truth confers prima facie justification on your belief. (This translates to all-things-considered justification, provided this initial justification is not undermined, or the undermining evidence is itself defeated). I’ll be assuming that the highly philosophical premises (HPP) in the arguments I consider are part of set M. That is, their truth is part of what grants justification for Spinoza’s anti-skepticism, according to his interpreters.

We can ask a further question about HPP—namely, whether they require conservative or liberal treatment. For Pryor, “a theory treats H \textit{conservatively} when it says that you need some justification to believe H in order to have a given kind of prima facie justification to believe P…[whereas] a theory that treats H \textit{liberally} denies that
having prima facie justification to believe P requires you to have antecedent justification to believe H. But it does count not-H as an undermining hypothesis; evidence against H undermines your prima facie justification to believe P.\textsuperscript{115} Since this distinction is crucial for understanding my argument, it’ll be helpful to see the contrast in action.

Conservative and liberal treatments are appropriate for different sorts of situations. Moreover, whether a hypothesis deserves liberal or conservative treatment can vary based on context. One of Pryor’s examples where \textit{liberality} is apt is a case of reading a proof of the Pythagorean theorem:

H1 is the claim that you understand and correctly follow the proof. Presumably, for you to be justified in believing the theorem [P], H1 does have to be true. But you don’t need to have evidence that H1 is true. It’s the proof itself that justifies you in believing the theorem [P]. H1 is just some condition that enables that to happen…So the right treatment of H1 seems to be a non-conservative one.\textsuperscript{116}

Similarly, Pryor argues, direct perception should be treated liberally. One need not antecedently ensure the reliability of one’s perceptual faculties in order to accord their deliverances some prima facie justification. Of course, strong evidence for some skeptical hypothesis might undermine belief in one’s perceptual faculties.\textsuperscript{117} Certainly, if Pryor receives evidence that he has been kidnapped and drugged by government agents, and then fed LSD, he should question this prima facie justification. So Pryor’s dogmatism about perception doesn’t stretch indefinitely far.

\textsuperscript{115} Pryor, “What’s Wrong with Moore’s Argument?” 353-354

\textsuperscript{116} Pryor, “What’s Wrong with Moore’s Argument?” 354

\textsuperscript{117} That is, Pryor’s reading of Moore sees skeptical hypotheses as potential defeaters for prima facie justified direct perceptual beliefs, but he denies that the anti-skeptic must rule out skeptical hypotheses antecedently, as a condition for according direct perceptual beliefs prima facie justification. The further thought, presumably, is that no skeptical hypotheses suitably undermine perception’s prima facie justification in enough cases.
But suppose that $H_1$ is a claim about a child’s or a schizophrenic’s understanding of some mathematical proof. Consider Anthony Hopkins playing a former genius mathematician slowly going insane in *Proof*, or John Nash in *A Beautiful Mind*. This changes the context and makes a conservative treatment more appropriate. For suppose I am totally incapable of judging whether the theorem $P$ is true, and thus whether Hopkins’, say, is justified in believing it. Still, if $H_1$ is the claim that Hopkins understands and correctly follows the proof, then I think some further evidence for $H_1$ is antecedently required, as a condition of my believing the theorem $P$. $H_1$ might have been amenable to liberal treatment in Hopkins’ younger days. But it now seems like both he, and I as an outside observer, need some antecedent assurance that his mind hasn’t failed as an antecedent condition for being prima facie justified in believing the theorem $P$. The context changes the reasonable treatment for the hypothesis. That should help clarify premise (3).

I introduce premise (4) by beginning with examples where conservative treatment of some hypothesis $H$ seems justified. This should help clarify the illicit epistemic dependence (*IED*) I see operating in the arguments given by Spinoza’s interpreters. As I noted earlier, and discussed at length in footnote 19, some initial attempts to formulate *IED* as brute question begging clearly go awry. At least, they seem to have catastrophic

---

118 One might try to interpret this shift from liberal to conservative treatment in different contexts as evidence that a conditional analysis is most appropriate. Perhaps, that is, the correct analysis is something like: If my faculties are reliable, then $H_1$ is satisfied and should be treated liberally. Then, in order to discharge the antecedent, one requires prior justification for the reliability of the faculties (thanks to Samuel Newlands for this objection). I think this amount to treating no, or at any rate very few, hypotheses liberally. Moreover, it’s quite difficult to see what one’s evidence for the reliability of one’s faculties could possibly be. After all, any evidence would have to be evaluated using those very faculties. This makes the skeptical problem basically insoluble, which makes establishing my thesis here—that Spinoza’s anti-skeptical arguments are either impotent or redundant—even easier. I welcome the result, but nothing I say relies on this being the correct interpretation.
consequences for all philosophical argumentation, not localized to discussions of skepticism. I will not be using these notions here—my notion is more demanding. Here is a homey example argument, containing a premise demanding conservative treatment, which as a whole displays IED.

1. I intend to walk to lot 15 and drive home.

2. So I will walk to lot 15 and drive home.

3. So my car will still be in lot 15 when I get there.\textsuperscript{119}

Pryor thinks that arguments like this exhibit what he calls the fifth sort of epistemic dependence, which occurs when “having justification to believe the conclusion is \textit{among the conditions that make you have} the justification you purport to have for the premise. That is, whenever you need antecedent justification to believe the conclusion, as condition for having that justification for the premise.”\textsuperscript{120} But in order to believe (8), as Pryor suggests, you must have antecedent justification to believe the conclusion, “and for that reason [(6)] can’t make [(9)] any more credible for you.”\textsuperscript{121} You must assume, and be justified in so assuming, that your car will still be in the lot in order to be justified in believing that you will drive it home.

This helps clarify the relation between treating a premise conservatively and illicit epistemic dependence. Let H be the statement requiring conservative treatment, and suppose H is a premise in one’s argument. If the other proposition J, which provides the antecedent justification for premise H, is either the argument’s conclusion, or part of

\textsuperscript{119} Pryor, “What’s Wrong with Moore’s Argument?” 359-360

\textsuperscript{120} Pryor, “What’s Wrong with Moore’s Argument?” 359

\textsuperscript{121} Pryor, “What’s Wrong with Moore’s Argument?” 360
one’s independent justification for believing the conclusion, vicious epistemic dependence results. In our example, (8) is the hypothesis H to be treated conservatively. The argument’s conclusion (9) provides the antecedent justification for (8), and thus (8) must be justified by (9) prior to its use in helping to prove (9), the argument’s conclusion.

This also shows why premise (4) of the argument for redundancy or impotence can be put disjunctively. I state premise 4 as: (a) “If having justification J for believing the conclusion is among the conditions required to antecedently justify H as an argument premise, the argument displays IED.” But an argument where (b) is true, where (b) says that “evidence against the argument’s conclusion undermines, to some extent, one’s putative justification for a premise,” also displays IED, in the cases I consider here. Pryor, by contrast, thinks that in general condition (b) names a permissible form of epistemic dependence, whereas condition (a) clearly restates the illicit fifth type of dependence.

But he thinks this only because he thinks that not all responses to skepticism have premises requiring conservative, rather than liberal, treatment. In particular, he thinks that direct perception prima facie justifies me in believing that I have hands. However, Pryor admits, “If you insist on treating all undermining hypotheses conservatively, then you can maintain that the two types of dependence coincide. Whenever evidence against a conclusion would undermine your justification to believe its premise, you’ll think your justification for the premise must already rest on, and require antecedent justification for, the assumption that the conclusion must be true.”

122 Pryor, “What’s Wrong with Moore’s Argument,” 360
Now I think Pryor is right that we must be willing to treat some hypotheses liberally if we want any hope of defeating the skeptic. So I concur when he says: “I suspect we’ll have no prospect of avoiding skepticism unless we agree to be liberal to some degree, about some kinds of hypotheses. This puts pressure on the non-skeptic to say what’s wrong with the particular kind of liberalism that the perceptual dogmatist espouses.”\textsuperscript{123} It’s important to keep this point about my argument in mind. My claim is NOT that all premises in anti-skeptical arguments should be treated conservatively. As I have already stated, if one wants Spinoza to have a straightforward response to skepticism, the commonsense route is the way to go—certainly, our commonsense beliefs deserve liberal treatment if any hypotheses do.\textsuperscript{124}

Rather, my claim is that a certain class of highly philosophical premises should be treated conservatively. But then for this class, as Pryor admits, the two sorts of dependence coincide. For a premise requiring conservative treatment, either interpretation of IED is appropriate, and one may be more or less intuitive for some readers.

Premise (5)—arguments displaying illicit epistemic dependence are either redundant or impotent—is just a definition, but the negative connotations of ‘redundancy’

\textsuperscript{123} Pryor, “What’s Wrong with Moore’s Argument?” 356

\textsuperscript{124} One might think that this view is in tension with the belief that, for example, our best scientific picture of the world is far more accurate than our set of commonsense beliefs. This is a mistake, for two reasons. For one, scientists themselves give theoretical arguments and present experimental evidence to justify the fundamental laws they use to characterize phenomena. So they themselves treat their claims conservatively, and rightly so, since at least one thing they are interested in is whether these fundamental laws are the right ones. Second, giving theoretical arguments and evaluating experimental evidence presupposes that at least some, and indeed most, of our commonsense beliefs are true or approximately so. Minimally, scientists must be able to trust the deliverances of their senses to some degree when using the various experimental instruments that produce the data used to justify their hypotheses. So the greater accuracy of the scientific world picture with respect to commonsense does not endanger the relative priority for being treated liberally that commonsense still maintains.
and ‘impotence’ can be given some defense. Here’s another example argument, this time more philosophical, with premises one should treat conservatively and the illicit sort of dependence.

1. Justified true belief is insufficient for knowledge.

2. If justified true belief is insufficient for knowledge, then subjects in the standard Gettier cases can have JTB without having knowledge.

3. Subjects in the standard Gettier cases can have JTB without having knowledge.

Suppose I want to know how to interpret the standard Gettier cases after someone describes fake barn country to me. Does this argument convince me that I should follow the mainstream interpretation and think that such subjects can fail to know while having justified true belief? Or suppose I want to know whether my car is still in lot 15. Does being told that I will walk to lot 15 and drive home do anything to make my car’s not being stolen (and thus remaining in lot 15) more credible? Certainly, neither argument sounds very good. But these arguments fail as arguments precisely because arguments are supposed to give us more reason to believe their conclusions than we had prior to being presented with them. Arguments displaying illicit epistemic dependence don’t do what arguments are supposed to do.\textsuperscript{125} I have no additional reason for believing their conclusions than I did before—this makes them redundant. Or, if one (correctly) thinks that illicit epistemic dependence makes an argument bad, then one will find such

\begin{flushright}
\textsuperscript{125} Suppose one rejects my loose characterization of what ‘arguments are supposed to do.’ One might find the notion unclear, or think that arguments are really doing something else, or nothing of the sort. All I need, I think, is that one sees arguments with vicious epistemic dependence as worse than those that lack the property. This might be enough to find something salutary in an interpretation of Spinoza that does not require that he make such arguments.
\end{flushright}

93
viciously epistemically dependent arguments impotent. Either such arguments merely confirm what I already knew about my car, or about the subjects in Gettier cases, and are thus redundant, or they do not establish anything, and are impotent for that reason.

My main claim then becomes: wanting more than dogmatism about commonsense beliefs gets Spinoza’s interpreters into trouble. This gives rise to the dilemma, which I can now state more precisely. By a redundant response to skepticism, I mean: an argument that, in its clearest form, exhibits illicit epistemic dependence (IED). So, the offending premises of some irrelevant response to skepticism (1) will be such that they should be treated conservatively, rather than liberally. Furthermore, (2) justification for the premise to be treated conservatively will require, among its conditions, justification for the argument’s conclusion or some sub-conclusion. Both our previous examples (car lot, Gettier cases) exemplify IED.

By an impotent response to skepticism, I mean: an argument that fails to secure its desired anti-skeptical conclusions. Of course, this is vaguely defined. Typically, however, philosophers otherwise sympathetic to the broad anti-skeptical project will identify impotent responses as failing on their own terms. They reject their central claims independently of any desire to defend the skeptic. Often, perhaps, an anti-skeptical argument might start with seemingly innocuous premises, which upon further reflection conceal a great deal of theoretical baggage. Suitably unpacked, these theoretical assumptions themselves suffice to refute the skeptic, because justification for the anti-skeptical conclusion is among the conditions that provide putative justification for the theoretical premise.
In this case, it seems natural to me to reason from the argument’s redundancy to its impotency, even if one continues, for commonsense reasons, to believe its conclusion. But I distinguish redundancy and impotency because one of the two might seem philosophically worse to some. Redundancy, some might argue, may not be the worst thing, since even straightforwardly question begging arguments can be sound if their inferences are good and their premises are true. On the other hand, one might deny that anti-skeptical arguments with highly philosophical premises displaying vicious epistemic dependence are wholly redundant. Nonetheless, one could still grant that they do very little to make their conclusions more credible. In either of these two cases, redundancy and impotence come apart, but having either feature is a negative quality that interpretations of Spinoza’s arguments, if they can avoid it, should do without.

2.4 Defending Premise 2 of the Redundancy-Impotence Argument

This completes my statement of the dilemma of redundancy or impotence for interpretations of Spinoza’s anti-skepticism. Before running these interpretations through my argument, however, I must defend premise (2)—these highly philosophical hypotheses (the premises of Spinoza’s anti-skeptical arguments) require conservative rather than liberal treatment. We can see the urgency for some defense here. For Spinoza’s interpreters see him as confronting a radical Cartesian skeptic, one who doubts the truth of all his most clear and distinct ideas due to doubts about the origins of his beliefs. Minimally, such a skeptic thinks we lack empirical knowledge of the external world, and probably, such skepticism extends to our reasoning abilities as well. Knowledge of, or justified belief in, the premises of some anti-skeptical argument will be
defeated if the radical Cartesian skeptical hypothesis is true. Said differently, the anti-skeptical conclusion will be among the set of conditions that provide the putative justification for the premises of the anti-skeptical argument. So if such premises should be treated conservatively, Spinoza’s anti-skepticism fails. But why should they be?

I think that starting with the unvarnished intuition is again helpful: highly philosophical premises deserve conservative treatment because, for any philosophical premise, we have good evidence that it is false. Evidence for its falsity consists partially in the fact that no philosophical view commands anything close to general assent. Of course not every philosophical premise can be false, but that does nothing to clarify which ones are. Moreover, there appears to be no reliable method for determining which views are true (not merely plausible or not), or even more likely to be so, at least when choosing among some reasonable contrast class of views. So to use some substantive philosophical claim as an argument premise, we require some antecedent justification of its truth, given evidence for its falsity and the apparent absence of a reliable method for evaluating its truth.\textsuperscript{126} This means we should treat all philosophical premises conservatively. This implies, of course, that the particular premises Spinoza’s interpreters attribute to him should be treated conservatively as well. In the absence of the required antecedent justification, philosophers who disagree about some issue should either ‘move toward’ the opposing viewpoint by adjusting their credence downward, or they should suspend judgment.

\textsuperscript{126} Some beliefs do not command general assent—that anthropogenic global warming is occurring, for example—but there is a method for determining their truth (bracketing global skeptical worries, obviously). Perhaps a belief in other minds is generally accepted, even if there is no good method for establishing its truth. But the combination of widespread disagreement with the methodological impasse worsens matters considerably.
The unvarnished intuition sounds extreme, and it cannot be quite right generally, for at least three reasons. First, commonsense, as roughly characterized earlier, clearly has some commitments plausibly described as philosophical. That’s at least part of what makes G.E. Moore’s fairly sophisticated argument recognizable as an extension of our commonsense beliefs. Even if commonsense is inconsistent or unreflective, it still shares something with more ‘philosophically polished’ anti-skeptical views. And I decidedly do not mean to challenge commonsense, which uncontroversially assumes, rather than argues for, its anti-skepticism.

Second, some truths of mathematics, and certainly logic, are good candidate philosophical truths, and there might seem to be obvious philosophical progress in these domains. This seems broadly correct, but none of Spinoza’s arguments employ premises of this sort, which moreover seem ill suited to doing the anti-skeptical work we are interested in here. Additionally, appreciating philosophy’s success in the logical domain often underscores just how different progress there is from progress on mainstream metaphysical and epistemological questions. As Lycan notes:

There is a corner in which the philosophical track record is good: logic. Otherwise, the history of philosophy is a disgusting mess of squabbling, inconclusion, dogma and counter-dogma, trendy patois, fashionable but actually groundless assumptions, vacillation from one paradigm to another, mere speculation, and sheer abuse. Nothing in that sordid history can be called progress, except what derives directly from developments in logic or in science.¹²⁷

¹²⁷ Lycan, “Bealer on the Possibility of Philosophical Knowledge,” as quoted in Rea, World Without Design: The Ontological Consequences of Naturalism, 149, who approvingly quotes the passage in discussing rational intuition. Since rational intuition is the plausible source of many of the principles discussed in this chapter, Rea’s arguments against its reliability seem particularly appropriate.
Third, there might be a worry that the argument’s second premise, if translated as the unvarnished intuition that all philosophical beliefs deserve conservative treatment, since we have good evidence that they are false, is self-undermining. For it itself appears to be a philosophical premise. So too does the claim that, absent the necessary antecedent justification for premises to be treated conservatively, we should suspend judgment about them or adjust our credence to more closely resemble our opponent’s. As Thomas Kelly and Adam Elga have forcefully argued, no crude ‘split the difference’ view that assigns equal weight to the disagreeing peer’s beliefs, and always recommends moving one’s credence down as a result of disagreement, can be rationally maintained as a completely general position.

Elga shows that such views, held completely generally, are self-undermining. So consider a series of dispute between a confident dogmatist and someone who maintains the conciliatory view (the compromiser) about the proper response to disagreement. The dogmatist’s refusal to adjust his credence downwards, on the subject of disagreement, can force the compromiser to adopt credences about the proper view of disagreement that are arbitrarily close to his. In dispute one, the dogmatist maintains credence 1 while the compromiser has credence .7 about the proper view of disagreement. Splitting the difference requires that the compromiser adopt .85 as his new credence, but if the dogmatist maintains his position, further iterations of dispute result in the compromiser’s adopting dogmatism with respect to the question. This is clearly an unacceptable result for any conciliatory view about disagreement, and thus the view must be restricted.

Perhaps the tension comes out clearer in the following example. “The magazine Consumer Reports rates appliances, and gives recommendations on which ones to buy.
But pretend that, in addition to rating appliances, *Consumer Reports* also rates and recommends consumer ratings magazines. Then it cannot coherently recommend a competing magazine over itself. (By a ‘competing magazine’ I mean a magazine that offers contrary appliance recommendations).”\(^{128}\) It’s easy to see why not. For consider the triad consisting of *Consumer Reports* recommendation of appliance X, a competing magazine’s recommendation of appliance Y, and the *Consumer Reports* assertion that their competitor is the top ratings magazine, whose recommendations should be followed. The first member recommends appliance X, and the conjunction of the latter two members entail *Consumer Reports* recommending Y, which is inconsistent.

Elga here identifies a general problem for epistemic methods—they must be dogmatic about their preferred way of coming to form beliefs, resolve disputes, whatever. *Consumer Reports* must be dogmatic about their method of resolving questions about appliances (to consult *Consumer Reports*). To put this issue in the terms relevant to this chapter, I must be somewhat dogmatic that highly philosophical premises of the class we are interested in should be treated conservatively. This is so even though *that claim itself* is a philosophical one, which requires antecedent defense, perhaps more than I’ve given here. But this—having a reasonably inflexible view about dispute resolution—is simply a demand that any internally consistent position must meet.

These two restrictions to the claim that philosophical premises require conservative treatment thus have very different justifications. I avoid impugning the philosophical commitments of commonsense because I do not wish to criticize a response to skepticism that does not even purport to give an argument for its assumption that

\(^{128}\) Elga, “How to Disagree about How to Disagree,” 180
skepticism is false. Moreover, since coming to believe some more complicated philosophical or scientific view itself often requires that one presuppose the truth of many commonsense beliefs, weakening commonsense only further weakens these more complex views. On the other hand, being at least somewhat dogmatic about the proper method for disagreeing about philosophical premises, by requiring antecedent justification for them or suspension of judgment, is substantially more controversial. But, it appears to be a requirement any internally consistent position must meet. So every view of how best to treat philosophical disagreements will court controversy in the same way, or will be straightforwardly self-undermining.

Nonetheless, I think more can be said in favor of the view that highly philosophical premises, of the sort Spinoza’s interpreters employ on his behalf, should be treated conservatively. At the very least, recent work in the epistemology of disagreement can illuminate that this contention is no radical minority view. It’s admittedly uncommon to grant brute appeals to authority much influence in philosophy. I think it’s informative to note, however, that of ten essays in a recent volume on disagreement, at least five contributors defend more or less pessimistic responses to philosophical disagreement. These range from those arguing that it’s most rational to suspend belief on all philosophical questions, to those who refuse to suspend belief but admit they cannot at all satisfactorily answer the arguments for doing so.129

129 Of the contributors to Feldman and Warfield, Disagreement, Hilary Kornblith, Catherine Elgin, Richard Fumerton, and Adam Elga defend fairly straightforwardly pessimistic responses to disagreement. Elga restricts his view to philosophical questions not containing questions about how to disagree philosophically (on pains of inconsistency, he thinks), and Elgin clarifies that certain practical costs of not accepting philosophical beliefs may be too high to bear, but these are ultimately minor concessions in light of the overall discussion. Peter van Inwagen (“We’re Right. They’re Wrong,” 28) admits he cannot answer those who object to his continuing to hold all his beliefs in the face of disagreement, saying, “I am unwilling to become agnostic about everything except empirically verifiable matters of fact… I am unable
Here, I aim to provide a taste of the reasoning in favor of treating all highly philosophical claims in the anti-skeptical arguments considered here conservatively. This is strictly weaker than the further claim that we should suspend judgment about them, since it merely requires that we have antecedent justification for such a premise in order to justify its use in further reasoning. Still, the combination of the requirement on conservative treatment with the problem of illicit epistemic dependence is a potent one, in the context of the radical sort of skepticism considered here.

This context is important, and differs even from that of ordinary philosophical debates, which further weakens the demand to treat Spinoza’s anti-skeptical premises conservatively. It is unlike the context of debates between causal theory and descriptivist views of reference, where there might be many mutually accepted premises, including a shared commitment to anti-skepticism. No questions are begged, and no argument need display vicious epistemic dependence, in a debate where shared commitment to anti-skepticism, which is required to assert knowledge of much of anything, is assumed by both sides. Causal theorists and descriptivists might still require that, for example, a justified belief in descriptivism not be among the set of conditions justifying any particular premise in the descriptivists argument. Thus, one could accept my claim that highly philosophical premises require conservative treatment, without thereby halting all philosophical investigation until some conclusive refutation of skepticism that avoids these problems is found.
Peter van Inwagen develops perhaps the best case for treating philosophical premises conservatively. He says:

David Lewis believed that infinitely many possible worlds exist, each of them just as real as the actual world. There is no denying that he believed this. Moreover, there is no denying that he was incredibly smart, philosophically gifted, and intellectually responsible. He examined the arguments for and against his position with enormous care. It is no false modesty for me to say that David Lewis was a far better philosopher than I am. Nevertheless, I think he was wrong. I cannot refute his position; it is admirably well defended. But, despite Lewis’s intelligence and arguments, I do not believe that there exist real possible worlds, consisting of material objects and inaccessible from the actual world.\(^{130}\)

Van Inwagen is considering whether the facts of his disagreement with David Lewis compel him to suspend judgment about the question of whether there are infinitely many, concrete possible worlds. Left implicit is the thought that David Lewis might make much the same statement about him. And it’s clear from his discussion that he means the verdict on this question to apply far more broadly, to philosophical disagreement generally. That’s why he later speaks of his reluctance to “become agnostic about everything except empirical matters of fact,”—agnosticism is the suggested stance towards intractable philosophical disagreement. Again, this proposal is substantially stronger than what my argument requires—conservative treatment for philosophical premises. Van Inwagen considers the stronger claim, I think, because he believes that any antecedent justification he offers Lewis, or Lewis offers him, will be insufficient for rationally compelling the other’s assent.

No easy outs are available here. Sometimes, for instance, disagreement presents no real obstacle to maintaining one’s belief. One party has background information that, 

\(^{130}\) Catherine Elgin discusses the version of van Inwagen’s view quoted above, which is presented in “We’re Right. They’re Wrong,” in her “Persistent Disagreement,” 58.
were his opponent to have it, would cause her to drop the dispute. Richard Fumerton
discusses a subtle example of this sort of asymmetric position by considering the well-
known Monty Hall problem. When most people are presented with the problem, they
come quickly to the erroneous conclusion. Suppose you’ve considered the problem
yourself, come to the erroneous conclusion, and only laboriously worked through
explanations for why your initial intuition is misguided. Then you might have evidence in
advance that someone who vehemently persists in believing the incorrect answer would,
were she to go through the same reasoning process as you did, come to hold your
(correct) beliefs about the solution.

No particular background knowledge, but only an experience of error and
subsequent disillusionment through painstaking reasoning, allows you to sanguinely
maintain your belief as to the correct response to the Monty Hall problem. Perhaps, then,
philosophical disagreements are more like this? Maybe philosophers who disagree with
my preferred views, ones I’ve come to after years of contemplation and careful
consideration of the available arguments for and against them, have some defect in
background knowledge. Maybe they’ve failed to go through some necessary process of
reasoning such that, were they to go through it, they would come to agree with me. Van
Inwagen considers this way of avoiding the problem, but he thinks that, at least as far as
many of his most intractable disagreements with David Lewis go, it does not help. Each
of them has painstakingly considered the other’s arguments for concrete possible worlds,
compatibilism about free will, etc., and it is implausible to attribute to either of them
some defect in reasoning or background knowledge.
Michael Rea, in his discussion of the unreliability of the rational intuitions that account for many of our philosophical beliefs, provides a nice supplement to van Inwagen’s argument. He says:

The force of these considerations might be more easily appreciated if we imagine a parallel case arising in connection with vision. Suppose we found that, in most of the visual circumstances we share with other people (i.e. circumstances in which we and our peers have the same regions of spacetime within our visual fields), we are in agreement with many people about our visual beliefs but in disagreement with many others…Suppose further that there is no vision-independent way of resolving the disagreement. As far as we can tell, we are employing the same concepts, but we are all isolated from one another so that our visual circumstances are the only circumstances we share. What would be a rational response? What should we think about RV (the thesis that vision is reliable) in light of this evidence? It seems clear that our rational degree of confidence that RV is greatly diminished by this evidence—unless we have some way of fleshing out the picture that explains the disagreement we observe in a way compatible with RV.131

Again, though, in the philosophical case, unlike with that of vision, we cannot easily insist that the disagreement is illusory in some way or that our opponents are the victims of some pernicious eye disease, as Rea suggests, while we are not. In the best cases, philosophical opponents like Lewis and van Inwagen have treated the other’s arguments with respectful, charitable seriousness, and they appear to have fully appreciated their nuances and motivations. Yet disagreement, with no philosophy-independent way of resolving it, still reins.

Catherine Elgin proposes a way of avoiding the most extreme conclusion, that one should cease holding all one’s philosophical beliefs. She nicely points out that some philosophical positions—Lewis’s concrete possible worlds, eliminativism about the mental, deniers of free will, perhaps—are simply too incredible for some agents to

131 Rea, *World Without Design*, 198
believe. But “since ‘ought’ implies ‘can’, that I cannot believe it entails that it is not the case that I ought to believe it. And that I cannot believe that it might be true entails that it is not the case that I should suspend belief or lower my degree of belief that [in the Lewis case] the only real world is the actual world.”¹³² Now I’m not sure that I accept each step of Elgin’s reasoning here, but for my purposes I can endorse it wholesale. Certainly, Elgin makes plausible the thought that we should focus less on our (probably involuntary) mental states or on the mental states of those with whom we disagree.

Rather, Elgin thinks we should frame the question in terms of ‘acceptance’, rather than belief, where “to believe that p is to feel that p is so. To accept that p is to adopt a policy of being willing to treat p as a premise in inferences or as a basis for action.”¹³³ Suitably reoriented, let’s assume that van Inwagen need not be agnostic about whether there are infinitely many, concrete possible worlds, simply because Lewis’s view is too incredible for him to believe. Still, should he accept that the actual world is the only real world (in some further dispute with Lewis, say, about the status of counterfactuals)? Should van Inwagen get to use his actualism as a premise in further reasoning? Van Inwagen admits that David Lewis has no background knowledge defects, and that Lewis has carefully considered all the reasoning that he, van Inwagen, has done. The question is whether, with these explanations for disagreement ruled out, van Inwagen remains rationally entitled to accept his philosophical positions for use in further arguments.

¹³² Elgin, “Persistent Disagreement,” 60

¹³³ See Cohen, An Essay on Belief and Acceptance. Presumably, accepting p involves treating p as a premise in inference and action, plus believing that p, as well. That is, acceptance is a more demanding notion.
Elgin rightly notes that, even though reframing the question in terms of acceptance versus belief might seem to trade one puzzle for a nearly identical one, this is not the case. Focusing on whether something can be used as a premise in reasoning, rather than on whether something should be believed, denied, or neither, turns our response to epistemic disagreements at least partially into a practical question. Now in some cases refusing to accept one’s beliefs as premises for further reasoning, in the face of stern disagreement from one’s epistemic peers or epistemic superiors, has unacceptable costs. Consider two astronomers who disagree about the correct interpretation of some recently released data. If, in the face of this disagreement, each refused to accept their belief as a premise in further reasoning, then neither would be capable of fleshing out the consequences that would be true, were her favored interpretations correct. This, however, would cripple scientific inquiry, since often the best way of resolving such first order interpretive disputes involves determining their consequences and designing experiments to test whether such consequences hold or occur.

By contrast, the anti-skeptical background shared by most philosophers allows philosophical reasoning to continue. This is so even if one refuses to accept, in Elgin and Cohen’s technical sense, the anti-skeptical premises as reasons for drawing further conclusions in arguments against skeptics. Failing to defeat skepticism has not been, as a matter of empirical fact, any barrier to our continuing to philosophize. This holds for at

---

134 I use ‘epistemic peer’ to echo a general term used in the disagreement literature, about which, however, there is (ironically) substantial disagreement. Some view epistemic peers as agents who have exactly the same total evidence for a given proposition. Others, noting that perhaps no two agents share the same evidence for anything, fall back to a looser notion, according to which one’s epistemic peer is at least as smart, careful, attentive to the arguments, and informed about the empirical facts as oneself is. This latter characterization more closely fits what I have in mind.
least two reasons. First, philosophical disputes can fruitfully proceed conditionally, in many cases. That is, often the descriptivist will be quite happy to establish: if skepticism is false, descriptivism is true.\textsuperscript{135} The causal theorist will resist the conditional conclusion because, of course, both theorists accept a shared commonsense anti-skepticism despite their differences on this question.

These crucial points help moderate the force of the intuition I expressed at the beginning of this section on disagreement. My argument does not require total suspension of judgment on philosophical questions. All I need for my purposes is that disputes about the highly philosophical premises of Spinoza’s anti-skeptical arguments are sufficiently similar to the van Inwagen-Lewis case 4. So, they should be treated conservatively (if one of the disputants is a skeptic), or mutually begged and bracketed, if neither disputant is. This shows why refusing to accept these premises in arguments against the skeptic cannot have unacceptable pragmatic costs for philosophical reasoning more generally. Combined, these considerations hopefully make my argument’s second premise, and thus the dilemma itself, substantially more palatable as a tool for evaluating the anti-skeptical arguments Spinoza’s interpreters attribute to him.

2.5 Unsuccessful Strategy I: Spinoza Directly Confronts the Cartesian Skeptic

Having clarified and defended the dilemma of redundancy or impotence, I now argue that almost all discussions of Spinoza’s response to skepticism founder upon it. Of course, I will not be able to discuss every argument with the detail it undoubtedly

\textsuperscript{135} Of course, it seems doubtful that the descriptivist can establish this, since often far more than the mere assumption of anti-skepticism is required to establish some particular philosophical view.
deserves, but I suspect that after a few examples the general pattern will be easy to spot. In each case, I’ll identify the philosophically loaded premises underlying the reconstructed anti-skeptical argument, those that require conservative treatment. Each time, I’ll argue briefly for this treatment, and then show how the illicit epistemic dependence arises. Willis Doney’s classic treatment of the problem of philosophical skepticism in Spinoza can serve as a promising starting point. Doney considers Spinoza’s responses to two sorts of skeptic, one of whom is basically the Cartesian skeptic as I have defined it. The other sort of skeptic professes doubt, but not because of principled philosophical worries arising within a specific theory of mind. Rather, Spinoza in the TIE (G II/18) considers skeptics who simply say “I know nothing,”—these skeptics might be considered ‘merely verbal’.

This class of merely verbal skeptics raises problems orthogonal to my main concerns here. Nonetheless, Doney’s interpretation of Spinoza’s response to them so vividly illustrates the dilemma that interests me that they serve as useful training examples for recognizing the pattern to follow. On Doney’s reading, Spinoza offers three arguments against “skeptics who persist in professing total doubt and who simply say, ‘I know nothing.’” First, according to Doney, Spinoza gives the following *reductio* argument.

Suppose that someone who utters the words ‘I know nothing’ is making a true assertion. If his assertion is true, he cannot know that he is making this assertion. On the assumption that a person cannot make an assertion without knowing that he is making an assertion, it follows that he is not making an assertion and *a fortiori* that he is not making a true assertion. Hence, someone uttering the words ‘I know nothing’ cannot be making a true assertion...someone uttering...

---

136 Doney, “Spinoza on Philosophical Skepticism,” 632
these words is merely uttering words—he is not asserting that he knows nothing, nor is he expressing doubt.137

The argument’s underlined portion is most relevant for my purposes. There, Doney appears to propose the following principle on Spinoza’s behalf—**Knowledge**: It is possible for S to assert that P only if S knows that she asserts P. **Knowledge** is philosophically contested. While Timothy Williamson suggests that “one must: assert p only if one knows p,” others like Jason Stanley have suggested one should only assert what is epistemically certain.138 Most others have defended substantially weaker norms of assertion, including asserting what it’s reasonable to believe, assert p only if one’s belief that p is evidence responsive, or assert only what is true.139 But even Williamson, who defends a knowledge norm for assertion, does not suggest that violating this norm is impossible, as Doney requires on Spinoza’s behalf here.

All this controversy suggests, based on our earlier reasoning, that **Knowledge** deserves conservative treatment. Plainly, however, an argument with **Knowledge** as a premise and the falsity of skepticism as its conclusion displays IED. For among the conditions that make you have justification for **Knowledge** as a premise—which it needs antecedently, since it should be treated conservatively—is the falsity of ‘I know nothing’ skepticism. For if this skeptical conclusion were true for all speakers, then no one could assert anything, contrary to our numerous experiences of assertion. Even if it were only true for the skeptic at issue, the only way one can evaluate the skeptic’s claim as an

137 Doney, “Spinoza on Philosophical Skepticism,” 632-633
138 Williamson, *Knowledge and its Limits*, 235; Stanley, *Knowledge and Practical Interests*
139 In order: Lackey, “Norms of Assertion”; Maitra and Weatherson, “Assertion, knowledge, and action”; Weiner, “Must We Know What We Say?”
assertion that he knows nothing, in order to defeat it, must presuppose its assertoric character, and thus the falsity of his knowing nothing, as a condition for justifying Knowledge. Otherwise, the skeptic’s own assertion would counterexample the premise directly. Since the requisite justification for Knowledge occurs in the argument’s conclusion, the argument displays IED. It is thus redundant or impotent.

Doney’s next two reconstructions proceed along similar lines. On Spinoza’s behalf, he endorses Transparency: “if an assertion ostensibly about one’s own mental state is true, it must be known to be true,” as well as Verification: “if [the skeptic] is in fact making a theoretical statement, there must be some way in which his statement can be supported or defended.”

Transparency faces devastating criticisms from other philosophical traditions that have never found the view at all plausible, empirical results that appear to cast doubt on it, and powerful vagueness arguments against it. It deserves conservative treatment. But then any argument with Transparency as a premise, against the ‘I know nothing’ brand of skeptic, either presupposes the falsity of ‘I know nothing’ skepticism, or has the absurd consequence that there are no true claims about such a skeptic’s mental states. Clearly, the ‘I know nothing’ skeptic appears to assert something. Transparency requires that if this is true, it must be known, and the justification for Transparency is, apparently, the presupposition of this brand of 140 Doney, “Spinoza on Philosophical Skepticism,” 633-634

141 Garfield, Engaging Buddhism: Why It Matters to Philosophy, makes this point about self-knowledge and the rejection of Transparency ubiquitously in his discussions of the Buddhist tradition. Schwitzgebel, “How Well Do We Know Our Own Conscious Experience?” summarizes some recent empirical work, and Williamson, Knowledge and its Limits, offers a phenomenal sorites case involving sensations of hot and cold as a general argument against a thesis very much like Transparency.
skepticism’s falsity. Otherwise, the skeptic’s assertion constitutes a straightforward counterexample.

**Verification** is similarly flawed. If the ‘I know nothing’ skeptic makes his claim but insists that he has no general philosophical reason for doing so, despite its being a theoretical claim, then how can Spinoza use **Verification** against him? It seems that the falsity of ‘I know nothing’ skepticism (the argument’s conclusion) must be presupposed in order to justify **Verification**, since again, the skeptic’s claim would appear to constitute a straightforward counterexample. Apart from the independent unattractiveness of both principles, therefore, any argument from either **Transparency** or **Verification** to anti-skepticism fails as redundant or impotent.

Perhaps, however, Spinoza’s responses to this sort of skepticism are underwhelming, precisely because the ‘merely verbal’ skeptical position is itself so uninspiring. Unfortunately, as Doney reconstructs them, Spinoza’s arguments against our Cartesian skeptic are similarly structurally defective, albeit more interesting. To reiterate, we are here concerned with a Cartesian skeptic expressing doubts about his clear and distinct ideas in the context of a broadly Cartesian theory of mind, for reasons related to doubts about the origins of his beliefs (conditions a-c from before). The Cartesian skeptic supports these doubts with reasons, with reference to modally nearby evil demon possibilities. And, as we mentioned earlier, such doubts about clear and distinct perceptions plausibly extend to empirical knowledge of the external world, skepticism about reason and judgment, and skepticism about modality.

To combat such skeptics, Spinoza first states conditions any genuine doubt must meet. According to Doney, these two conditions arise from Spinoza’s own theory of
mind. First, Doney thinks “according to Spinoza, [Auto Acceptance] to have or conceive an idea is eo ipso to accept that idea—provided that there is no cause or reason for doubting or rejecting it…This statement is in keeping with Spinoza’s rejection of Descartes’ theory of judgment, which distinguishes having an idea (an act of intellect) and assent to that idea (an act of will).”\(^{142}\) Second, for Spinoza, “[Competition] doubt will arise through another idea, not clear and distinct enough for us to be able to draw any certain conclusion with regard to the matter under consideration (G II/29-30).

**Auto-Acceptance** and **Competition** are plausible and interesting principles, even more so within the context of Spinoza’s sustained and holistic defense of them, or principles like them, in the *Ethics*. Moreover, Doney thinks, they allow Spinoza to mount an interesting defense against the Cartesian skeptic, via the following two arguments.

First, Spinoza can respond to this sort of objection:

> It can also be objected that, when someone attains a clear and distinct perception of his origin and perceives that he is not the victim of a deceiving God, he cannot claim to know that this is so without begging the question at issue; that is, without assuming that at least one clear and distinct perception—the clear and distinct perception leading him to this conclusion—is true.\(^{143}\)

Doney’s reconstruction of Spinoza’s response to the charge of question begging is especially interesting for our purposes, since my charge of redundancy or impotence is closely related.

Here’s how Spinoza sees the dialectic, according to Doney’s reconstruction. The skeptical objector notes that Spinoza’s non-deceptive conception of God is doubtful,

---

\(^{142}\) Doney, “Spinoza on Philosophical Skepticism,” 622-623

\(^{143}\) Doney, “Spinoza on Philosophical Skepticism,” 626
since God might instead have the characteristics of an evil demon. This objection, as
Doney construes it, violates *Competition*. As he puts it, “Letting \( p \) stand for the
proposal that I am the creature of a deceiving God, the skeptic’s argument is of the
form *not*-\( p \) is doubtful because it is possible that \( p \). The object of doubt and the reason for
doubt both contain (though in different senses or ways) the same proposition.”\(^{144}\) This
fails to respect *Competition*, and so fails as a skeptical argument. Without respecting
*Competition*, for Doney, the skeptic fails to express a principled reason for doubt, and so
becomes merely an instance of the ‘I know nothing’ skeptic in another guise.

We might clarify Doney’s thought that the object and reason for doubt both
contain the same proposition as follows. The skeptic argues that (\( P \)) ‘I am the creature of
a non-deceptive God’ is doubtful because, possibly, (\(~P\) `I am the creature of a
deceptive God.`) The object doubted (*I am the creature of a non-deceptive God*) is
doubtful because of a reason (possibly, it is not the case that *I am the creature of a non-
deceptive God*) that simply adds ~ and the possibility-operator to the same proposition.
But *Competition* requires that the reason for doubt be some other idea, not the same idea
with ‘possibly ~’ affixed—the skeptic’s doubt does not respect this restriction.

Doney’s background assumption here seems to be that Descartes’ evil demon
hypothesis is so under-described, so vague and nebulous, that it amounts to the mere
possibility that God might be other than he is claimed to be. Some such assumption
seems necessary, since otherwise there are far more natural ways of characterizing the
object of and reason for doubt. One might say that the object of doubt \( P \) (*I am the
creature of a non-deceptive God*) and the reason for doubt \( Q \) (*My beliefs are given to me

\(^{144}\) Doney, “Spinoza on Philosophical Skepticism,” 626
by an evil demon intent on deceiving me) seem like wholly distinct propositions. If one assumes that this Q is just some under-motivated way of describing (possibly, ~P), then Doney can apply Competition to rule out this skeptical charge of question begging.

These two principles, for Doney, can combine with a further Spinozistic claim to show, in a second argument, that the radical Cartesian hypothesis—that one is the victim of a radical deceiving God—is not just false, but impossible. Doney thinks we should spot Spinoza what he states in 2p44 (Reason): “It is not of the nature of reason to consider things as contingent, but as necessary.” Doney says:

Someone who has a clear and distinct idea of his origin concludes not just that not-p, but that p is impossible [by endorsing Reason, which precludes seeing such truths as contingent, rather than necessary]...Given the very plausible principle that the denial of a proposition cannot serve as a reason for doubting that proposition, it follows that the denial of the proposition ‘p is impossible’ cannot be a reason for doubting that this proposition is true. And it is precisely this principle that is violated when the skeptic questions Spinoza’s conclusion that p is impossible and gives as his reason that he may after all be the creature of a deceiving God, viz, that p is possible.145

Here, it should be stated, we do not have a single principle whose acceptance renders the remainder of the argument superfluous. Rather, Spinoza’s trio of principles (Reason, Auto-Acceptance, Competition) sketch the rough outlines of a theory of mind according to which the reasoning mind considers certain principles as necessary not contingent (Reason). The reasoning mind, merely by considering them, accepts them (Auto-Acceptance). And because necessary truths are clear and distinct, for Spinoza (see especially his comments on ‘common notions’ in 2p38-39), any reason for doubt will inevitably be less clear and distinct than the original ideas, thus lacking the power to

145 Doney, “Spinoza on Philosophical Skepticism,” 626
overcome the first idea (as *Competition* demands).\textsuperscript{146} Even more, *Reason* guarantees that competing ideas are not just false, but impossible. Given this theory of mind, then, the Cartesian skeptic’s reasons for doubt, upon further inspection, are inadequate to the conditions Spinoza establishes for rendering them genuine.

But Spinoza’s argument from *Auto-Acceptance, Competition, and Reason*, to the falsity of Cartesian skepticism, fails. To start, all three principles are clearly highly philosophical, and none enjoys uncontested acceptance. *Reason*, for instance, seems to imply that scientists studying indeterministic physics are using their minds in fundamentally unnatural ways, on the plausible assumption that indeterminism is incompatible with necessity. *Auto-acceptance* seems to render a whole area of contemporary epistemology and action theory senseless, since it precludes suspension of belief and seems incapable of explaining cases where we hesitate about what to believe.\textsuperscript{147}

Contemporary defenders and critics of *Competition* are less obviously identifiable, since Spinoza’s theory of ideas enjoys almost no widespread acceptance now. But it’s not obviously compatible with certain natural descriptions of hesitation in acting in morally difficult situations. Suppose Truman thinks about whether to drop the atomic bomb. After considering various specific pros and cons so many times, he no longer considers them explicitly. Rather, he simply believes: It’s all-things-considered

\textsuperscript{146} See Della Rocca, “The Power of an Idea,” for an extended explication of this process of one idea’s overpowering another.

\textsuperscript{147} For a summary of recent discussions on the ‘ethics of belief’, which appears to presuppose that there is at least some propositional attitude like belief or something relevantly similar that we can control (and hence, plausibly, refrain from), see, Chignell, "The Ethics of Belief.”
best to drop the bomb. Later that day, he doubts his decision, and his reasoning might seem to be: perhaps, it’s not all-things-considered best to drop the bomb. This apparently acceptable reasoning involves one proposition as both the object of and reason for doubt, violating **Competition**. Spinoza, of course, would reject this plausible-sounding story. My point is simply that the philosophical, contested nature of these premises, plus their interdependence (i.e. Spinoza needs at least two of the three to mount an interesting argument), demands their conservative treatment.

If these three principles require antecedent justification, we can ask what the conditions for our justifiably believing them might be. Clearly, the falsity of Cartesian skepticism as defined here must be among them. For to be justified in believing these three principles, we must have some confidence that no one really doubts as the Cartesian skeptic appears to. That is, we must assume that the numerous putative counterexamples to **Auto-Acceptance** fail to be so. Moreover, we seem to require *knowledge* of very many things about the mind (i.e. that it forms beliefs, that it is capable of assigning both contingent and necessary causes to events) in order to be justified in asserting that the mind’s nature is to see things as necessary not contingent. These bits of knowledge about the mind, and these dismissals of putative counterexamples, are just included in the argument’s desired conclusion, however—that we have the knowledge the Cartesian skeptic claims we lack. So any argument built from them displays illicit epistemic dependence, and is thus either redundant or impotent. To be justified in accepting such strong premises, the skeptic would already have to be who he is not—someone who already accepts the anti-skeptical conclusion.
Doney is not alone in reconstructing Spinoza’s arguments so that they are vulnerable to these sorts of criticisms. Martha Bolton defends a non-Cartesian, distinctively Spinozistic response to skepticism. She thinks that a distinctive feature of Spinoza’s metaphysics allows him to argue for thesis T: “There is a way to know (rationally affirm, pending some reason for doubt) p that does not involve recognizing that you know p and concluding p.” This distinctive feature is Spinoza’s metaphysics of ideas. Bolton continues:

To bring out that knowing is epistemically independent from knowing that you know, Spinoza relies on the metaphysics of the idea/object distinction…In order to affirm something of an idea (in particular, that it constitutes knowledge) you must make that idea the object of a second idea…[the first idea] constitutes knowledge if you can affirm [for example] that Peter is a man, whether or not you arrive at the affirmation by deduction from the fact that you know you can affirm it.”

How exactly is this metaphysical point supposed to serve as the fulcrum for a successful anti-skeptical argument? Bolton’s basic thought is that the skeptic has illicitly restricted the range of possible ways of coming to know a given proposition p, for example that ‘two and three are five’. The skeptic has made it seem like knowledge of p depends first on seeing that p is a self-evident proposition and concluding on the basis of the principle that ‘self-evident propositions are true’ that p is true. But if thesis T, along with Spinoza’s broader metaphysics of ideas, gives reason to think that there might be other ways of coming to know this proposition, then the skeptic’s argument is shown to be invalid.

---

148 Bolton, “Spinoza on Cartesian Doubt,” 388

149 Bolton, “Spinoza on Cartesian Doubt,” 389
Bolton’s argument is sophisticated, but there are numerous problems with it. Most importantly for my purposes, thesis T, and its accompanying basis in Spinoza’s idea/object metaphysics, is such that accepting it, along with its corresponding metaphysics, by itself suffices for refuting the Cartesian skeptic. At least, it does so if one licit way of coming to know propositions is simply to ‘immediately and evidently see them,’ as Bolton maintains on Spinoza’s behalf. For thesis T is itself a highly philosophical premise deserving of conservative treatment. Certainly, thesis T has this going for it. The bare statement of thesis T has far greater plausibility than the other philosophical premises we have considered thus far. After all, who would have thought that the only way to come to know propositions was to deduce that they were true from the fact that they were self-evident (given that self-evidence suffices for truth)? That seems like a wildly implausible understanding of knowledge acquisition!

Unfortunately, thesis T doesn’t do all the work by itself. The claim—that alternatives to the deductive account just sketched include immediately perceiving that X is so in a manner sufficient for knowledge—also provides substantial support in Bolton’s argument. Moreover, Bolton clearly means for X to range over mathematical and other more complicated examples, rather than simply over directly perceived external objects,

---

150 For one, Bolton appears to forget that rationally affirming or truly believing that p is true does not constitute knowing that p is true. So even a faculty that directly intuits truths without inference, which Bolton seems to believe Spinoza thinks we have, does not yet help with the skeptical problem. Brueckner, “Semantic Answers to Skepticism,” 141 makes a similar point about many extant anti-skeptical arguments, many of which seem to forget, at crucial stages of the argument, that true or rational belief is often insufficient for knowledge, for any number of good (and non-skeptical) reasons, including but not limited to Gettier cases and mysterious but reliable belief forming faculties.

151 Questions concerning the scope of rational intuitions, skeptical doubts, and their overlap are notoriously difficult. It is by no means always clear what sorts of propositions are supposed to be immune from doubt, or which ones (only a priori truths, only the cogito, etc.) are susceptible to being rationally intuited. See Schaffer, “The Debasing Demon” for a particularly pessimistic take on the scope of propositions susceptible to skeptical doubts.
for example. But this claim about our ability to see and thus come to know such mathematical truths, say, also seems to require antecedent justification.

To be justified in thinking that some subset of one’s self-evident perceptions suffices for mathematical knowledge, one must have already decided that the Cartesian skeptic’s doubts about the truth of his clearest and most distinct perceptions are unfounded. Otherwise, there seems to be no justification for believing in this alternative knowledge acquisition method, of immediate and evident perception, as applied to mathematical statements not even known to be true. Surely, in order to be justified in believing a story about some human capacity for directly intuiting that Goldbach’s Conjecture is true, I must at least be justified in believing that the conjecture is true. But I lack that justification here unless I presuppose its truth, and thus the falsity of Cartesian skepticism. So Bolton’s reconstruction, too, seems either redundant or impotent.

Doney and Bolton well represent one natural way of interpreting Spinoza, according to which he offers direct arguments against the Cartesian skeptic. The draw of this sort of interpretation, naturally, increases significantly if it seems like these arguments are successful. But the class of highly philosophical beliefs used as premises in these anti-skeptical reconstructions simply does not deserve dogmatic status. In the language of our dilemma, these premises do not deserve liberal treatment. They are highly philosophical claims, widely contested even by non-skeptics, which require antecedent justification as precondition for using them in further reasoning. Arguments based on them display a vicious sort of epistemic dependence, since their antecedent justification occurs only by presupposing justification to already believe the arguments’ conclusions.
2.6 Unsuccessful Strategy II: Spinoza Preempts the Skeptic?

Perhaps though, we’d do well to see Spinoza as a more efficient thinker. The second class of arguments I discuss seems to recognize the deficiencies inherent in the first group of interpretations. Instead of focusing on local features of Spinoza’s philosophy of mind or epistemology, which are unlikely to have the requisite justification independently, these interpreters locate Spinoza’s anti-skepticism in highly general features of his metaphysical system. Spinoza dispels skepticism right from the start, arguing from some single, powerful principle. According to this second group of interpretations, Spinoza never even tries to confront the skeptic directly, from mutually acceptable premises about knowledge, justification, or the metaphysics of ideas. Rather, he seeks to preemptively or preventatively undermine the skeptic, shaping the context of debate so that certain questions can’t even coherently arise.

One might note that this general description seems to fit, for example, Doney’s characterization of Spinoza’s strategy (based in Auto-Acceptance, Competition, Reason) as well. This is a fair observation, and admittedly, classing groups of more or less systematic reconstructions of Spinoza’s reasoning is not a neat and tidy business. Still, Doney’s treatment is illustrative as much for the sheer number and variety of implausible, highly philosophical principles that he attributes to Spinoza, as for their content. The attitude he exhibits reveals as much as the particular content of the principles he attributes to Spinoza does. He assumes that Spinoza might plausibly be read as relying on some observation from the philosophy of language about the conditions for assertions of self-knowledge, despite those concerns being wholly foreign to Spinoza’s
philosophy generally. By contrast, this second group of interpreters relies on fundamental tenets of Spinozistic metaphysics.

For now, it’s most crucial to see the arguments discussed in this chapter as lying on a spectrum. At one extreme, Spinoza appears to accept the framework in which the skeptical problem arises and to argue from this neutral starting point that the argument can be defeated. At the other, Spinoza recommends a set of foundational metaphysical principles his interpreters find incompatible with even the statement of the Cartesian skeptical problem. On this reading, one focuses on identifying the theoretical presuppositions of the skeptic’s own position, rather than on combatting his arguments. Accepting the foundational principles makes this direct confrontation unfruitful, since doing so already commits one to having deep knowledge of the world’s structure. In one straightforward sense, these global anti-skeptical arguments, based on foundational metaphysical principles, fare no better than our first group of interpretations. They too fall prey to the dilemma of redundancy or impotence, and their global scope makes them perhaps even more difficult to justify than the more limited principles already considered.

Michael Della Rocca has offered some of the most prominent arguments of this sort. In early work, he concedes that the radical Cartesian skeptic can posit virtually limitless explanations, or even no explanation or cause at all, for why a system of clear and distinct ideas might fail to correspond to reality. He says: “The skeptic need not even assume that there is any cause at all which brings about the failure of the correspondence here, nor need the skeptic assume that the ideas themselves have a cause.”\(^{152}\) Nonetheless, he thinks that Spinoza can combat this radical Cartesian skepticism by

---

\(^{152}\) Della Rocca, “Mental Content and Skepticism in Descartes and Spinoza,” 24
endorsing a plausible view of what *merely having a mind* entails. Della Rocca starts with what Spinoza says, in *TIE* 48:

> If someone proves something to them, they do not know whether it is a proof or not. If they deny, grant, or oppose, they do not know that they deny, grant, or oppose. So they must be regarded as automata, completely lacking a mind.

On this alternative, Spinoza does not start with the premises of Descartes’s evil demon argument, and work from there to combat one or another brand of radical skepticism. Rather, he recognizes that even making a show of ordinary argument evaluation, much less engaging in philosophical discussion, requires that both parties have a grasp of *what it would be* to give a philosophical proof. They must know what it would be to deny a premise or support a claim. That is, any worthwhile discussion occurs between *minded* individuals, and, as Della Rocca puts it, “Spinoza’s [claim] is that, by virtue of what it is to be a mind, each mind grasps truths. There cannot be a mind that fails completely to have mental contents that are true.”\(^{153}\) This bare grasping of some truths cannot be enough, I don’t think—otherwise, Della Rocca has only refuted the reason-less skeptic who claims he knows nothing. However, we might interpret Della Rocca as arguing that the mind’s nature is to have knowledge of some of its own contents, as well as of some basic logical and inferential principles. If these sorts of knowledge are required to give voice to the Cartesian skeptical challenge, then Della Rocca’s argument can be seen as directed at the skeptical position defined here.

At the paper’s conclusion, Della Rocca notes that, despite avoiding Cartesian inconsistency, Spinoza does seem to beg the question against the skeptic in a rather

\(^{153}\) Della Rocca, “Mental Content and Skepticism in Descartes and Spinoza,” 37
straightforward way. If one argues from the claim that the mind’s nature is to have at least some true thoughts, in the way I’ve interpreted this claim above, to anti-skepticism, one’s argument is redundant or impotent in the sense of this chapter. The argument employs a philosophical premise amenable to conservative treatment, one condition for the justification of which is that skepticism be false. Moreover, this claim about the mind’s naturally having true beliefs has been powerfully challenged in the contemporary literature. So while Della Rocca doesn’t put the point precisely this way, he does admit: “This does not mean, however, that Spinoza has refuted or silenced the skeptic.” Presumably, he says this because accepting Spinoza’s preferred view of minds up front is equivalent to declining to engage the most radical skeptical argument.

Diane Steinberg develops a similar global anti-skeptical strategy by examining Spinoza’s response to skepticism in light of his more general account of justification. For Steinberg, Spinoza holds “a coherientist or non-linear account of justification.” She finds the central argument for this view in 2p43, its demonstration, and the accompanying scholium. Here, she reads Spinoza as beginning where Della Rocca’s interpretation leaves off, with the assumption that the human mind has some true ideas:

---

154 There is further slippage too. Della Rocca seems to make the same mistake that Bolton does, and which Brueckner cautions against: for the falsity of skepticism does not follow from the fact that all minds have some true thoughts, but only from the claim that they have some true thoughts known as such.

155 Rosenberg, “Disenchanted Naturalism,” denies that we have beliefs with intentional contents, and a fortiori that we have true beliefs. Patricia Churchland, Neurophilosophy and Paul Churchland, Plato’s Camera, have defended similar claims.

156 Della Rocca, “Mental Content and Skepticism in Descartes and Spinoza,” 39

157 Steinberg, “Knowledge in Spinoza’s Ethics,” 155
1. For the human mind to have a true (adequate) idea [which it has] is for there to be an idea which is adequate in God insofar as he constitutes the human mind (2p11c).

2. Ideas of ideas follow in God in the same way (the same order and connection) as the ideas themselves (2p20, which itself derives from 2p7).

3. The human mind has true (adequate) ideas of its true (adequate) ideas (from 1 and 2).

4. Certainty is having a true (adequate) knowledge of one’s knowledge (definition).\(^\text{158}\)

5. A person who has a true (adequate) idea will be certain of her knowledge.

Steinberg thinks that Spinoza is concerned with normative, rather than merely psychological, certainty, and I will always have this justificatory reading in mind.\(^\text{159}\)

Steinberg’s interpretation of Spinoza’s account of justification owes a deep debt to his metaphysics, in particular to the explanatory barrier between the attributes.\(^\text{160}\)

Steinberg justifies (2) with 2p20, which she says follows from 2p7, Spinoza’s Parallelism doctrine, which stems, as I argued at length in chapter one, from the EB. Premise (1) follows from the fact that the human mind, as Spinoza writes in 2p11c, “is a part of the infinite intellect of God…[God] insofar as he constitutes the essence of the human mind,

\(^{158}\) Steinberg, “Knowledge in Spinoza’s Ethics,” 156-57

\(^{159}\) Steinberg, “Knowledge in Spinoza’s Ethics,” 157; Della Rocca makes the point in “Spinoza and the Metaphysics of Skepticism,” 860, and Perler accepts it also.

\(^{160}\) It’s important to note that both Steinberg and Perler use Spinoza’s metaphysics of ideas in a slightly different way than Bolton did earlier. Whereas Bolton moves upwards (we might say), from ideas to ideas of ideas, Steinberg and Perler move down (or down and up), from the idea to the material object of that idea, to secure their anti-skeptical aims. Both are aspects of Spinoza’s strategy, but it’s helpful to keep them separate.
has this or that idea.” And while the chain of dependence is longer here, we learn in 1p25 that God is the efficient cause of the existence and essence of things. That proposition cites 1p15, which states the total conceptual and causal dependence of all things on God. But Spinoza only manages to argue for this consequence of substance monism, as I suggested in chapter one, thanks to his reliance on the EB. Spinoza’s metaphysics fixes the epistemological options for him, and the EB rears its head in a surprising place much further into the Ethics.

This thoroughgoing coherentist account of justification reveals deep interconnections between foundational metaphysical knowledge and ordinary truth claims. As Steinberg observes: “We have an adequate idea or knowledge of a thing when we understand it as following from its ultimate cause, or when we have a complete explanation of why it exists and has the properties it has.” Adequate ideas, as Steinberg rightly notes, result from our knowing their efficient causes, and that such causal explanations must be ultimate or complete seems dictated by the PSR. So a certain account of explanation plus the PSR’s guaranteeing the existence of all explanations for any phenomenon gives each adequate idea a radical interconnectedness with all other such ideas.

Steinberg uses this interpretation to make sense of the puzzle we are concerned with here: why exactly does Spinoza’s engagement with skepticism seem cursory at best and question begging at worst? Steinberg mentions Spinoza’s discussion of a sort of skepticism in the TIE, but ultimately, she thinks, Spinoza’s engagement with skepticism is simply radically opposed to that of Descartes. She writes:

161 Steinberg, “Knowledge in Spinoza’s Ethics,” 158
According to Descartes, the way to answer the skeptic is to begin by doubting everything that can be doubted, and then build the metaphysical system from the foundation of self-evident, hence self-justifying simple truths that even the skeptic cannot doubt…[But] that is not Spinoza’s way. We know our ideas are true because we have adequate knowledge of them; and adequate knowledge of an adequate idea involves (adequate) knowledge of the entire basic metaphysical system. No single idea is justifiably prior to any other, and there is no way into the system from self-justifying simply truths. For Spinoza, the skeptic is refuted at the end, not the beginning, of the knowledge enterprise.\textsuperscript{162}

Tracing the relations between Spinoza’s metaphysics and his epistemology shows how Steinberg’s anti-skeptical solution runs comically afoul of the dilemma of redundancy or impotence. Steinberg uses Spinoza’s coherentism to give an argument against the skeptic. But the argument for coherentism relies on Parallelism and Substance Monism, each of which ultimately depends, as we saw in chapter one, on the explanatory barrier for its justification (among other controversial premises). But the explanatory barrier is a highly philosophical premise, widely rejected by physicalists and dualists alike, which demands conservative treatment.

I argue in chapter one that no antecedent justification for the explanatory barrier has been proposed in the literature, and certainly Steinberg does not provide one.\textsuperscript{163} Moreover, as Steinberg makes clear, Spinoza’s substance monism requires both 2p3—“In God there is necessarily an idea, both of his essence and of everything that necessarily follows from his essence,”—and 2p20—“There is also in God an idea, or knowledge, of the human Mind.” So if the EB does entail substance monism, then justification for the

\textsuperscript{162} Steinberg, “Knowledge in Spinoza’s Ethics,” 160

\textsuperscript{163} Steinberg does argue that 1ax4 requires that every conceptual relation be a causal relation (“Knowledge in Spinoza’s Ethics,” 140). Thus, in combination with 1p10’s claim that each attribute must be conceived through itself, she does give some explanation for why causal and conceptual isolation go together. But she makes no effort to justify the overarching thesis.
falsity of Cartesian skepticism seems necessary to justify the EB and its myriad metaphysical consequences, many of which presuppose that we already have a great deal of knowledge. So Steinberg’s argument, which employs one of these myriad consequences—coherentism—to argue against the Cartesian skeptic, seems either redundant or impotent. It comes too late in the justificatory game to do any justificatory work.

Dominic Perler, like Steinberg, believes that Spinoza’s coherentist account of justification plays an important role in his response to skepticism. But he understands Spinoza’s foundational principles and overall strategy slightly differently, thus presenting an interesting contrast with the strategies surveyed thus far. For Perler, the skeptic’s argument reveals a tacit commitment to three claims—semantic atomism, dualism, and anti-naturalism—that Spinoza’s own system highlights by its opposition to them. Spinozism, by contrast, begins by assuming semantic holism, anti-dualism, and naturalism, and its development reveals the oddity or unnaturalness of the skeptic’s presuppositions.164

 Appropriately, Perler sees Spinoza departing from Descartes, as did Bolton and Steinberg, in his understanding of ideas. Like Descartes, Spinoza understands ideas as mental tokens—acts of thinking, rather than abstract objects like propositions—with contents, or objective being. But unlike Descartes, Spinoza focuses immediately on what 2p7 calls the order and connection of ideas. Spinoza never takes note of a lone idea and excavates its contents. Rather, Perler says, “there is always an order and connection of

---

164 Spinoza thus gives what Perler calls, following Michael Williams, a theoretical diagnosis of the skeptical problem, which “concedes the seriousness of skeptical problems while questioning their naturalness” (“Spinoza on Skepticism,” 3, online).
ideas’ (E2p7), just as there is an order of material things, and Spinoza spells this connection out in causal terms. If one intends to explain the content of an idea in a specific mind, one needs to examine how it is causally connected to other ideas. Here, Perler closely mirrors Steinberg’s interpretation of Spinoza’s coherentism. But he makes more explicit, rightly in my view, that this coherentism is wholly epistemic. Perler attributes semantic holism to Spinoza while declining to label him a coherentist about truth. An idea’s coherence within the larger system of causally interconnected ideas provides a criterion of its truth, but does not constitute it. Truth’s nature consists in an adequate idea’s agreement with its object.

This observation about truth serves nicely to introduce a further claim, which Perler shares with Steinberg but helpfully makes more explicit. Spinoza’s Parallelism doctrine, interpreted crucially in both its conceptual and causal forms, guarantees an idea’s correspondence with some bodily state. Spinoza thus reserves some truth to all ideas. Perler uses Spinoza’s example of the idea of a horse to make his point clear.

So when a horse is running by, it affects my visual system and causes visual and other sensory inputs, which give rise to a brain state. And this item cannot be considered only under the attribute of extension but also under the attribute of thinking. Taken under this second attribute, it is nothing but a first idea of a horse. Therefore, whenever a brain state is caused, an idea shows up—no additional causation is required. The decisive point is that I do not get this first idea of a horse simply by introspection or by means of an actualization of an inborn horse idea, but only by being in empirical contact with a horse…And whenever I have

---

165 Perler, “Spinoza on Skepticism,” 4, online

166 Perler and others, including Curley, “Spinoza on Truth,” have discussed whether this notion of ‘some truth’ is coherent, but I will not do so here. Suffice to say that I think there are nearby properties more amenable to gradation than truth is—these properties may even be more important, ultimately, but I think we should leave truth binary.
such affections, I also have an idea or even many ideas because of the token identity of modes under the attributes of extension and thinking.\textsuperscript{167}

Perler states Spinoza’s underlying justification admirably clearly, and he forthrightly notes that Spinoza’s complete view, as he states in 2p16c2, is “[that] the ideas which we have of external bodies indicate the condition of our own body more than the nature of external bodies.” That is, Perler does not rule out a great deal of confusion in my idea of a horse, and he acknowledges the common experience of making a great many invalid inferences from one’s horse perceptions. (Inferring that whales are fish might have been an even better example from the animal kingdom). Nonetheless, “no matter how poor and confused the first sensory idea may be, no matter how inadequate it is, it is linked to a horse in the material world. If there were no horse in the chain of physical causes to start with, there would be no corporeal affection of it and, consequently, no idea in the chain of mental causes that stands in a relation of token identity to this affection.”\textsuperscript{168}

Perler, like Della Rocca, addresses the natural skeptical objection. The accusation, of course, is that Spinoza has simply begged the question by assuming that any brain state’s being caused automatically gives rise to a corresponding idea, which is caused by something outside the person whose brain is in that state. For this already precludes global error: “It cannot be the case that all our ideas completely miss their target and do

\textsuperscript{167} Perler, “Spinoza on Skepticism,” 9-10, online; notice that to respect the EB, Perler claims that the first idea of a horse only arises within the causal chain of ideas exactly when, in the chain of physical causes, my body and brain make empirical contact with an actual horse.

\textsuperscript{168} Perler, “Spinoza on Skepticism,” 10, online
not represent anything in the material world.”169 And we can correct local errors, Perler thinks, by calmly proceeding along broadly coherentist lines, justifying our ideas with reference to their nice conformity within our total ideational system.

For Perler, naturalism provides the key that unlocks the Spinozistic puzzle and answers the skeptic’s countercharge. Perler cites Spinoza’s preface to Part III of the Ethics, where Spinoza claims that the laws and rules of Nature are always and everywhere the same. He argues that for Spinoza, the skeptic’s attempt to see human beings as outside the single interconnected natural world makes no sense. “If one intends to explain things and events in the world, one should always speak about one world and one set of laws ruling all things.”170 If one starts with this framework, pausing to wonder whether human ideas, clear and distinct or not, have any direct purchase on reality seems insane. Human beings share the one world they occupy with other material things that causally affect them. Of course, one might think that Perler’s argument moves too quickly here, for surely there are interesting skeptical possibilities that do not require that I have no actual contact with an external world at all.171 I might have an internally coherent but systematically false set of ideas, all caused by bodily things outside me that nonetheless have nothing like the natures or properties I take them to have.

The more important problem is that naturalism, semantic holism, and antidualism, the foundational metaphysical principles Perler identifies as operating at the base of Spinoza’s system, are highly philosophical principles. All of them are widely

169 Perler, “Spinoza on Skepticism,” 11, online

170 Perler, “Spinoza on Skepticism,” 13, online

171 Brain-in-the-vat skepticism, for instance, requires an external world that houses the vats to generate its problem.
contested, and each individually, as well as the trio collectively, deserves conservative

treatment. Perler recognizes something like this problem, and it might initially seem like
he thinks that the requisite antecedent justification for these principles can be provided.

Why should [the skeptic] accept [Spinoza’s] framework...First, [Spinoza] could
point out that he gave detailed arguments for his theoretical assumptions, which
are not just ad hoc claims in a skeptical debate. They are embedded in an entire
metaphysical system. Any reader who has studied the first two parts of the Ethics
should be familiar with this system and with the arguments adduced for, say, the
claim that ideas are not detachable from bodily states.172

It quickly becomes clear that this is not the case, since the ‘detailed arguments for
his theoretical assumptions’ that Perler mentions themselves depend, as did Steinberg’s
reconstructions, on the explanatory barrier (among other Spinozistic principles).

Perler’s actual view seems to be that Spinoza is in a sort of default dialectical
position. He thinks Spinoza’s theoretical assumptions are prior to any specifically anti-
skeptical moves, and thus do not require more than defense, as opposed to ground-up
justification. He concludes by arguing that ultimately Spinoza’s system is an intricate
machine, justified by its prodigious explanatory power. Spinoza can only offer it to the
skeptic to try out, like a particularly cool toy. The assumption that a certain radical
Cartesian skepticism is false and that Spinozistic naturalism is true allows for the
construction of powerful metaphysical machinery. But this assumption vitiates the need
to then train these starting assumptions on the skeptic, since the falsity of skepticism was
presupposed in order to justify the Spinozistic metaphysics in the first place. While Perler

172 Perler, “Spinoza on Skepticism,” 17, online
might teach us something about skepticism, his argument against the skeptic is either redundant or impotent.173

Della Rocca’s later work exemplifies a final instance of this preventative strategy. He too appeals to Spinoza’s metaphysics to craft an anti-skeptical argument. The fulcrum of the debate becomes the skeptic’s brute assertion of the possibility that even our most consistent system of ideas fails to match reality, without any well-defined reason or cause for doing so. Why, exactly, should we spot the skeptic this much? In a fascinating set of interrelated papers, Della Rocca argues that the principle of sufficient reason (PSR) suffices for defeating the radical Cartesian skeptic.174

Without evaluating the argument in detail, we can see how this will go. The PSR is a highly philosophical premise that deserves conservative treatment. Philosophers have offered powerful arguments against the PSR, alleging that it entails necessitarianism, and that necessitarianism is false.175 Moreover, the PSR, which states that every fact has an explanation, is crucially ambiguous. On one reading, it seems to have no epistemic implications at all.176 This is not how it is typically used, however. Often, the PSR plus

173 Bernard Williams, *Ethics and the Limits of Philosophy*, suggests that the real question is not what we should say to the amoralist (who shares no common moral commitments), but what we should say about him. Perler’s solution for skepticism looks quite similar structurally.


175 See van Inwagen, *An Essay on Free Will*, 202-204 for this argument against the PSR, as well as Bennett, *A Study of Spinoza’s Ethics*, 115. Many others exist in the literature. See Levey, “The Paradox of Sufficient Reason,” for an interesting defense against the conjunction argument Bennett and van Inwagen endorse.

176 I discuss this interpretation more below.
the fact that we do not know or cannot think of any possible explanation for some phenomenon is used to argue that the phenomenon does not obtain.\footnote{Della Rocca, “PSR” gives many examples of this sort of thing, including arguments surrounding brute dispositions, fission cases in the personal identity literature, etc.}

This only makes sense as an argument strategy if both parties presuppose that we have some cognitive access to the explanations whose existence the PSR guarantees. Our \textit{presumed knowledge} of the guaranteed explanation in many cases makes its absence in the circumstances of the discussion problematic. But then our justification for believing the epistemic interpretation of the PSR hinges on our having prior justification for skepticism’s falsity. We must be antecedently confident about our knowledge possession in a wide variety of explanatory contexts, in order to employ such a sweeping premise. We thus cannot be justified in doubting the truth of all of our most clear and distinct ideas, since these clear and distinct ideas are the shared basis from which we can then evaluate whether a PSR-based argument succeeds. So an argument from the epistemic version of the PSR to the falsity of Cartesian skepticism already presupposes some justification for believing the conclusion, as a condition for employing the PSR as a premise. Della Rocca’s argument is thus redundant or impotent.

\section*{2.7 Lessons from the Preventative Strategies}

Preventative strategies face the same dilemma of redundancy or impotence that affects more limited reconstructions of Spinoza’s anti-skeptical arguments. Nevertheless, I think they represent a significant advance over arguments employing more limited anti-
skeptical principles, and I want to conclude by drawing four lessons from the preventative framework.\textsuperscript{178}

The first lesson concerns the structure of Della Rocca’s later, PSR-centric argument against the skeptic, which proceeds first by showing how the PSR defeats skepticism and then argues for the truth of the PSR directly.\textsuperscript{179} I actually think both stages are unsuccessful. But the latter stage does teach the lesson that there is something puzzling about standard philosophical practice, which endorses widespread demand for explanation, but only inconsistently. Obviously, most philosophers are far more sympathetic to their own failures of explanations than to those of other—this is a psychological truism. The more interesting way of stating the lesson is this: it isn’t clear whether the PSR is a philosophical principle deserving conservative treatment, or a liberal principle that we must accept as a precondition for reasoning generally. This suggestion comes out in Della Rocca’s attempt to use the PSR to defeat the skeptic.

His argument there also has two stages. The first stage tries to show that the skeptic accepts a primitive bifurcation between two features of ideas, their internal representational credentials and their truth. The second stage recognizes the need for an argument for the PSR itself, which would, if successful, rule out this primitive bifurcation and thus undermine the skeptic’s arguments.\textsuperscript{180} That is, he treats the PSR as a

\textsuperscript{178} Limited principles considered here include Knowledge, Transparency, Verification, Auto-Acceptance, Competition, Thesis T. More generally, I find anti-skeptical strategies that employ principles specific to human thought or language similarly defective. In my book, either commonsense anti-skepticism or global anti-skepticism is the way to go.

\textsuperscript{179} Stage one occurs in Della Rocca, “Spinoza and the Metaphysics of Skepticism,” while stage two is found in Della Rocca, “PSR.”

\textsuperscript{180} Della Rocca argues more directly for the PSR, on Spinoza’s behalf, by applying what he calls the two-fold use of the PSR to the question of existence. Here, Spinoza wonders, according to Della Rocca,
philosophical principle that initially deserves conservative treatment. Most interestingly, however, it becomes clear that Della Rocca’s ultimate aim is actually to argue that the PSR deserves liberal, rather than conservative, treatment. He is arguing that the PSR should share the status that commonsense occupies in this chapter. Unlike most of the other principles considered here, accepting the PSR is a precondition for engaging in reasoning at all, according to Della Rocca.

Della Rocca’s clearest statement of this argument for the PSR proceeds mostly ad hominem and is too complex to do full justice to here. Others have ably criticized it elsewhere. Its basic structure consists in noting the widespread use of explicability arguments across philosophy. Explicability arguments, simply put, are arguments where “a certain state of affairs is said not to obtain simply because its obtaining would be inexplicable, a so-called brute fact.” Examples of such arguments include the case of brute dispositions, fission cases in the personal identity literature, rejection of Aristotelian substantial forms, and requests for explanation of causality and modality to name a few. Since many philosophers grant that explicability arguments have considerable power, they can be backed, awkwardly, into the following corner. Given their what existence is, and finds that “to exist is simply for that thing to be intelligible or conceivable.” The detailed textual case for attributing this line of reasoning to Spinoza emerges in Della Rocca’s “Spinoza and the Principle of Sufficient Reason: A Rationalist Manifesto,” but the short version is this. God’s essence is identical to his existence (1p20), and God’s essence (and thus existence) is to be self-conceived or self-caused (1d3). But if God and his modes ‘play by the same rules,’ as Spinoza’s thoroughgoing naturalism demands, then “God must called the cause of all things in the same sense in which he is called the cause of himself” (1p25). So the existence of any finite mode essentially depends on God’s conferring his own intelligible existence on that mode. For God and finite modes, to be is to be intelligible.

See, in particular, Boehm, Kant’s Critique of Spinoza, for a detailed dissection of the following argument although Melamed, “The Sirens of Elea,” and elsewhere has also provided trenchant criticism along the lines that I favor.

Della Rocca, “PSR,” 2

The relevant examples are discussed in Della Rocca, “PSR,” 2-6.
employment of explicability arguments against those defending brute causation, brute modality, or brute ascriptions of personal identity, they face considerable pressure to endorse similar non-brutality about the question of existence. But to reject brute existence is just to accept the PSR—no thing’s existence is a brute fact means that nothing exists without an explanation. This is Della Rocca’s direct argument.

But there is a second-order problem here too, for if the explicability argument in the case of existence must be rejected, it must, presumably, be rejected for some reason. And this demand for explanation is actually harder to meet, since most philosophers have no plans to cease and desist wielding explicability arguments against their opponents when it suits their purposes to do so. But that policy, Della Rocca insists, is hopelessly ad hoc, and can be rejected as untenable for that reason. Since no line can be drawn between classes of such arguments, one is forced either to accept the full-blown PSR or reject all explicability arguments, a nuclear option Della Rocca thinks most will want to resist.

Bracketing other problems, this seems to beg the question against the PSR’s opponent.\[184\] After all, absent a sound argument for the truth of the PSR, we might expect that the divide between licit and illicit explicability arguments would be brute. But that

\[184\] In addition, philosophers have criticized Della Rocca for invalidly referring, from a claim about how we are forced to reason, to a claim about what is true of the world. They have also pointed to the negative consequences of accepting PSR for our explanatory and ethical projects. Boehm, *Kant’s Critique of Spinoza*, 181-186, writes: “Rationalists, speculative or critical, are committed to the view that one ought not claim to know that a given metaphysical position is true unless one has sufficient grounds for making that claim…Strictly speaking, Kant, who does not understand himself as a non-rationalist but as a critical rationalist, embraces (2) [of Della’s Rocca’s options: that no explicability arguments are legitimate]…Kant would argue, first, that no explicability argument is known to be true—call this 2*. The reason behind this (very costly) claim is that in order for an explicability argument to be known to be true, existence itself must be known to be explicable. And in order for us to believe that existence is explicable, we need to believe the ontological argument, etc. However, the Kantian will at the same time accept that we are rationally committed to striving to explain everything, to eliminate all brute facts.” Later in that section, he points out some of the costs of accepting PSR for our moral and explanatory projects. Suffice to say that Boehm appreciates, even in the context of denying that Della Rocca’s argument is successful, the potential (skeptical) costs of arguing that no explicability argument is legitimate.
does not mean that there is no puzzle here. Surely some demands for explanation are legitimate, and neither philosophy nor science progresses much if appeals to brute relations, powers, or influences are allowed to go unchallenged. In fact, the opposite seems most accurate. A philosopher who criticizes an opponent’s account of causation for being brute but then denies that his account of modality faces the same criticism seems to be doing something wrong. Questions concerning the scope of appropriate application of explicability arguments seem to be legitimate ones. So one can block Della Rocca’s direct argument for the PSR by expanding the dilemma to include three choices: accept the full-blown PSR, reject all explicability arguments, or grant that the line between legitimate and illegitimate explicability arguments is brute and unknown (and potentially unknowable!).

But pending some principled explanation for why some explicability arguments rather than others hold water, there is a distinctly skeptical flavor to this way of avoiding Della Rocca’s argument. And it’s worth remembering that question begging arguments, like Della Rocca’s, can still be sound. So while the skeptic should have little to fear from the argument for the PSR, non-skeptical philosophers dissatisfied with the skeptical tenor of the mechanism that blocks it most naturally face additional pressure to develop a competing response. Surprisingly, Spinozistic arguments for the PSR, often interpreted as anti-skeptical, most immediately challenge non-skeptical philosophers uncomfortable with brute facts in their opponents’ positions but incapable of eradicating them from their own. Such philosophers must decide why they believe that the PSR should be treated conservatively, thus requiring a successful argument that Della Rocca tries and fails to provide, rather than being treated liberally. They must examine why they believe other
philosophical principles are more deserving of this liberal treatment, and must consider whether these other principles tacitly presuppose the truth of the PSR as well.

The second lesson emerges from reflection on similar tensions arising within Perler’s naturalism. The issue is that Perler simply assumes that the radical skeptic will not be a naturalist. Since Descartes, the original purveyor of such skepticism, has been interpreted as an anti-naturalist, it’s understandable that Perler shares this interpretation and attributes this assumption to Spinoza as well. He says:

Whatever [humans] think, whatever ideas show up in their minds, happens because they are in contact with other things. They are permanently exposed to external influences and therefore also to a change in their physical states that inevitably goes along with a change in their thoughts. Anyone who denies this has to explain how there can be a causal order for things in a material world and another, completely different causal order for human thoughts. It does not help to appeal to a special order as long as one cannot make it intelligible, explaining its relation to the causal order governing the rest of the world. If one fails to provide this explanation, one simply introduces an unexplained explainer: something that is merely postulated in order to grant a special place to human thoughts but that cannot be understood.

The underlined sections of Perler’s discussion provide evidence that he has Descartes in mind—he sees Cartesian skepticism as arising from Descartes’s insistence on preserving a special place and brand of causal activity for human thoughts. Actually though, problems similar to those of the Cartesian skeptic arise within a metaphysical system very much like Spinoza’s. They do so for reasons wholly independent of any

185 There is a sense in which reflection on modern science engenders fairly radical skepticism in some of its practitioners. For example, Williamson, “Philosophical Intuitions and Skepticism about Judgment,” 150 argues: “That the skeptic about perception wears the traditional garb of the philosopher while the skeptic about judgment dresses up in a scientist’s white lab coat should not blind us to the underlying structural similarity of their arguments.” And scientific arguments for skepticism are widely discussed in the contemporary literature. See, for example, Frances, Skepticism Comes Alive.

186 Perler, “Spinoza on Skepticism,” 13, online
possibly misguided attempt to preserve human mental activity as unique. A second lesson to keep in mind is that skeptical consequences develop predictably from basic tenets of certain naturalist projects.\textsuperscript{187}

Further reflection on naturalism and the PSR suggests that the critical discussion of Spinoza’s anti-skepticism is fundamentally misguided. Interpreters have failed to recognize a third lesson; namely, naturalism and the PSR are deeply compatible with a radically skeptical position on human knowledge. Recall Spinoza’s version of the PSR in 1p11d: “for each thing there must be assigned a cause, or reason, as much for its existence as for its nonexistence.” Contemporary versions often state the principle in some related way: “there are no brute facts, for every fact F, there is an explanation for why F is the case.”\textsuperscript{188} Spinoza’s naturalism is just the view, expressed in the Preface to Part III, that the “laws and rules of nature, according to which all things happen, are always and everywhere the same” (G II/138, 14-16).

Compatibility among naturalism, the PSR, and radical skepticism manifests on several levels. Most superficially, we may note that there is no logical incompatibility between each fact F’s having an explanation and our not knowing, or justifiably believing, for any F that explanation E is its cause or reason. The principle states an existence claim about explanations, but formally lacks any epistemic import at all. Naturalism makes a similar existence claim about a single set of laws, again without any

\textsuperscript{187} Plantinga, Where the Conflict Really Lies and Rea, World Without Design: The Ontological Consequences of Naturalism, as well as many others, make versions of this point that I find quite persuasive at least generally.

\textsuperscript{188} Lin and Melamed, “The Principle of Sufficient Reason,” sec. 1
direct implications for epistemology. Even more, the skeptic may be rationally committed to the PSR while preserving her universal doubt. Finally, as I’ll show next chapter, the skeptic can tell a detailed, coherent story, compatible with thoroughgoing naturalism as well as formal statements of the PSR, that seriously impairs our ability to provide the requisite explanations whose existence the PSR guarantees.

The fourth and final lesson draws out the implications of the preceding three. Any brand of skepticism with doubt as its central feature, including the more specific Cartesian variety characterized earlier (with conditions a-c), evades anti-skeptical arguments from naturalism or the PSR. Perler seems to recognize this implicitly, when he observes that Spinoza doesn’t adopt naturalism solely in order to respond to the radical Cartesian skeptic. Spinoza finds naturalism independently plausible, uses it to explain a wide variety of phenomena, and its success in offering such explanations then serves as some justification for its invocation against the skeptic too. But since skepticism is deeply consistent with both naturalism and the PSR, the latter two cannot be used to

189 As I’ll discuss more in chapter 3, the implications of Spinoza’s naturalism, in particular, are further complicated by the fact that the single set of natural laws must hold across all attributes, even though humans only have knowledge of two of infinitely many of them.

190 At least, that is, if the PSR appears true to her. Remember, the ancient skeptics are, as Sextus Empiricus famously puts it, “still investigating”—the skeptical practitioner’s task of seeking alternative explanations to oppose to common beliefs seems guided by some underlying commitment to such causes’ existence. I cannot defend the view here, but I think that most of the Cartesian skeptical conditions (a)-(c) outlined here can be fruitfully incorporated into a Pyrrhonian skeptical view. See Sextus Empiricus, Outlines, I.3; Barnes, “The Beliefs of a Pyrrhonist,” gives a representative reading of the way in which the skeptic can retain certain beliefs in the context of universal doubt about theoretical questions.

191 Actually, I’ll introduce a further twist, since Spinoza, rather than any skeptic, will be the one telling this detailed, coherent story, and drawing out its consequences. These consequences might be seen as skeptical, but in any case, they seriously call into question our ability to provide certain sorts of explanations. Spinoza will be developing this detailed story to provide motivation for introducing the EB, but I’ll say more about that in chapter three.

192 Perler, “Spinoza on Skepticism,” 17, online speaks of the overall, as opposed to the specifically anti-skeptical, comprehensiveness and explanatory power of Spinoza’s metaphysical system.
respond to skepticism. Certainly, they can’t be used to do so without adding auxiliary epistemological premises that end up making the argument viciously epistemically dependent.

I think Spinoza recognizes this, which suggests to me that Spinoza’s interpreters have misidentified the target of Spinoza’s arguments from naturalism and the PSR. The dilemma of redundancy or impotence gets going because Spinoza and the skeptic share too little philosophical common ground for the anti-skeptical arguments to avoid vicious epistemic dependence. But against a shared background of anti-skepticism, such as Spinoza shares with Cartesian dualists as well as contemporary physicalists, Spinoza can use the PSR, naturalism, plus the threat of being unable to explain their anti-skepticism, given their ontology, to drive his opponents towards his position. In the next chapter, I will argue that Spinoza identifies fundamental explanatory defects in competing dualist or (certain forms of) physicalist metaphysics. Spinoza then exploits the natural philosophical discomfort with these explanatory failures, against a backdrop of shared anti-skepticism, to argue for introducing the explanatory barrier between the attributes. The most interesting and philosophically charitable interpretation thus identifies interactionist dualists and non-reductive physicalists, rather than the skeptic, as Spinoza’s primary targets. Spinoza tries to show that the EB must supplement PSR-based naturalism in order to secure the twin explanatory goods of the philosophical naturalist and natural scientific projects.
CHAPTER 3:
A READING OF SPINOZA’S SHORT TREATISE AS AN ARGUMENT FOR THE EB

3.1 The Shape of the Problem

The explanatory barrier between the attributes serves as Spinoza’s response to his deep puzzlement with the Cartesian interactionist system. In the first chapter, I argued that the EB plays a fundamental role in deriving most of Spinoza’s central theses in the first books of the Ethics, as well as in resolving longstanding problems within his metaphysics. Regrettably, many of Spinoza’s interpreters have failed to locate any compelling justification for introducing it. Then, I suggested without explicit justification that this explanatory problem might be remedied by reading the EB as the best response to a new sort of skeptical or explanatory problem. The second chapter then sought to show that Spinoza hasn’t solved the problem of specifically Cartesian skepticism by other means. I claimed that if Spinoza intends to engage the Cartesian skeptic directly, as most commentators have thought, then his arguments aren’t very good. Moreover, recent attempts to fashion Spinoza’s strategy along broadly preventative lines, as a way of

193 In the remainder of what follows, I’ll often refer to ‘interactionism’, by which I have in mind primarily interactionist forms of dualism, of the broadly Cartesian variety. Some of what I say in chapter 5 touches on epiphenomenalist or emergentist forms of dualism, but interactionist varieties, rather than these forms, are the primary target of what follows, naturally because of their rejection of the EB. However, note that even idealist or physicalist systems that want to allow for 1) distinct mental and physical properties and 2) some mental-physical causation qua mental to physical properties, or mental-physical cross-attribute explanations, face a version of the argument I develop here on Spinoza’s behalf.
precluding skepticism’s arising, are similarly ineffective if directed as arguments against the skeptic. Both sorts of arguments face a dilemma of redundancy or impotence.

Still, this preventative interpretation represents a step in the right direction. The key, I think, is to sharpen our appreciation of the target of Spinoza’s reasoning, to turn our attention away from the skeptic and towards rival systematic philosophies. These targeted systems include Cartesian interactionism, of course, but also other naturalistic versions of physicalism that lack the explanatory barrier. Shared anti-skeptical commitments unify these otherwise disparate metaphysical positions, whether interactionist dualist or physicalist, and this gives Spinoza a crucial argumentative opening. Last chapter I concluded that naturalism, guided by the principle of sufficient reason, could not secure Spinoza’s anti-skeptical aims, precisely because naturalism and the PSR are compatible with many skeptical scenarios. But, once we no longer interpret Spinoza as in dialogue with the skeptic herself, we can see Spinoza’s strategy take shape.

His crucial argumentative opening takes the form of an explanatory challenge. With skepticism presumed false, as it is for Descartes and for most physicalists, Spinoza argues that his system, built around the PSR, naturalism, and the crucial explanatory barrier, can much better explain why anti-skepticism is true, while other systems cannot. More particularly, I will argue that Spinoza’s isolationism provides a compelling response to what I call the problem of pathological causation, which rival interactionist systems can explain only very clumsily, or not at all. This strategy might seem

194 ‘Spinozistic Isolationism’ names the conjunction of the EB and Spinoza’s other central metaphysical theses. It is at least epistemically possible to have a non-Spinozistic isolationist system, since I haven’t defended Spinoza’s argument for substance monism fully in chapter 1, though the only ready examples of such systems that come to mind are versions of Leibnizianism, which, notoriously, seems rather precariously placed to deny many of its apparently Spinozistic implications.
excessively risky, for it might suggest that, if any serious skeptical problems remain on the Spinozistic Isolationist account, then no explanatory advance has been made. Further, it might be urged, whatever one’s metaphysical system, skeptical problems can always arise, and so no argument of the form I propose could possibly be successful.

This objection is fundamentally misguided, confusing as it does two versions of the way that a skeptical or explanatory challenge might be developed and used. Tim O’Connor explains the confusion well in his discussion of a common response to criticisms of Humean modality:

A better reply claims the problem [of induction] is insoluble whatever one’s metaphysics of modality, and so it cannot be wielded against one who would reject necessities, whether causal or logical (Lewis 1986: 117; Loewer 1996: 101-27). But I think that this reply rests on a misunderstanding of the problem, or, if one likes, it conflates two problems of induction, one insoluble but resulting from unreasonably high standards of justification, and the other quite sensible and insoluble only for the impoverished metaphysician. The unreasonable problem, posed by the skeptic, is to give non-question-begging proof or evidence that proper inductions will (mostly) pan out in our world…The other ‘problem’ allows that we may reasonably assume induction to be a fundamentally reliable practice, but it goes on to ask what sort of metaphysical conception of the world (as regards its basic dynamics) meshes with—makes sense of—that epistemological assumption.195

John Heil makes an analogous point in a sphere more directly relevant to the one we are concerned with here, in an introduction to the problem of mental causation. He says:

The question I shall address, however, is not whether mental goings-on cause, or are ‘causally relevant’ to, physical goings-on, but how this could be possible, how it might work. In answering this question, it is not the point to call up examples of accepted causal explanations in psychology or neuroscience as philosophers of science have sometimes done (e.g. Woodward, 2008). Princess Elizabeth does not doubt that minds and bodies interact. Her challenge to Descartes is to account for

195 O’Connor, *Theism and Ultimate Explanation*, 14
mind-body interaction given Descartes’ commitment to dualism. Nowadays the
difficulty is, more often than not, taken to be that of accounting for mental
causation, or the ‘causal relevance’ of the mental, given the truth of non-reductive
physicalism.\textsuperscript{196}

Both Heil and O’Connor make perfectly clear that interesting problems remain
even after one has left the skeptic behind. So too, in the present chapter, I do not develop
a problem of pathological causation that I regard as a serious skeptical challenge capable
of undermining our ordinary justification for causal \textit{claims independent of our other
metaphysical commitments}. But given shared anti-skepticism about what I’ll call the
problem of pathological causation, the dualist or physicalist’s inability to explain how
such anti-skepticism is possible, either in principle or given the laws of the actual world
and her other metaphysical commitments, undermines her justification for believing it,
given those commitments. Spinozism arises from the shaky epistemic foundations of rival
systems, illuminating metaphysical commitments we might not be aware of having.
Spinozism explains something important—why causation is not pathological—that the
interactionist cannot explain well, if at all.

Spinoza obliquely examines skeptical threats arising from interactionism in his
sustained discussions of mind-body relations in the \textit{Short Treatise on God, Man, and His
Well-Being}, an early work that exemplifies a transitional period in his thought.\textsuperscript{197} At this
point, Spinoza is still grappling with his Cartesian inheritance. This is typical—
philosophers of his time and the present are, of course, well aware of the traditional

\textsuperscript{196} Heil, “Mental Causation,” 19

\textsuperscript{197} In what follows, all quotations from Spinoza’s \textit{Short Treatise on God, Man, and His Well-
Being} will be denoted using KV and the associated Gebhardt pagination.
interaction problem of how mental and physical substances could possibly interact even in principle. But few interpreters have fully appreciated that Descartes saddles later philosophers with a set of interaction problems beyond this traditional, perhaps too strong appeal to the metaphysical impossibility of interaction. This set of interrelated problems, Spinoza seems to realize, threaten philosophy with a relapse into skepticism. Responding to these non-traditional interaction problems, it might seem, motivates Spinoza’s introduction of a strict explanatory barrier between the attributes in the Ethics, an apparent change from his stance in the Short Treatise. Relatedly, I want to use these interaction problems to construct arguments, on Spinoza’s behalf, against both Cartesian dualism and physicalism that lacks the EB.

In the Ethics, Spinoza concludes that effective naturalism requires causal and conceptual isolation between the attributes. Moreover, this causal and conceptual isolation guarantees the methodological autonomy of the physical sciences, thereby vouchsafing the success of the project that is dearest to Spinoza’s heart. As the previous chapter’s opening quote suggests, Spinoza strives to create a stronger and more enduring nature in himself, by understanding “as much of Nature as required for such a nature,” and to strive that other inquirers gain it with him (TIE, 13-14, G II/8-9). The EB serves a dual purpose, solving several serious interaction problems from Spinoza’s Cartesian inheritance while ensuring scientific investigation’s methodological autonomy. Only after installing the explanatory barrier between the attributes does one receive the keys to the twin kingdoms of Spinozistic metaphysics and natural science.
3.2 Interactionism in the *Short Treatise on God, Man, and His Well-Being*

Spinoza’s *Short Treatise* is a confusing text. Though Spinoza remains preoccupied with its questions in the *Ethics*, often presenting certain sections of argument very similarly to how they appear in the *Short Treatise*, the early work is much stranger. In particular for our purposes, Spinoza maintains a quite *unstable* explanatory barrier between the attributes. This instability in Spinoza’s reflections on the mind-body union, more than anything else, will be my focus here. Like other interpreters, in this instability I see Spinoza grappling simultaneously with his Cartesian heritage and his maturing, naturalistic metaphysics. But we can make this simple observation more precise.

I will first provide some evidence that Spinoza’s position on interactionism in the *Short Treatise* is inconsistent with the one he takes in the *Ethics*. This inconsistency arises against the background of Spinoza’s consistent assertion of the infinitude of the attributes, I position I think he maintains from the *Short Treatise* through the *Ethics*. This combination of causal interaction and attribute infinitude spells trouble for Spinoza’s framework. I show this with detailed discussion of several places in the earlier treatise where Spinoza seems to become aware of what I call the *problem of pathological causation*. Though I’ll say much more about Spinoza’s understanding of this problem in what follows, it’s useful to have Spinoza’s suggestive statement of it in mind throughout.

Spinoza raises this worry for his position in the *Short Treatise*:

The second possible objection is this: we see that the soul, though it has nothing in common with the body, nevertheless can bring it about that the spirits, which would have moved in one direction, now however move in another direction—

---

198 Actually, as I’ll argue later in the chapter, even attribute plurality (i.e. more than a single attribute) combined with causal interactionism gets the problem off the ground. But attribute infinitude, I think, makes it worse, and I provide additional arguments in support of this further claim as well.
why, then, could they not also make a body which is completely at rest begin to move? Similarly, why could it not also move wherever it will all other bodies that already have motion? (*KV* I/97).

The problem of pathological causation, it turns out, can profitably be interpreted as a generalization or extension of the problem of causal pairings that is widely discussed in the contemporary philosophy of mind literature. In broad outline, the threat of pathological causation occurs when, in a context where two attributes differ in their essential properties, we wonder how to provide principled limits for the scope—both spatial and temporal—of the action of modes of one attribute on another. Why does the sphere of my *immediate, intentional* control only include my body, and not the bodies of others, or of inert objects? Is this always the case, or just mostly so? Under what conditions are cross attribute causal predictions justified—I will to raise my arm, and am confident it will rise, but not justified in supposing that some spatially distant body will move when I will it to do so? In the *Short Treatise*, Spinoza seems confident that he can resolve the trouble that the threat of pathological causation creates for his system, but the solutions he adopts there are far from satisfactory. Indirect confirmation that they may be conceptually unsatisfying on Spinoza’s own terms comes from his replacement of these positions in the *Ethics*. There, he forestalls the appearance of pathological causation through the imposition of a rigid explanatory barrier, wedding his PSR-based naturalism to causal and conceptual isolation between the attributes.

---

Before examining pathological causation, I note a background point of consistency between the two works, one that receives, however, far greater attention in the *Short Treatise*: the cardinality of the attributes. In chapter one, I suggested that some interpreters, most notably Jonathan Bennett, have suggested that Spinoza functionally embraces an ontology containing only two attributes, even while paying lip service to their infinitude. Others, including Newlands, Della Rocca, and Curley, have suggested that God is simply the being with *all possible* attributes, however many there are. This is right, but uninformative and overly noncommittal. At least in the *Short Treatise*, where these problems first rear their ugly heads, Spinoza expresses preemptory dissatisfaction with any Bennett-style interpretation of him, as maintaining the existence of only Thought and Extension. He says:

> From what has now been said, it is clear that the Idea of infinite attributes in the perfect being is no fiction. But we shall still add the following: After the preceding reflections on Nature we have not yet been able to find in it more than two attributes that belong to this all-perfect being. And these give us nothing by which we can satisfy ourselves that these would be the only ones of which this perfect being would consist. On the contrary, we find in ourselves something which openly indicates to us not only that there are more, but also that there are infinite perfect attributes which must pertain to this perfect being before it can be called perfect (*KV* I/17, 34-42).

According to Spinoza, expressing infinite attributes, where ‘infinite’ specifically delimits far more than two, is a condition for anything’s being called perfect, and so for

---

200 Bennett’s extended treatment of this question, in *A Study of Spinoza’s Ethics*, especially 74-79, is that the doctrine of more than two attributes was such that “Spinoza was under no pressure to accept it, nothing in the *Ethics* unquestionably means it, none of the work’s structures reflect it, and if it were added to the *Ethics* and fully developed it would create an impossible problem for Spinoza’s epistemology” (78). He also discusses several ways of reading ‘infinite’ as ‘all’ or ‘all possible’, though he does not discuss the passage cited immediately above the present footnote. Moreover, we have already seen, in chapter one, that Melamed offers a satisfying solution to the problem Bennett claims is impossible.
being God. Spinoza makes this point clear, and repeats it several times elsewhere. For instance, he states:

Therefore, the essence of the soul consists only in the being of an Idea, or objective essence, in the thinking attribute, arising from the essence of an object which in fact exists in Nature. I say of an object that really exists, etc., without further particulars, in order to include here not only the modes of extension, but also the modes of all the infinite attributes, which have a soul just as much as those of extension do (KV I/19, 6-13).

Spinoza endorses the same claim in the Ethics, without giving any indication that he has changed his mind or the sense of the words that appear to commit him to it.

But the relevant problem is philosophical, not merely interpretive. In what follows, I will focus primarily on cases of pathological causation between the modes of Thought and Extension. It will thus be easy, and tempting, to forget that such coordination problems occur across all attributes, all infinitely many of them. Moreover, the human mind cannot know more than the two attributes of Thought and Extension, even though Spinoza thinks we can infer their existence (a point I discussed in chapter one). But while Spinoza attempts to justify this lack of knowledge with the EB, it’s important to remember that he regards it as a fact of our experience that we know nothing of any further attributes. But knowing nothing about the properties, particulars, or natural laws of other attributes seriously limits our ability to determine what would count as evidence that a particular mode of some unknown attribute was interacting with mental and physical objects with which we are more familiar. So if coordination problems due to
the threat of pathological causation arise between Thought and Extension, they become even more serious when other attributes are considered.\footnote{Melamed, “The Building Blocks of Spinoza’s Metaphysics,” 37, online perceptively notes: “Another philosopher who seemed to ascribe to God infinitely many attributes (and with whom Spinoza was somewhat familiar) is none other than Descartes, who claims that God has ‘countless’ attributes that are unknown to us.” This consonance is not surprising, for it seems to represent a basic commitment for most serious, traditional theists: she must acknowledge that God has far more fundamental ways of being, or attributes, than she happens to fathom. If this seems at all compelling, that one might wonder whether both Descartes and Spinoza have the more serious, infinite-attribute version of the problem of pathological causation. The thought, briefly, would be this: if God has other attributes X, Y, Z, then it’s at least possible that he created beings of those attributes. These beings could then interact with mind and matter, at least in principle. Moreover, the Cartesian theist seems to have some reason for thinking that beings of the unknown attributes were actually created—traditional ‘plentitude’, Great Chain of Being considerations—and no obvious reason for thinking that God did not create such beings (since his reasons for any creation at all, and thus most of his reasons for refraining from creating, are unknown to us). But now, suitably committed to these additional attributes and thus to the possibility of such beings, along with some reason for thinking God actually created them, we wonder at our confidence that such interaction does not already occur. A convincing contemporary parallel might be this: dark matter, which does not interact with the electromagnetic force, as ordinary matter does, is hypothesized on the basis of its gravitational effects on ordinary matter (CERN, “Dark Matter,”). This suggests one model of limited interaction—responding to one but not all fundamental forces—but by no means the exclusive one. The possibility of other sorts of limited interaction, I am claiming, should perhaps worry someone like Descartes, who is antecedently committed to the existence of “dark attributes” X, Y, and Z.}

With this background of attribute infinitude in mind, let’s examine what Spinoza thinks about the causal relationships between modes of two attributes, the mind and the body, in the \textit{Short Treatise}. As I suggested earlier, he is often unclear. Several times he offers arguments that seem to point quite clearly to the strict explanatory barrier that we find in the \textit{Ethics}. For example, he says: “When we consider extension alone, we perceive nothing in it except motion and rest, from which we find that all its effects derive. And such are these two modes in body, that \textit{there can be no other thing which can change them, except themselves}” (\textit{KV} I/91, 10-14). This sounds like a rough statement of the causal closure of the physical—all effects in bodies arise from preceding proportions of motion and rest, and no non-bodily outside force influences bodily properties.

\footnote{Melamed, “The Building Blocks of Spinoza’s Metaphysics,” 37, online perceptively notes: “Another philosopher who seemed to ascribe to God infinitely many attributes (and with whom Spinoza was somewhat familiar) is none other than Descartes, who claims that God has ‘countless’ attributes that are unknown to us.” This consonance is not surprising, for it seems to represent a basic commitment for most serious, traditional theists: she must acknowledge that God has far more fundamental ways of being, or attributes, than she happens to fathom. If this seems at all compelling, that one might wonder whether both Descartes and Spinoza have the more serious, infinite-attribute version of the problem of pathological causation. The thought, briefly, would be this: if God has other attributes X, Y, Z, then it’s at least possible that he created beings of those attributes. These beings could then interact with mind and matter, at least in principle. Moreover, the Cartesian theist seems to have some reason for thinking that beings of the unknown attributes were actually created—traditional ‘plentitude’, Great Chain of Being considerations—and no obvious reason for thinking that God did not create such beings (since his reasons for any creation at all, and thus most of his reasons for refraining from creating, are unknown to us). But now, suitably committed to these additional attributes and thus to the possibility of such beings, along with some reason for thinking God actually created them, we wonder at our confidence that such interaction does not already occur. A convincing contemporary parallel might be this: dark matter, which does not interact with the electromagnetic force, as ordinary matter does, is hypothesized on the basis of its gravitational effects on ordinary matter (CERN, “Dark Matter,”). This suggests one model of limited interaction—responding to one but not all fundamental forces—but by no means the exclusive one. The possibility of other sorts of limited interaction, I am claiming, should perhaps worry someone like Descartes, who is antecedently committed to the existence of “dark attributes” X, Y, and Z.}
In a footnote to the already cited passage outlining the problem of pathological causation, which I discuss more below, he makes a similar claim, this time explicitly inferring parallelism between modes of Thought and Extension, much as he does in the *Ethics*. He argues:

Since man had a beginning, no attribute is to be found in him other than those which were already in Nature. And since he consists of a body such that there must necessarily be an Idea of it in the thinking thing, and that Idea must necessarily be united with the body, we affirm without hesitation that his soul is nothing but this Idea, in the thinking thing, of this body of his. And because this body has a proportion of motion and rest, which is determined and continually changed by external objects, and because no change can occur in the object, unless the same thing also actually occurs in the Idea, the result is that people have reflexive ideas. I say 'because it has a proportion of motion and rest,' because no action can occur in the body without the two concurring (\(KV\ I/98, 24-36\)).

Here, Spinoza asserts that the soul is nothing but the idea of the body, and, more importantly, that every bodily change has a corresponding mental change. Only the last line represents some small slippage characteristic of the *Short Treatise*, rather than the *Ethics*. In that line Spinoza appears to explain the Idea’s properties with reference to the proportion of motion and rest in the body (“people have reflexive ideas…because the body has a certain portion of motion and rest”). This violates 2p7s’s dictum in the *Ethics*: “Hence, so long as things are considered as modes of thinking, we must explain the order of the whole of nature, or the connection of causes, through the attribute of Thought alone.” Still, sympathetic readers might interpret Spinoza’s words in the *Short Treatise* as

---

202 Wolf, in his commentary on this footnote, glosses the penultimate sentence as follows: “And as this body has a motion and rest (which has its proportion determined, and is usually altered, through external objects) and as no alteration can take place in the object without occurring also immediately in the Idea, the result is that people have reflexive ideas” (Wolf (trans), *Short Treatise*, 129). I use the Wolf translation here because it is clearer than Curley’s in its final lines.
saying that the overall proportion of motion and rest determines the body’s properties.

Then, *since mind and body are parallel modes* ("because no change can occur in the object, unless the same thing also actually occurs in the Idea"), we can infer that there is an idea that has properties that parallel the bodily property of having a certain proportion of motion and rest. This feature of parallelism causes the idea and its object, the mind and the body, to concur, as Spinoza describes it, though there need not be an actual causal interaction between them. This more careful statement does not violate the *Ethics*’ stricter explanatory barrier.

In addition to passages like these, which appear to closely prefigure Spinoza’s views in the *Ethics*, there are contrary hints that Spinoza holds something much closer to the Cartesian conception of mind-body interactionism. So, immediately after concluding that no modes of thinking can produce motion in a resting body, Spinoza continues: “But according to what we perceive in ourselves, it can indeed happen that a body which is now moving in one direction comes to move in another direction” (*KV* I/91, 22-24). Here, bodies come to move in different directions according to internal (presumably mental) perceptions. This quote might be read as inconclusive, but later, Spinoza envisions a sort of war for causal control of the motion of the body, and suggests that this conflict can result in anxiety. This discussion of anxiety paints a clearer picture:

> And because these spirits can also be moved by the body, and so determined [in their direction], it can often happen that having their motion in one direction because of the body, and in another because of the soul, they bring about those anxieties which we often perceive in ourselves, without knowing the reasons why we have them (*KV* I/92, 19-25).

This passage clearly contradicts the view of mind-body isolationism we find in the
Ethics, as well as passages immediately preceding this one, where Spinoza explicitly attributes all the properties of bodies to the overall proportion of motion and rest.

Spinoza seems sensitive to these tensions between his inherited Cartesian interactionist commitments and the parallelism he endorses in the Ethics. Several times, cited below, Spinoza makes appeal to the close union between mind and body to resolve the apparent contradictions in what he says about their interaction. This move too, of course, has Cartesian sources. In the Sixth Meditation, Descartes observes: “Nature also teaches me, by these sensations of pain, hunger, thirst and so on, that I am not merely present in my body as a sailor is present in a ship, but that I am very closely joined and, as it were, intermingled with it, so that I and the body form a unit.”203 This intermingling, of course, has puzzled commentators for centuries. I want to examine what Spinoza says about this union, and to investigate what he might mean in developing the pathological causation objection, as well as how he proposes to combat it. Crucially, his response to the objection in the Short Treatise fails to resolve the issue, a fact this response’s disappearance in the Ethics seems to confirm. This sets the stage for considering how this failure might be turned into an explicit argument for Spinoza’s preferred view, which he adopts in the Ethics.

Immediately after arguing both that all effects of bodies derive from proportions of motion and rest, but also that we perceive in our soul an ability to move our own body, in flagrant violation of his preceding claim, Spinoza tries to resolve the inconsistency. He says: “The cause of this is, and can only be, that the soul, being an Idea of this body, is so united with it, that it and this body, so constituted, together make a whole” (KV I/91, 29-
A few pages later, having suggested that the body’s only effect on the soul is to cause the soul to perceive the body, and through its mediation, other external bodies, Spinoza explains this union by connecting love with knowledge. He says: “So whatever else apart from this perception happens to the soul cannot be produced through the body. And because the first thing the soul comes to know is the body, the result is that the soul loves the body and is united to it” (KV I/93, 10-13). This allows him to speculate that we might overcome our passions—love, hate, and jealousy. Their origin, Spinoza thinks, can only be in the soul, since he fails to see how such passions could arise from motion and rest alone. We can overcome them by coming to clearly know, and thus to love, God more than our own body, thereby transferring the mind-body union to union with God and freeing ourselves from the former. Needless to say, this mystical vision does little to explain our more pedestrian puzzlement concerning the nature of the former mind-body union.

The basic idea is clear enough, however. Spinoza intends to argue that the problems that would arise if a wholly thinking thing were to cause motion or changes in a motionless external body, like a stone, do not actually arise within a mind-body union. The same problems do not come about because the mind and the body are, in a sense, the same thing, one whole. As Wolf puts the point: “Similarly, mind and body are really one whole, merely a double-faced mode of substance; mind does not affect body, nor body mind; the one simply is the other—that is, a parallel aspect of the same reality.”

Wolf describes Spinoza’s mature view in this way, but I think he would agree that this cannot be the view that Spinoza endorses univocally throughout the Short Treatise, for at least

---

204 Wolf, *A Short Treatise on God, Man, and His Well-Being: Commentary*, 228-229
two reasons. First, wholes have questionable ontological status in the *Short Treatise*. Spinoza characterizes them in this way: “Part and whole are not true or real beings, but only beings of reason; consequently in Nature there are neither whole nor parts” (*KV* I/24, 19-21). So either the whole composed of mind and body does not exist to do the explanatory work required of it, or mind and body form a whole in some different, as yet unexplained sense. Second, even after introducing the tight union between mind and body and characterizing the two as bound by the soul’s love, Spinoza continues to speak of each affecting the other. This lingering interactionist commitment, I suggest, renders Spinoza incapable of satisfactorily answering the objection he puts himself. This objection—the threat of pathological causation—eventually serves as the foundation for Spinoza’s argument for the EB. I now examine this threat in detail.

### 3.3 The Threat of Pathological Causation

Spinoza outlines the threat of pathological causation as an objection with two parts ([A] and [B][C]). He says:

The second possible objection is this: [A] we see that the soul, though it has nothing in common with the body, nevertheless can bring it about that the spirits, which would have moved in one direction, now however move in another direction—why, then, [B] could they not also make a body which is completely at rest begin to move? Similarly, why could [C] it not also move wherever it will all other bodies that already have motion? (*KV* I/97, 1-8)

It’s crucial to recognize the structure of Spinoza’s objection here. The problems begin when we grant [A]—that the mind can direct the bodily spirits despite having nothing in common with them. Bewilderment at the bare possibility of such mental-physical interaction, of course, is the traditional interaction problem pressed most
forcefully by Princess Elizabeth. In the *Ethics*, this view comes in for particularly harsh criticism. In the Preface to Part V, Spinoza suggests that it is almost unfathomable that a philosopher of Descartes’s caliber could have believed such a thing. He says: “Such is the opinion of that most distinguished Man—as far as I can gather it from his words. I would hardly have believed it had been propounded by so great a Man, had it not been so subtle” (G II/279, 17-19).

Of course, this criticism of Descartes is either convenient intellectual hypocrisy or self-serving forgetfulness, since Spinoza clearly holds much the same thing in the *Short Treatise*. The short argument he gives in the Preface to *Ethics* Part V is more interesting, since he takes it to rule out, apparently independent of his other commitments in the *Ethics*, mind-body interaction at all. It relies on a general principle about causation, which requires that there be some common measure between cause and effect. He notes: “Of course, since there is no common measure between the will and motion, there is also no comparison between the power, or forces, of the Mind and those of the Body. Consequently, the forces of the Body cannot in any way be determined by those of the Mind.”

This short argument jives well with contemporary ones in the literature on the problem of causal pairings, which I’ll discuss next chapter. It’s also a version of what Robb and Heil call the Causal Nexus thesis: “Any causal relation requires a nexus, a means by which cause and effect are related.”205 Some contemporary philosophers argue further that only a network of spatial, presumably wholly physical relations provide sufficient common ground to underwrite unique causal relations, which allow for the

205 Heil and Robb, “Mental Causation,” 8
avoidance of pairing problems between mental and physical events and states. Spinoza, by contrast, is not a physicalist. But his related thought in the Preface to *Ethics* Part V is that, absent such a network, “the forces of the Body cannot in any way be determined by the Mind.” Determination and determinacy are the crucial concepts—lack of a common measure prevents precise, unique causal interaction between things otherwise wholly distinct, a point I will return to next chapter.

Nonetheless, in the *Short Treatise* Spinoza, quite interestingly, does grant [A]—that the soul can direct motion in the body despite having nothing in common with it. That is, he sets aside the traditional interaction problem, assuming that such interaction is at least epistemically possible. This gives rise to two further possibilities, Spinoza’s objector thinks. First, it seems then that the mind should be able not only to redirect motion, but also to cause new motion in a motionless body. Notably, Descartes denied the mind this ability to cause motion in the previously motionless, thinking it the only way to preserve conservation laws in physics, as he understood them. Second, it seems like the soul should be able to redirect “where it will all other bodies that already have

---

206 Jaegwon Kim, in *Physicalism or Something Near Enough*, and elsewhere, has perhaps done the most to develop this line of argument, as we will see next chapter.

207 Of course, Spinoza’s final position in the *Ethics*, which contains the EB, affirms the impossibility of mind-body interaction and thus endorses the traditional formulation of the interaction problem. My claim is simply that Spinoza does not merely accept this traditional formulation either without argument, or based only on the short argument in the *Ethics* (whose causal principle just seems to beg the question against the interactionist dualist). Rather, Spinoza sees his system’s ability, and the interactionist’s inability, to respond to the problem of pathological causation as the explanation for the truth of the traditional formulation, which might otherwise just seem to beg the question against the dualist.

208 See, for example, Garber, *Descartes Embodied*, 133-167. Central to Leibniz’s objection to Cartesian conservation laws is that the preservation of the quantity of motion is insufficient, and thus Descartes’ solution—to limit the mind to redirecting preexisting motions, rather than causing new ones—fails, previously because the quantity of motion and direction of motion must be conserved. Leibniz defends what we now describe as the classical conservation laws of energy and momentum.
motion.” We might add a third possibility as the combination of the first two: namely, why the soul could not cause new motion in all other bodies where it will. For ease of exposition later in the chapter, I’ll simply focus on one example of this sort. But we can define **Pathological Causation** as a broad class of variations on this theme, whereby the natural assumption that minds have only immediate, intentional control over one and only one body, the body with which they are commonsensically associated, is violated. Pathological causation involves a mind’s immediately causing the motion, or changes in motion, of bodies distinct from one’s own.

Spinoza’s abstract language obscures the radical nature of the pathological causation he envisions here. But we can make matters more concrete. Spinoza’s imagined objector wants to know how, once we have granted mind-body interactionism, we can preclude my mind’s being able to cause my colleague’s bike, currently chained to the fence over there, to move. That would be an instance of pathological causation. He wants to know how we could have principled reasons for ruling out my mind’s ability to immediately cause or redirect motion in any physical object whatsoever, once something like soul-body causation is granted. Why should such abilities be seen as impossible in principle, or weaker, simply non-actual given the actual laws of nature plus the dualist’s commitment to mental-physical interaction. That is the general epistemic threat of pathological causation. What justification do we have for thinking such pathological instances are non-actual or even particularly uncommon, given the dualist’s metaphysical commitments? What good explanation can the dualist give to justify the denial of actual pathological causation, given her interactionist metaphysics? I state the problem here epistemically, but, depending on how we interpret Spinoza’s commitment to the principle
of sufficient reason, such pathological epistemic possibilities may quickly acquire metaphysical import. Possibilities impossible to account for epistemically become, on some readings, impossible tout court. The precise nature of the problem deserves further scrutiny, and I’ll discuss it more once I’ve examined Spinoza’s own inadequate solutions.

In the Short Treatise, Spinoza thinks that we should be able “to remove this difficulty very easily” (KV I/97, 10). This is further confirmation of the development of Spinoza’s views by the time he writes the Ethics. There, this problem not only remains difficult; Spinoza claims that it is insoluble. As the Preface to Part V makes clear, Spinoza finds nothing worth salvaging in Cartesian interactionism. In the Short Treatise, however, Spinoza tries to split the difference by gesturing heavily to his position in the Ethics while retaining interactionism. His attempt is worth quoting at length:

But if we recall what we have already said about the thinking thing, we will be able to remove this difficulty very easily. We said that that although Nature has different attributes, it is nevertheless only one unique Being, of which all these attributes are predicated. We added that the thinking thing is also unique in Nature and that it is expressed in infinite Ideas, according to the infinite things that are in Nature. For if the body receives one mode, such as, for example, Peter’s body, and again another, such as Paul’s body, the result of this is that there are two different ideas in the thinking thing. One Idea of Peter’s body, which makes the soul of Peter, and another of Paul’s body, which makes Paul’s soul. So then, the thinking thing can indeed move Peter’s body, through the Idea of Peter’s body, but not through the Idea of Paul’s body. So Paul’s soul can indeed move his own body, but not that of someone else, such as Peter. And for this reason it also cannot move a stone at rest. For the stone makes another Idea again in the thinking thing. Hence it is no less clear that it is impossible for a body completely

209 Of course, one need not use the PSR to get this conclusion, and in fact the bare PSR will not give it to you, as I argued before. Only some further premise making the postulated explanations accessible to human knowers will do that. But still, any contemporary philosophical view that is anti-skeptical, and that embraces tight connections between conceivability and possibility, has the same consequences, if some principled limitation on mind-body interaction is actually inconceivable.
at rest to be able to be moved by any mode of thought, for the reasons given above.

The density of Spinoza’s prose obscures the steps in his reasoning, but moving through it slowly can uncover some of the assumptions underlying this argument, many of which Spinoza gives up in the Ethics. So I will first make clearer what Spinoza intends in this passage, before proceeding to evaluate his claims.

Spinoza’s first thought is that Nature is a single unified thing, albeit with infinite attributes which are expressed in, among others, infinite modes of Thought and Extension. For each mode of extension, there is a corresponding idea (“the result of this [there being two bodies, Peter’s and Paul’s] is that there are two different ideas in the Thinking thing”). Paul’s body has a corresponding soul, and Paul’s soul is an idea in the infinite thinking aspect of Nature. Similarly, Peter’s soul is a different idea in the same infinite thinking thing. Thus far things are reasonably straightforward, but they soon begin to move quickly.

The chain of reasoning, upon which I’ll focus my criticism, seems to be the following. Since Nature is unified in this way, there is some sense in which all causation looks a bit pathological, at least from God’s perspective. God or Nature can move Peter and Paul’s body in various and perhaps unexpected ways wherever it wills. Spinoza just cheats a bit, limiting this divine power by assuming that extended and thinking modes are paired 1-1, and that God’s power is expressed only while respecting these 1-1 pairings. With this implicit restriction in mind, then, since each extended mode has a thinking counterpart, we can say that each thing’s soul (the thinking correlate of its body) can causally affect its very own body, but not that of another. So if Paul’s body moves, it
moves because of Paul’s idea in the thinking thing, and not because of Peter’s idea, and vice versa. Moreover, Paul’s idea only exercises causal power over Paul’s body, and not over, for example, some stone at rest. For any movement in the stone presumably comes about either from physical interaction with other bodies, or from its own idea (“For the stone makes another Idea again in the thinking thing”).

Spinoza’s last line remains puzzling on any reading though—why does it follow that “it is impossible for a body completely at rest to be able to be moved by any mode of thought?” The focus on a body at rest is not puzzling—this refers, as mentioned earlier, to Descartes’ mistaken belief that physical conservation laws could be respected only if minds redirected preexisting motions, rather than causing new motion from rest. The reference to ‘any mode of thought’, however, still is puzzling. Perhaps Spinoza means here that a stone at rest cannot be moved by anything traditionally conceived as a mode of thought. That is, the stone at rest is moved by its corresponding idea in the thinking thing—here we see intimations of Spinoza’s maturing panpsychism—but not by the thoughts of any old Peter or Paul, that is, by anything typically conceived of as mental. Or perhaps he means that a stone really at rest would have a thought that was isomorphic in some way—structurally ‘at rest’, whatever that may mean. Any subsequent motion in the stone thus requires some outside influence, though this influence could only come from some physical interaction, rather than from some immediate mental cause. This last reading would partially duplicate the first, focusing our attention on the preservation of Cartesian conservation laws, while adding that any immediate changes in the motion of the stone could not have immediate mental precursors.
In any case, it’s reasonably clear what Spinoza wants here. Unlike in the *Ethics*,
the explanatory barrier is not yet fully in place (“the thinking thing can indeed move
Peter’s body, through the Idea of Peter’s body”). However, Spinoza envisions seriously
limited possibilities for interaction; the only mental-physical causation explicitly allowed
for here moves *from* a body’s corresponding idea *to* that body. It’s even unclear whether
the reverse—from a body to its corresponding idea—occurs. Regardless, this should be
enough to puzzle us, for these limited possibilities for interaction are supposed to be
*inferred or derived* from Spinoza’s starting point, which helps answer the objector: the
unity of Nature. To repeat: “But if we recall what we have already said about the thinking
thing, we will be able to remove this difficulty very easily. We said that that *although
Nature has different attributes, it is nevertheless only one unique Being, of which all
these attributes are predicated.*” My claim is simply that Spinoza lacks the metaphysical
resources here to restrict cross-attribute (i.e. between minds and bodies) causation in this
way.

There are numerous controversial steps here. For one thing, Nature’s unity does
not seem to require, in any obvious way, that there be a one-one correspondence between
each and every thinking and extended mode. Commonsensically, for instance, one might
suppose that within unified Nature there are many wholly extended things (rocks,
sailboats, etc.), and a much smaller subset, within the same natural unity, of thinking ones
(i.e. people, some animals). Suppose, though, that for this smaller subset, a one-one
correspondence between thinking modes (person or animal souls) and extended modes
(animal and person bodies) at least seems more natural. Still, nothing about Nature’s
unity, from anything Spinoza argues for here, seems to require that this relationship
between soul and affected body be one (soul)-one (body), as opposed to one soul-many bodies.

3.3.1 Uniqueness and Privilege as Crucial Principles, and Their Evaluation

If Spinoza wants to allow for cross attribute causation without pathological causation, he needs something like the following two principles to be true. Ideally, they would also follow from what he argues here, or at least be plausibly inferred from other resources in the *Short Treatise*.

**Uniqueness**: A given finite mind, throughout its career, causally interacts with just one body.

**Privilege**: A given finite mind, throughout its career, is the only mind that can directly causally influence a given body.\(^{210}\)

I have reciprocal interaction in view here, rather than what is sometimes called downward or upward causation, and the conjunction of Uniqueness and Privilege entails

---

\(^{210}\) I take these terms from Audi, “Primitive Causal Relations and the Pairing Problem,” 13, which I will discuss at greater length next chapter. However, I state them differently, removing the modal operators, as the modality with which they’re stated makes them difficult to employ for my purposes. I also add a reference to ‘finite minds’, rather than minds in general, though I am not always careful about this. This is to rule out the entailment from traditional theism, whereby God causally interacts with my body in creating and sustaining it, to epiphenomenalism, whereby my mind’s interaction with my body is outcompeted by God’s. If we bracket traditional theism we can ignore this in what follows. One might think that Spinoza can secure one-one, mind-body correspondence with an identity theory, but this seems like a problem for two reasons. First, nothing Spinoza says here provides any argument for an identity theory, and much of what he says cuts against it. For example, it’s hard to make sense of a soul coming to love one particular body, as we saw Spinoza argue earlier, and thus coming to be united to it, in any non-metaphorical sense that also allows them to be one identical thing. One identical thing cannot interact causally with itself, at least in a way that allows for any change. (Spinoza’s God as *causa sui* does not change—see 1p20c2). To stick with an earlier example: one identical thing cannot give motion to the spirits in one direction, qua soul, and give them motion in another direction, qua body (*KV* I/92, 19-25). So I take it that talk of interaction here rules out an identity theory at this point, though of course Spinoza will adopt such a theory in the *Ethics*. The point is, however, that he can’t yet have that theory because he hasn’t yet got the explanatory barrier between the attributes—I argued for this in chapter one.
that one and only one mind directly causally influences a single body.\textsuperscript{211} Of course, indirect causal influence among many minds and bodies is pervasive. On this model, if my friend slaps my arm, my brain registers the pain and my mind thinks, “That guy is a jerk,” but this foreign hand’s effect on my thoughts does not thereby violate Uniqueness. Rather, the principles attempt to codify the especially close relationship that obtains between the mind, as the dualist understands it, and exactly one body with which it interacts causally in an intimate way.\textsuperscript{212}

If Uniqueness and Privilege are true, then, even assuming mind-body interaction, Spinoza’s envisioned objector has nothing to worry about. There is no threat of pathological causation. The soul could not “move wherever it will all other bodies that already have motion.” Instead, Uniqueness guarantees that each soul interacts with just one body, and Privilege guarantees that no other mind can manipulate the body Uniqueness reserves for this particular soul. Unfortunately, Spinoza does not make his commitment to something like these two principles explicit, much less argue for them, and he certainly cannot simply assume that they are true. Their truth is precisely what his

\textsuperscript{211} Of course, there might be dualists who embrace only downward (mind to body) or only upward (body to mind) causation. Just upward causation gives a sort of mental epiphenomenalism, while just downward causation implies that none of our sensations, caused by bodily interaction with our environment and conveyed to our brain, causally influence our thoughts. This essay cannot, however, rule out every possible dualist position, nor does it seek to do so. The standard Cartesian, and inherently more plausible, interactionist picture involves causal interactions as reciprocal, even if one ignores general worries about whether one causing thing could ever cause some caused thing to occur without being affected by the caused object in return. Kant’s view on this general question, for instance, was that all causal interactions were reciprocal, something that appears to have substantial support from modern physics, at least for those things, whatever they are, within the physical domain.

\textsuperscript{212} Other commentators have noted Spinoza’s apparent commitment to something like Uniqueness and Privilege in the Ethics. There, of course, this commitment depends on the logic of mind-body identity, as, for example, Elhanan Yakira, \textit{Spinoza and the Case for Philosophy}, 123 makes clear: “Its idea, or soul, is its idea no matter what (or in every possible world)...One consequence of the necessary identity of the body with its idea is that a certain-idea/certain-body relation is unique in a particularly strong sense of the word.” See also, Renz, “The Definition of the Human Mind and the Numerical Difference Between Subjects (2p11-2p13).”
imagined objector is worried about, especially given that the soul and body have no common measure, and thus that we seem to lack all concept of how they might interact at all. That is, given that minds and bodies have nothing in common, how could it be (and, relatedly, how could we know) that such precise and restricted conditions on cross-attribute interaction obtain?

One might think that the facts of experience are all that’s required to secure the truth of Uniqueness and Privilege. Paul Audi, for example, whose use of these principles I discuss next chapter, suggests: “It seems a plain fact that I have a single body that I alone control.”213 This claim is crucially ambiguous, for it might mean either that we directly experience our control over one and only one body, in some sort of direct realist way. Or it might be taken as elliptical for an appeal to our experiences as inductive grounds that provide evidence for the truth of the principles. Perhaps our experiences of exercising conscious mental control over a single body (our own), of trying to exercise control over other spatially disconnected bodies and failing, and of never experiencing any other mind’s controlling of our body, are enough to motivate the two principles. This inductive generalization would not make such interaction metaphysically impossible, but perhaps it would justify the general claim that such pathological interactions do not occur, given the actual laws of our world.

Of course, appealing to the facts of experience would not explain why the two principles are true, but perhaps we shouldn’t expect that from a theory of mind-body interaction. Nevertheless, this strategy will not work. In fact, at least three sorts of experiences cast doubt on the truth of Uniqueness and Privilege (Privilege Violations, 213 Audi, “Primitive Causal Relations and the Pairing Problem,” 12 166
Uniqueness Violations, and Temporal Oscillations). If either principle is undermined, then the whole suite of cases of pathological causation can be developed in terms of the violated principle. So all that’s required on my part is that one of the three classes of experiences undermines either the direct experience of the sort of control that Uniqueness and Privilege stipulate, or the inductive generalization from experience to the truth of the two principles and thus to the non-actuality of cases of pathological causation. Because of how the cases are constructed, if any class of counterexample is successful, it follows that there will be many actual violations of Uniqueness or Privilege, rather than a single isolated incident, even if the many violations are still relatively rare. I discuss each briefly in turn.

**Privilege Violations:** Spinoza is an honest and perceptive psychologist of the emotions, and honest, perceptive emotional psychology renders any simply appeal to the facts of our experience unlikely to succeed here. Spinoza is not overly skeptical about our ability to understand our own mental states. But he does observe that competition between purely bodily causes of motion, and the soul’s causal contribution to the motion of its own body, “brings about those anxieties which we often perceive in ourselves, without knowing the reasons why we have them” (KV I/92, 24-25). Healthy individuals commonly find themselves in mental states they cannot explain, or find themselves appearing to act from unconscious desires they did not realize they had.

Depression, anxiety, and more severe forms of mental illness, such as schizophrenia or disassociative identity disorder, represent magnified versions of this typical experience. Those suffering from mental illness at least sometimes report experiencing negative emotions that appear to arise from somehow external, yet totally
opaque, causes outside of the person. The connection to Uniqueness and Privilege is this. According to the defender of the two principles, our experiences are supposed to provide evidence for an easy one-one correspondence between our mind and body, such that our mind, and only our mind, affects this (and only this) body. But we often observe ourselves in mental states that we do not experience as caused either by our own mental states, or by our bodies, or by any external bodies. Some suffering from depression report serious versions of this: a desire to kill themselves, for instance, which seems to come from nowhere. Of course, not experiencing our own mental states as caused by our own mind or body is not yet to experience them as caused by some other mind or body, but given the contrast with the ordinary case it does seem to undermine our confidence in the principles somewhat.

The phenomenon of ‘thought insertion’ is perhaps the most extreme putative violation of Privilege. So while I think that depression and inexplicable anxiety are common and relatable forms of this general class of phenomena, I mean for thought insertion to bear the brunt of the argumentative load here. Ordinary instances of thoughts arising randomly, as well as more serious cases of depression and anxiety, are meant to introduce violations of Privilege gradually. But again, for my purposes only the full-blown thought-insertion cases carry weight in the overall argument. So consider two examples of the thought insertion phenomenon, a classic symptom of schizophrenia:

I look out the window and I think that the garden looks nice and the grass looks cool, but the thoughts of Eamonn Andrews come into my mind. There are no other thoughts there, only his … He treats my mind like a screen and flashes thoughts onto it like you flash a picture.
Thoughts are put into my mind like ‘kill God.’ It is just like my mind working, but it isn’t. They come from this chap, Chris. They are his thoughts.\textsuperscript{214}

As these patients describe their experience, Privilege is violated in the actual world. The minds of Eamonn Andrews and ‘this chap Chris’, respectively, project foreign thoughts into the mind of the patient, and these thoughts cause them to voice certain thoughts, to take certain actions, or minimally to report on their experiences.\textsuperscript{215} Of course, I do not argue that this apparently uncaused emotion or mental state is really as fundamentally mysterious as it seems, though it may have lots of mysterious aspects. Almost certainly, there are bodily or psychological explanations available, though it is questionable whether such explanations can be appealed to without first presupposing Uniqueness and Privilege.

Rather, the idea is that such patients do not recognize the cause of their ideas as internal. This differs from ordinary cases. For example, in ordinary cases we might feel a flame’s warmth and experience our pleasant mental state as caused by the desired heat source, or might ingest something sweet and experience a happy feeling as arising from eating. In the normal, solely mental case, we often experience one thought as arising only from previous thoughts of ours, rather than from some other mental source.

\textsuperscript{214} The first example comes from Mellor, C. H. “First Rank Symptoms of Schizophrenia,” 117, while the second comes from Frith & Corcoran, "Exploring ‘Theory of Mind’ in People with Schizophrenia." I am very grateful to Kate Finley for providing these specific examples of thought insertion, and for drawing my attention to the relevance of the phenomenon to my project more generally.

\textsuperscript{215} Note that as stated previously, Uniqueness and Privilege are technically silent on the question of direct mind-mind causation, provided that there are no physical effects of these mind-mind causes. This is not Spinoza’s actual view in the Ethics, which does seem to preclude finite mind-mind causation, but it is sufficient for our purposes. Here, however, there clearly are physical effects—certain actions, or certain verbal reports—that occur as a result of the thoughts projected into the patient’s mind from the outside. So we can describe it as a putative violation of privilege even without ruling out telepathy by definition in characterizing Uniqueness and Privilege.
By contrast, when most people experience desires and thoughts that seem to come from nowhere, they nonetheless assume that the causes of such confusing, conflicted states are either in their bodies or elsewhere in their mental life. Arguably, for example, those suffering from depression or anxiety presume that this is the case. But that is an assumption we make, not something that experience presents us with directly. But these schizophrenic patients do not always assume it—indeed, they positively deny it. And part of what’s required for our assumption to be justified, which the schizophrenic does not always share, it seems, is that Privilege be true. Otherwise, we should also check to see whether some other mind—perhaps someone who wishes us ill—is causing our state of mental anguish and subsequent physical responses to it. That we do not, at least in contemporary western culture, usually check for such causes exemplifies our commitment to Uniqueness and Privilege, among other principles, but it does not represent our experience of the truth of either. The schizophrenic’s identification of some other mind as the cause of her thoughts merely makes our ordinary assumption manifest as such. If we reject the schizophrenic’s description of the case, as we might, it cannot simply be on the grounds that Privilege is a plain fact of all our experiences.

It might be less obvious why such cases of schizophrenic thought insertion undermine the defender of Uniqueness and Privilege who wishes to use ordinary experience as inductive support for the truth of the principles. Here, I do need to appeal to something like symmetry considerations. If two agents—the schizophrenic and someone who does not suffer from the disease—have radically different inductive bases, one of which justifies belief in the two principles and the other of which undermines it, then some additional reason, which does not simply presuppose the principles, should be given
for thinking that the principles are true. For evidence of the same sort now seems to both
support and falsify the principles. For those unhappy with this line of argument, the
violations of Uniqueness given below seem better suited to making my overall point,
since unlike with thought insertion, the experiential base for Uniqueness Violations can
be publicly shared in a way that the preceding examples cannot be.

**Uniqueness Violations:** Recent advances in medical technology have allowed
those recovering from serious injuries involving lost limbs to exert mental control over
bionic limbs, which seems like a *prima facie* violation of Uniqueness. Relatively, new
research has allowed patients with locked in syndrome to communicate responses to yes
or no questions via a machine that uses a combination of measures to detect changes in
brain activity associated with each response. Of course, issues are still complicated
here, and we might want to argue that direct neural connections just entail, for example,
that the bionic limb is now a part of the amputee’s body, rather than an external body.
However, this seems less plausible when the electrical signals are detected over some gap
of empty space (i.e. via antennae using near-field technology, rather than direct neural
connections to the bionic limb). And it certainly seems plausible that a sufficiently
sensitive detector could be located across the room while nonetheless responding directly
to neural firings near the amputated limb. Signal interference, rather than some in
principle impossibility, seems like the major stumbling block here. Suffice to say that
automatically inferring the bionic limb’s integration into the original body, simply from

---

216 Hargrove et. al, “Robotic Leg Control with EMG Decoding in an Amputee with Nerve
Transfers”

217 Chaudhary et. al, “Brain-Computer Interface Based Communication in the Completely Locked
In State”
the amputee’s ability to communicate with it, seems tendentious at best and question-begging at worst. Moreover, unlike with Privilege Violations, for such cases we do not need to take the schizophrenic’s word for it—though I see no reason how the dualist can explain, in a principled way, why we should not take the schizophrenic’s description of her experiences as veridical, given her metaphysics. Instead, we can simply visit a hospital where amputees are operating bionic limbs or locked-in patients are communicating through machines and see for ourselves. We can thus fill our inductive base with experiences that undermine Uniqueness.

Temporal Oscillation: Suppose finally that Paul’s soul can indeed, as Spinoza suggests, move one and only one body, and suppose further that this is (typically) his own body. Still, it doesn’t follow that Paul’s soul is never capable of directing Peter’s body. A maximally concessive objector to Spinoza’s position might admit that when Paul’s soul controls Peter’s body, it cannot simultaneously direct its own (Paul’s) body. But even popular culture has a story about why this might be. Telekinesis takes work, and one can easily imagine needing to remain perfectly still, deep in concentration, in order to assume control of another’s body. As time passes, one’s mental control of one’s own, and then another’s, and then one’s own, body might oscillate, responding perhaps to one’s concentration levels or one’s goals in such manipulations.

Evaluating the examples: Privilege Violation is an experience that many schizophrenics report as occurring in the actual world. On one way of appealing to plain facts of experience, then, the control over one and only one body stipulated in Uniqueness and Privilege is not directly experienced by all of us. This, combined with symmetry considerations, should undermine our confidence in the principles or should at
least prompt some additional, non-experiential argument for them. For those not persuaded by such cases, *Uniqueness Violation* cases represent a more direct threat to the principles. For unlike with schizophrenic thought insertion, evidence undermining the truth of Uniqueness is easily, publicly available. Such actual cases, I have acknowledged, remain difficult to interpret. Still, I want to insist that Uniqueness Violation becomes harder to render consistent with the principle as the mechanism of interaction between the patient’s mind and the body controlled (i.e. bionic limb) takes place at greater distances or without direct, ordinary, material connection (i.e. electromagnetically). Finally, *Temporal Oscillation* is obviously somewhat fanciful. But this is basically beside the point, since my goal is not to engender skepticism—skepticism that telekinesis is perhaps really possible—but to seek some explanation for our confidence in its impossibility, given the laws of the actual world and the interactionist’s metaphysics.

This failure of explanation is the defining feature of the problem of pathological causation. Spinoza’s imagined objector wants some principled explanation for why Paul’s soul controls and remains affected solely by Paul’s body, and so too for Peter’s soul and body. The objector wants to know which facts ground our attributing this body to Paul in the first place, rather than to Peter, given the non-spatiality and fundamental difference in kind between Paul’s soul and Paul’s body. In the *Short Treatise*, as we have seen, Spinoza responds confidently to his objector, while developing an ontology in many ways similar to the metaphysical system of the *Ethics*. Still, even Spinoza’s appeals to Nature’s unity, plus an objector willing to grant either one soul-one body causal interaction (for any particular time), or Privilege, do not resolve all the difficulties. For Spinoza still cannot necessarily explain *Temporal Oscillation* (for the first objector), or
Uniqueness Violations, for the second. The threat of pathological causation remains even in the context of a mostly naturalistic system, of the sort Spinoza develops in the *Short Treatise*, so long as Spinoza allows any inter-attribute causation.

What one wants to say in response to these cases, and what Spinoza does say in the *Ethics*, is something like this. Prior states of Paul’s body give rise to its present states, and so too for Peter’s body. Each of Peter and Paul’s bodily states is correlated with exactly one mental state, whose structure is isomorphic to the corresponding bodily one. The reason why Peter’s mental states cannot cause Paul’s current bodily states is not a sensible question, because interactionism, in the *Ethics*, is not just false but impossible. No mental states cause any physical states because the explanatory barrier between the attributes is true.

There is a closely related question that makes sense, however. The extended correlates of Peter’s mental states, Peter’s bodily states, cannot cause Paul’s bodily states because Paul’s current bodily state has a cause X, or a cause of type X. But Peter’s mind did not have an extended correlate, Peter’s body, with that precise cause or cause of type X previously, and so could not have given rise to that state in Paul. Peter’s mind must be so configured that Peter’s body is in the right state for causing Paul’s current bodily state, at which point, of course, we should say that it was Peter’s body, not Paul’s.

For having such a configuration—and this is obviously controversial—is a unique property of Paul’s past bodily states, and the reason why we attribute the current bodily states to him rather than to Peter. Peter is only in a position to mentally influence the exact mental correlates of what he can influence physically; that is, necessarily, Peter’s mind can only causally influence Peter’s ideas. And the bodily correlate to Peter’s soul is
not positioned to have the right sort of influence on Paul’s body that Paul does, because Peter’s body lacks the right sort of physical properties. No other body, besides Paul’s, even could have the right set of properties to act as Paul does.218

Spinoza does not take this route in the Short Treatise, and he suggests he adequately counters the objection he raises there without it. But the textual fact remains that Spinoza both abandons causal interaction, even between minds and bodies taken to be identical, and reserves some of his sharpest invective for Descartes on this point. I cannot provide conclusive evidence that this line of thinking actually caused Spinoza to change his mind on these points, nor do I intend to try to. Rather, I want to reconstruct the argument Spinoza’s imagined objector might have pressed to cause him to so thoroughly give up on interactionism in the Ethics. I happen to think that something like this argument actually did convince Spinoza to institute the EB in the Ethics, but this psychological claim floats free of my argument here. What I’m offering is a Spinozistic argument for the explanatory barrier between the attributes, regardless of how Spinoza’s commitment to the EB actually came about. Still, this argument for the EB is both grounded in Spinoza’s texts, and provides a serious response to the causal demarcation difficulties that I have identified as the threat of pathological causation. It thus represents

218 Also, cases like Uniqueness Violation no longer pose a threat, since the EB precludes all causal interaction between attributes. The residually surprising fact brought to light by such cases is that one’s body, and that body’s corresponding mental correlates, is vaster than it initially appeared, or can be made to be so, given recent technological advances. Relatedly, it may turn out, because of some strange features of Spinoza’s monism plus his sophisticated machinery of concepts or ways of conceiving, that when I adopt extremely broad conceptions of my self, for purposes of transcending my narrowly conceived self interest in order to act altruistically, I will end up conceiving of certain states of other individuals’ bodies as my own. This will occur because, on some extremely general conceptions, the rigid identification of particular individuals within God breaks down. See Newlands, Reconceiving Spinoza, Ch. 8 for more on this, but note that 1) this surprising ownership of others’ states lacks the same sort of skeptical or explanatory consequences I’ve been pushing here, and 2) occurs within a system explicitly committed to the EB.
a plausible reconstruction of one line of thinking Spinoza might have actually pursued to resolve these fundamental problems.

3.4 A Spinozistic Argument for the Explanatory Barrier Between the Attributes

In what follows, I want to gather up the major conclusions of my examination of the *Short Treatise*, in the process codifying Spinoza’s imagined objector’s insight into an argument, in two versions, for the explanatory barrier between the attributes. If we label Spinoza an isolationist—someone who accepts the truth of the EB—then dualists and physicalists are Spinoza’s primary non-isolationist targets. Chapters four and five will discuss the extent to which the considerations offered here should be persuasive to each non-isolationist camp. Predictably, my reconstructed argument for the explanatory barrier between the attributes will not convince any and all rational non-isolationist opponents. But this is not a special failing of Spinoza’s thinking, to the extent that I capture it with my reconstruction here. For I know of no philosophical argument that even approaches this standard. Still, I hope to show where Spinoza can put pressure on his non-isolationist opponents, especially against a backdrop of shared premises and a broad commitment—among dualists, physicalists, and Spinozists alike—to anti-skepticism.

The most powerful form of the argument, then, starts with the presumption that skepticism is false. As I argued last chapter, Spinoza’s commitments to the PSR and

---

219 Physicalists might be said to be non-isolationists in two senses. First, we could stipulate that if one denies Attribute Pluralism, one is a non-isolationist. Or, second, we could ask about mental and physical properties. Many non-reductive physicalists think that mental properties are in some sense distinct from physical properties (see the Causal Closure discussion in chapter 5), but they nonetheless accept interaction between them, thereby denying isolationism, which holds between things and properties of different attributes. I do not address the question of idealism in what follows, as the position has far fewer contemporary defenders, but some version of the argument to come might work if the idealist requires a distinction between two sorts of properties that nonetheless interact.
Naturalism are insufficient for avoiding skepticism. To further support this claim, I
develop a serious explanatory problem for causal interactionism between attributes while
remaining within the confines both of the PSR and of Spinoza’s naturalism. For only then
can we fully appreciate the extent to which neither the PSR nor Naturalism alone is
capable of resolving all the explanatory shortcomings (or skeptical problems, in the
language of last chapter) for which, according to other interpreters, they are introduced.
To repeat the Spinozistic Isolationist’s basic claim: Isolationism can explain why
pathological causation does not occur far better than rival systems, who can provide only
ad hoc explanations, or no good explanation at all.

The desired conclusion, which Spinoza reaches in the *Ethics*, is that effective
naturalism requires causal and conceptual isolation between the attributes. Moreover, this
causal and conceptual isolation guarantees the *methodological autonomy* of the physical
 sciences. Spinoza’s *Short Treatise* can be read as the record of his searching for a sure
path to these conclusions. Before attempting to collect these searches into an argument
for the EB, however, it’s worth first restating Spinoza’s version of the PSR and providing
a clearer definition of naturalism against which we can measure our argument.

**3.4.1 Preliminaries: PSR and Naturalism**

Spinoza states the PSR in several places, but perhaps the most explicit version
occurs in 1p11, when he says: “For each thing there must be assigned a cause or reason,
as much for its existence as for its nonexistence.” 1ax2 complements this definition by
exhaustively sorting all beings, facts, states of affairs, events, etc. into two groups,
suggesting, “What cannot be conceived through another, must be conceived through
itself.” For our purposes, I interpret the PSR as stating that every fact, state of affairs, event, or being has an explanation for its existence or truth, if it does exist or obtain, and an explanation for its nonexistence or falsity, if it does not. Recently, there have been several exciting attempts to defend the PSR against some of the most promising objections to it, including the objection that it entails necessitarianism. But I do not think I need to endorse these defenses wholesale to employ the principle, since Spinoza clearly endorses it, and since its role in the argument is primarily to demand some explanation of mind-body interaction phenomena most philosophers are interested in explaining anyway. Moreover, a much weaker demand for explanation in this context should suffice for the argument, even if Spinoza himself always accepts the stronger explanatory demand. In this section, though, I want to try to reconstruct, as closely as possible, the line of thought I take Spinoza to have pursued in his development from the Short Treatise to the Ethics, which speaks in favor of working with the full-blown PSR.

Spinoza’s embrace of naturalism, which is closely connected with his use of the principle of sufficient reason, can initially sound rather metaphorical. But as we saw in the last chapter, both Perler and Della Rocca take naturalism to be doing important work in Spinoza’s arguments against skepticism. I thus think it imperative to have some more precise definition to work with, one whose compatibility with the problem of pathological causation can be more clearly assessed. Leibniz, I suggest, provides precisely this notion of naturalism, which I think can helpfully illuminate what Spinoza has in mind when he writes, in the Preface to Part III:

Nothing happens in nature which can be attributed to any defect in it, for nature is always the same, and its virtue and power of acting are everywhere one and the same, i.e., the laws and rules of nature, according to which all things happen, and change from one form to another, are always and everywhere the same. So the way of understanding the nature of anything, of whatever kind, must also be the same, viz. through the universal laws and rules of nature.

For Spinoza, both the natures of things, as well as their laws of interaction—how they “change from one form to another”—are universal and exceptionless, in keeping with his reputed necessitarianism. Universality, or freedom from exceptions, is the key characteristic of Spinoza’s naturalism. This definition initially appears to pose serious problems for my claim that cases of pathological causation are in principle compatible with both PSR and naturalism, and thus that these principles alone cannot be enough to rule such cases out. For isn’t the exercise of telekinetic powers—my ability to mentally control physical modes unconnected with my own body, or for some other mind to control my body—one of the more intuitive examples of a case where the laws of nature would be violated?

On the contrary, Leibniz sees clearly why Spinoza’s naturalism cannot do the work of ruling out cases of pathological causation without further supplementation. In his *Discourse on Metaphysics*, Leibniz argues:

> For everything is in conformity with respect to the universal order. This is true to such an extent that not only does nothing completely irregular occur in the world, but we would not even be able to imagine such a thing. Thus, let us assume, for example, that someone jots down a number of points at random on a piece of paper, as do those who practice the ridiculous art of geomancy. I maintain that it is possible to find a geometric line whose notion is constant and uniform,

---

following a certain rule, such that this line passes through all the points in the same order in which the hand jotted them down. And if someone traced a continuous line which is sometimes straight, sometimes circular, and sometimes of another nature, it is possible to find a notion, or rule, or equation common to all the points of this line, in virtue of which these very changes must occur. But, when a rule is extremely complex, what is in conformity with it passes for irregular. Thus, one can say, in whatever manner God might have created the world, it would always have been regular and in accordance with a certain general order (DM 6).

Leibniz goes on to appeal to God’s employment of the simplest hypotheses to produce the richest phenomena to explain why God chooses the world that he does. But what is striking here is Leibniz’s claim that a violation of naturalism would be not just unimaginable, but impossible. Any phenomena, no matter how irregular, which respect logical laws of non-contradiction, can occur in accordance with natural laws. A forteriori, pathological causation, of the sort considered here, can do so as well. Naturalism, we might say, is the view that all phenomena and their changes are governed by universal developmental rules. But, as Leibniz reminds us, even the ridiculous geomancers’ predictions, and thus certainly cases of pathological causation considered here, satisfy this requirement.

This point becomes even easier to see in the context of Spinoza’s metaphysics, where God or Nature is seen as the being with infinitely many attributes. To continue with Leibniz’s example, it’s as if we jotted down a series of points, some of which were written in invisible ink inaccessible in principle to our cognitive faculties. These invisible

222 Something like David Lewis’s appeal to natural properties in, for example, On the Plurality of Worlds, will not help us here, most obviously because of the infinite number of properties of attributes 3, 4, 5…n of which we know nothing. These are not alien properties, possibly instantiated in other worlds but not in the actual world. Rather, they are actually instantiated properties of which we know nothing, actual aliens, and whose naturalness we therefore cannot evaluate. The more interesting question is whether a similar appeal is useful when we consider only Thought and Extension.
points, of course, are modes of the other attributes besides Thought and Extension. We would then be assured that some law or function described these points in the order in which they were written, but we would have absolutely no idea what it was, and no way of figuring it out. Once this example is made more concrete, and stated in the language of mental causation, its relevance to our problem should be apparent.

3.4.2 The “Progressive” Argument for EB: Compatibility with Foundational Principles

In the first stage of my formal reconstruction of Spinoza’s reasoning, I want to show a bit more clearly how pathological causation can arise within a system explicitly committed to Spinoza’s foundational principles. To do this, I’ll first present a case where pathological causation occurs without any violation of Plenitude (there are infinitely many modes of infinitely many attributes), PSR, or Naturalism. To facilitate easy transition to the next chapter, I’ll use a case similar to those that occur frequently in the literature on the causal pairings problem for dualism, which I’ve suggested can fruitfully be seen as a limited version of the pathological causation objection. Of course, at this stage we cannot assume that the explanatory barrier is true, since this is an argument to motivate it. Nor can we, therefore, assume that a mind-body identity theory obtains. As I argue at length in chapter one, Spinoza justifies this identity thesis with his Substance Monism and his Parallelism, both of which depend on the EB and thus cannot be assumed at this juncture either.

223 I haven’t motivated this suggestion really, but I reserve that task for next chapter.
There’s one final point worth reemphasizing. Recall from section 3.3 that I argued that, if the Spinoza of the *Short Treatise* wishes to remain committed to causal interactionism between Thought and Extension, Uniqueness and Privilege, at least, must be true. Of course, ultimately he does not want to remain so committed. But it’s important to see the structure of the following argument as moving from Spinoza’s foundational principles (Plentitude, PSR, Naturalism) to a putative case of pathological causation that violates Uniqueness and Privilege. Of course, all parties involved—physicalists, dualists, and Spinozists—want to deny that pathological causation, so described, is possible, either in principle or given the laws of the actual world. Within an interactionist framework, however, verbally denying that causation is pathological is much easier than explaining why pathological causation does not occur. The further thought worth reemphasizing is that this represents an asymmetry between Spinozist and interactionist accounts. Spinoza, armed with the explanatory barrier and the metaphysical results that follow from it, can explain something important—why causation is not pathological—that the interactionist cannot.

**The Case:** Consider Adam and Bob, each of whom wills to raise his arm at time $t_0$. For Spinoza, each of these acts of willing strictly corresponds to some thinking mode, which we can label $tm_1$ (for Adam) and $tm_2$ (for Bob). Superficially, $tm_1$ and $tm_2$ are quite similar. Adam and Bob are similar agents who describe their acts of willing to raise their arms, from the inside, in similar terms. We can make $tm_1$ and $tm_2$ alike in a

---

224 In the next chapter, we will see that many participants in the debate surrounding causal pairing assume that we must presume that Adam and Bob are qualitatively identical, mentally, in order to get the pairing problem between them going. At the appropriate time in the next chapter, I’ll suggest some reasons for doubting this presumption. Here it’s worth noting simply that the case can be presented smoothly without the assumption of qualitative mental identity between $tm_1$ and $tm_2$, and that pathological causation
variety of other ways as well. They both occur at $t_0$, both Adam and Bob can be located in rooms that are arranged similarly, etc. Recall too that modes of thought are not located at any particular point in space, since thoughts, for the dualist, are not extended, and space is a property of Extension (or space just is Extension). Finally, recall the assumption of Plentitude, PSR, and Naturalism. Then there seem to be three broad ways things could go:

**Ordinary:** Whenever neither the direct causal antecedents of $tm_1$ nor those of $tm_2$ are modes of some unknown third attribute or thinking modes of type X, ‘ordinary’ causation occurs.\(^{225}\) This is the sort with which, we naturally say, we are most familiar. Adam wills to raise his arm at $t_0$ ($tm_1$ occurs at $t_0$), and Adam’s mental event causes Adam’s arm (not Bob’s) to move at $t_1$. Similarly, Bob wills to raise his arm at $t_0$, and Bob’s mental event causes Bob’s arm (not Adam’s) to move at $t_1$.\(^{226}\)

**Pathological I:** However, whenever any thinking mode at $t_0$ has an immediate causal antecedent of an unknown third attribute, then at $t_1$ some pathological causal interaction occurs. Perhaps when Adam wills to raise his arm at $t_0$, Adam’s mental event immediately causes Bob’s colleague’s bike to begin to move at $t_1$. Perhaps when Adam wills to raise his arm at $t_0$, Adam’s mental event immediately causes Bob’s arm to rise instead, while Adam’s arm remains inert at his side.

\(^{225}\) Thinking modes of type X are introduced immediately below, in Pathological II.

\(^{226}\) I will try to remain somewhat neutral on what the causal relata are here. So, for example, it might be that Adam himself, as an agent, causes the arm movement at the later time, or it might be that some mental event or mental state causes the later physical event or state. I will speak mostly in terms of events for convenience, and because the contemporary literature tends to do so, but keep this qualification in mind.
Pathological II: Whenever any thinking mode at t₀ has an immediate causal antecedent that’s a thinking mode of type X, then at t₁ some pathological causal interaction occurs. Perhaps when Adam wills to raise his arm at t₀, Adam’s mental event immediately causes Bob’s colleague’s bike to begin to move at t₁. Perhaps when Adam wills to raise his arm at t₀, Adam’s mental event immediately causes Bob’s arm to rise instead, while Adam’s arm remains inert at his side.

The only difference between Pathological I and II, of course, concerns the thinking mode’s immediate causal antecedent—some unknown attribute mode in I, and some thinking mode of a given type, in II. Providing both cases shows that the problem arises regardless of whether one endorses Attribute Plentitude or merely Attribute Plurality, since Pathological II requires only two attributes for its development.

There are several notable features of these perhaps silly-seeming cases. First, there are no violations of the PSR here. Every action that occurs at t₁ in any of the three cases has a sufficient reason, its immediate causal antecedent at t₀ as part of a whole chain of causes we might envision as 1) necessary and 2) stretching back infinitely far in time. Similarly, there is no violation of Naturalism. Each event at t₁ occurs in accord with universal, exceptionless laws, specified in part with reference to the type of mode (of attribute 3, or mental but of type X). Nevertheless, the events in Pathological I-II are highly dissonant with our ordinary experiences, and might even be presumed impossible. This might not be so surprising, however, for if attribute 3 modes, or thinking modes of type X, rarely precede particular volitions in the causal chains of human actions, we
might be acquainted with them only rarely. And again, while Plentitude is required for Pathological I, Pathological II need not assume Plentitude, which shows that pathological causation threatens even if we restrict ourselves to dualism.

We might try to marshal our thoughts on such cases into an argument as follows:

1. **Not EB**, for reductio: That is, it is not the case that every attribute and its modes is causally isolated from every other.

2. **PSR**: Every fact, state of affairs, event, or being has a cause or explanation for its truth or existence, if it obtains or exists, and a cause or explanation for its falsity or nonexistence, if it does not.

3. **Plentitude**: There are infinitely many modes of infinitely many attributes.

4. **Naturalism**: Universal and exceptionless laws describe the nature and behavior of every mode.

---

227 Strawson, “Real Materialism (with new postscript),” has a great aside, in fn. 91, along precisely the lines I’m thinking of here, which he connects to views that require some sort of dualism for theoretical reasons (as Spinoza’s seems to require attribute pluralism, even in the context of substance monism). He says: “There may be phenomena in the universe that can’t interact causally with other things given their nature (rather than their position in spacetime) or that do so only on the first Thursday of every seventh century, in a highly peculiar way.” Perhaps attribute 3 modes directly preceding mental acts in causal chains is an every seventh century sort of phenomenon, or a phenomenon that only occurs in regions of spacetime radically separated from ours—certainly, there is no empirical evidence to rule this out. And lest one think the example too quixotic to be believable, consider the physicists’ genuine search for magnetic monopoles, dark matter and energy, or for gravitational waves, the latter of which are only detectable after certain rare, particularly energetic astronomical collisions occur.

228 Note that this is not, strictly, how EB has been defined thus far. The EB requires causal and conceptual isolation, but for ease of exposition, and because it better accords with the way the pairing problem is discussed in the contemporary literature, I’ll stick with causation here. Nonetheless, this is not to be interpreted as any betrayal of Spinoza’s views, since I endorsed, in chapter 1, the position Newlands calls CDM: all dependence relations (i.e. causation, conceiving through, inherence) are the same dependence relation. Thus I’m choosing simply to talk about one relation, with three names, in one consistent, causal, way.

229 Similarly to the footnote above, the PSR will be stated in causal and explanatory language, but the relevant guarantee for this statement of the argument is that every mode has a cause (which Spinoza takes to explain its effect).
5. **Anti-skepticism:** At least in many ordinary cases, we know the explanation the PSR guarantees.

6. Without EB, at least some universal exceptionless laws hold between causes and effects from different attributes [Not EB, Naturalism].

7. Some physical event ‘A’ occurs, where A names some ordinary physical event where Not EB holds, as for example Adam’s arm rising at t1 [from experience].

8. A has a cause or explanation [PSR, 7].

9. The immediate cause or explanation of A is some mental mode C or some mode of attribute 3,4,5 etc. governed by a universal law [Not EB, 8, Plentitude, Naturalism].

10. **Ignorance:** Human beings lack all knowledge of modes, including anything about such modes intrinsic natures, typical causal antecedents or effects, or even whether they are generally causally ordered at all, besides those of Thought and Extension [fact of experience, though Spinoza can provide theoretical justification with the EB in hand].

11. If the immediate cause C of event A at t₁ is some mode of an attribute other than Thought or Extension, then we do not know anything specific about C’s intrinsic nature, typical antecedent causes or effects, or whether C participates in a generally causally ordered structure at all [9, Ignorance].

12. **Discrimination:** Without some sufficient reason for thinking that cause C of A (Adam’s arm rising) should be restricted to modes of Thought, the
probability that C is Thinking is either 0 or infinitely small.

13. **No Restriction:** Given the cross-attribute causation in 7, there is no sufficient reason for restricting C to modes of Thought.

14. The probability that C is Thinking is either 0 or infinitely small [12, 13].

15. The probability that we do not know anything about C [intrinsic nature, typical causal antecedents or effects, whether C participates in a causally ordered structure at all] is either 1 or arbitrarily close to 1 [11, Not EB, 14].

16. **Explanatory Constraint:** If the probability that we do not know anything about the nature of the cause of some event E, its typical causal antecedents or effects, or whether it participates in a causally ordered structure at all, is 1 or arbitrarily close to 1, we cannot explain E.

17. We cannot explain A, Adam’s arm rising at t₁ [15, 16].

18. But we can explain A, for a contradiction [17, Anti-Skepticism].

19. EB is true.

Of course, this argument strictly shows only that there is at least one event A that we cannot explain, though our anti-skeptical assumption assures that we can know the explanation of A that the PSR guarantees, even if we cannot know the complete or ultimate explanation for A. So strictly only one contradiction has been shown, which might not seem to threaten anti-skepticism generally. But since A was chosen wholly arbitrarily, there will be many violations of our anti-skeptical assumption, on the assumption that Not-EB is true. So if the argument’s other independent premises are true, then Not-EB is false.
This version of Spinoza’s argument is ‘progressive’ in the sense that has often been applied to one reading of Kant’s transcendental deduction, and to which Karl Ameriks opposes a ‘regressive’ reading. Roughly, a progressive reading of the transcendental deduction starts from premises a skeptic will accept and seeks to provide a proof of objectivity and empirical knowledge, while a regressive argument assumes some empirical knowledge and seeks to show its necessary preconditions. Relatedly, in this first reading of Spinoza’s argument, the assumption shared by both Spinoza and his opponents, that we can somehow explain A, is used to generate a contradiction, which Spinoza attributes to interactionism, or Not EB. Admittedly, this shared presupposition is much more substantive than that of other famous progressive arguments—i.e. P.F. Strawson’s formulation of the reidentifiability of objects premise, for example, in Kant. But this is at least partially because Spinoza and his opponents are mutually committed to anti-skepticism. Nonetheless, the moves are structurally similar. This progressive reading sees Spinoza as moving from anti-skeptical presuppositions he shares with his interlocutors to provide a proof of the truth of EB.

3.4.3 Challenges for the Progressive Version of the Argument

I think this line of reasoning captures many of the intuitions driving Spinoza’s imagined objector in the Short Treatise, but in its present form I do not think it is overwhelmingly compelling. On which of the independent premises, given in bold, should Spinoza’s interactionist opponent focus her attention? Not-EB is just the thesis of

230 Ameriks, “Kant’s Transcendental Deduction as a Regressive Argument,” 273

231 Strawson, The Bounds of Sense
interactionism, assumed for the sake of argument. Rejecting Naturalism is not a tempting move, even if most contemporary naturalists, following physicists, believe that universal probabilistic laws should replace universal exceptionless ones. And it might seem that reformulating the problem with probabilistic laws would make the explanation for A even less secure than before.

Another obvious strategy rejects the PSR. But, as should be clear from considering the argument, the full-blown PSR actually does very little work here. Getting the conclusion that A, Adam’s hand rising, has some explanation obviously requires something much weaker than the full PSR, since even philosophers who reject the principle want some explanation for ordinary events like A. There may be serious difficulties, as we saw last chapter, with drawing the line between places where explanations succeed and where they run out, but we needn’t confront them at this stage. As long as we are unhappy making A, plus a whole range of ordinary physical happenings, brute, a much weaker explanatory demand suffices. Anti-skepticism then simply guarantees that, at least in many cases where this weaker explanatory demand is met, we know the relevant explanations, even if our explanatory knowledge is not complete or total.

The PSR might be seen as operating in the background of No Restriction. It suggests that, given cross attribute causation (i.e. mental-physical causation), there’s no sufficient reason for thinking that C is a mode of any attribute in particular, and thus no sufficient reason for restricting C to modes of Thought. But No Restriction might also be seen as a version of Humean Recombination, whereby, prior to experience, causes and effects might be paired arbitrarily. As Hume states: “And as the first imagination or
invention of a particular effect, in all natural operations, is arbitrary, where we consult not experience; so must we also esteem the supposed tie or connection between the cause and effect, which binds them together, and renders it impossible that any other effect could result from the operation of that cause.”

Given cross-attribute causation between items with radically different natures, and pairwise arbitrary recombination possibilities *a priori*, it seems equally possible for cause C of A to be from any attribute. Of course, our experience seems to drastically reduce the scope of possibilities compatible with the actual laws of nature, but given *Privilege* and *Uniqueness Violations*, as well as *Temporal Oscillations*, it’s as yet unclear what weight we should attach to this. For those cases suggested, first, that some agents do not experience the sort of control stipulated in Uniqueness and Privilege, and that symmetry considerations should cause us to take others’ experiential evidence seriously, and thus modify our inductive base, if no direct appeal to experience is being made. Or, second, and if this first line of reasoning is rejected, Uniqueness Violation cases can be experienced by ordinary agents willing to seek out the right sorts of experiences, and can thus not be dismissed in the way that Privilege Violations might be. This all suggests that the PSR need not motivate No Restriction, and thus that resistance to the full-blown PSR does not necessarily resolve the issue.

**Discrimination** then claims that, if it’s possible for C to come from any attribute, and if there are infinitely many attributes, then the probability that C is thinking is infinitely small or zero. (We have already supposed that C is not extended, by presuming Not-EB.) The PSR might seem to operate here by suggesting that the absence of some

---

232 Hume, *Enquiries*, 29
intelligible reason for restricting C to Thought, in light of inexplicable mind-body causal
interactions between fundamentally unlike things, there is no such restriction. Indeed, this
appears to be a weak point in the argument, which we should consider alongside a way of
rejecting Plentitude.

In the contemporary philosophical environment, one natural way of rejecting
Plentitude is endorsing physicalism. I will postpone discussion of physicalism until
chapter five, however, since I think Spinoza can offer compelling arguments against the
physicalist challenge. However, these arguments differ from those considered here;
considering them now would take us too far afield. One might also reject Plentitude in
favor of just dualism. It does not appear that this alone will work, since Pathological II
appears capable of generating the worry without employing Plentitude. Of course, this
would require modifying Ignorance to include the same or at least comparable ignorance
of one part of the causal network of Thought—that concerning effects of “thinking modes
of type X.” Debating the plausibility of this expansion of Ignorance is beside the point for
the version I consider next, however, so I will not go into the details of it here.

Overall, the further thought behind the objection developing here is that, once we
restrict our focus to Thought and Extension, we gain much more insight into explanations
for cross attribute causation than the argument allows. For instance, if Uniqueness and
Privilege are true, then we can know directly that no other mind’s willing caused Adam’s
hand to rise spontaneously at t₁ (by Privilege). Further, Adam’s willing at t₀ could not
have caused some other body, besides his own, to move at t₁ (by Uniqueness). Of course,
Adam’s arm might rise if I push it, but this would be some body-body interaction (my
hand→his arm→his mind) that we could account for straightforwardly.
We have wandered into a nest of problems for this version of the argument. Spinoza’s interactionist opponent should frame her response as a rough dilemma: either explaining why pathological causation does not occur is not a serious challenge, or it gets no easier when one adds the explanatory barrier between the attributes. Therefore, the problem of pathological causation remains even with the EB, and so avoiding it cannot be the motivation for introducing EB in the first place.

More specifically, one way of characterizing the major objection seems to run as follows. There are not only infinitely many modes of different attributes, but there are infinitely many modes of each single attribute (infinite Extended modes, infinite Thinking modes, etc.). So there are versions of Discrimination and No Restriction that hold within a single attribute, like Thought (or Extension). Given that there are infinitely many thinking (or extended) modes, the probability of identifying the correct thinking (or extended) cause C, or the correct class of possible causes that includes C, of A is zero or arbitrarily small (Discrimination*). This is because there is no sufficient reason for restricting C to tm1. That is, there is no sufficient reason for restricting the cause of Adam’s arm rising to Adam’s willing to raise it, or even to Adam’s mental states, at all (No Restriction*).

This seeming parity in the probability calculus within one attribute with infinite modes, as opposed to between infinite attributes of infinitely many modes, suggests that the real problem is not the sheer number of candidates for cause C of A, but something else. This is one aspect of the horn of the dilemma that claims that the EB makes things no easier, and thus that these pathological causation cases cannot motivate its introduction. The objection is that, if numerical infinitude were the real issue, then the EB
would make things no better, since the same cardinality of problem remains within the single attribute of Extension (i.e. remains on a wholly physicalist, not to mention Spinozist, picture).

In response, I should note two things. First and most importantly, I think the interactionist objector is right that, for the most part, the sheer number of candidates for \( C \) is not central to the problem. But second, it can plausibly be shown that the problem is strictly, mathematically worse for the interactionist, even if it is debatable whether this greater magnitude is one that makes an epistemic difference. Alexander Pruss, in arguing for the PSR, gives an argument that if we allow for PSR violations to be metaphysically possible, then such violations are overwhelmingly (infinitely) more likely than non-violations. In a structurally similar way, we can argue that there are infinitely more pathological causal scenarios than ordinary ones in a one-attribute (i.e. purely physical) case, and second, that there are infinitely many more than this already infinitely many pathological causal scenarios than ordinary ones in a two-attribute (i.e. interactionist dualist) case than in a one-attribute case.\(^{233}\)

\[^{233}\text{Pruss, The Principle of Sufficient Reason, 267-270, gives his argument for the PSR, rather than for the claim I defend here, that interactionism makes the problem of pathological causation strictly worse, in a mathematically demonstrable way. So some details of my presentation naturally differ, though mostly it follows his quite closely.}\]

Define a \textit{regular universe} as a maximal aggregate of physical entities that obeys our laws of physics where causation is ordinary.\(^{233}\) Let \( S \) be the set of regular universes, which is infinite, but is nonetheless, as Pruss notes, a set of fixed cardinality, unlike the set of all possible universes.\(^{233}\) “Now assume the Axiom of Choice. Let \( a \) denote the cardinality of \( S \). Let \( n \) be any cardinality greater than \( a \), for example, \( 2^a \) (the cardinality of the set of all subsets of \( S \), which Cantor’s diagonal argument has shown to be greater than the cardinality of \( S \)).” We will let \( U \) be a set of possible universes in which the effect of one of Adam’s brain events (remember, we are in the purely physical case) is pathological rather than ordinary (i.e. his arm doesn’t rise as normal). Specifically, it is pathological in that, after the relevant brain event, rather than Adam’s arm rising, “the Earth and Sun disappear…and are replaced by a big cloud of photons all dancing the polka, with the number of these photons having some value between \( N_0 \) (the cardinality of the set of integers), inclusive, and the \( (n+1) \)st infinite cardinal number \( N_n \) exclusive, which photons then disappear, the Earth and Sun return, and everything returns to normal nomic order the day after. The number of different cardinalities \( m \) satisfying \( N_0 \leq m < N_n \) is equal to \( n \), and so we can choose the
Still, I think the more fundamental idea here is that the size of the class of candidates only seems worrisome if the identification of C is a \textit{random search} through the space of possible modes (either of one attribute, or of infinitely many). Said differently, identifying C in a principled way might seem fruitless if there are no features that make certain modes more likely candidates for C than others. But, Spinoza’s interactionist dualist opponent might suggest, if there are only two possible attributes, Thought and Extension, and if we can reliably sort causal interactions into 1) purely extended causation, 2) mind-body causation, and 3) purely mental causation, we are set. If we have a purely extended or purely mental causation interaction, the explanation will be some immediate extended or mental causal antecedent, respectively. And for mind-body causation, Uniqueness and Privilege vouchsafe the explanations required in those instances, at least to some extent. This would be a way of defending the other horn of the worlds in U so that U will have cardinality at least n and hence greater than a. Moreover, since n is an infinite cardinality, it must be that n \textit{is infinitely many times} greater than a.” This establishes our first conclusion, which is that pathological scenarios are infinitely more likely than ordinary scenarios if we allow any pathological violations at all.

The second conclusion is that there are even more pathological scenarios in the two-attribute (interactionist) case than in the prior, purely physical one. This can be shown in a variety of similarly goofy ways. Here is one. First, again let S be the set of universes where causation is ordinary, and so where Adam’s willing to raise his arm causes his arm to rise (note that we are now in the interactionist case, where willings in Thought cause bodily movements in Extension). Let U be a set of possible universes in which the cause of an apparently ordinary arm raising comes, not from Adam’s willing, but from some other mind’s willing. Let the cardinality of these other possible intruding minds be defined as in the case above. Then by the same reasoning we can show that the cardinality of such pathological cases is infinitely greater than a, the cardinality of S. But now \textit{for each of the pathological cases} just considered (i.e. intruding mind m1 causes Adam’s arm raising, intruding mind m2 causes Adam’s arm raising…), we can let the set of universes where that particular pathological event occurs be S, with a cardinality a as before. And we can then add to the pathological cause (the intruding mind, rather than Adam’s mind) a pathological effect (i.e. some weird cloud of photons dancing the polka, ala Pruss, or something more reasonable, like Adam’s arm raising in some m number of different ways, with the ways m such that No \text{m} \leq m \leq Nn as in Pruss’s argument above. Since the ordinary case is infinitely swamped by the set of pathological scenarios in Pruss’s original argument, and since every single pathological scenario considered there can itself be infinitely swamped in precisely the same way with a pathological effect, the problem is strictly worse for the interactionist. More generally, for any physical event and any set of purely physical pathological causes or effects, the interactionist faces a Pruss-style argument where every possible physical pathological cause or effect can then be paired with a mental cause or effect of a given cardinality as defined by Pruss. And each successive addition of a further attribute makes the problem strictly worse.
dilemma. Our experience would provide good inductive evidence in both sorts of cases for restricting possible causes to the “convenient” sets of thinking and extended modes. For instance, in seeking to explain Adam’s arm rising, only Adam’s mental states or bodily states would be considered as candidate causes. Explaining pathological causation’s non-occurrence would be revealed as easier than it initially appears.

This gets us closer to the argument’s fulcrum, I think, but we are not quite there yet. For one, notice that the interactionist is not yet out of trouble. Interactionists seem to require three further assumptions. First, experience generally, and Uniqueness and Privilege in particular, greatly assist the interactionist theory to restrict the space of possible candidate causes. So we must have some good reason for continuing to rely on these principles in light of the counterexamples to them proposed earlier.

Second, there must either be straightforward pairing relations between causes and effects in the purely physical and purely mental cases, or there must be some argument for why such pairing relations are not required to secure purely physical or purely mental causal explanations. In the next chapter, I discuss these two possibilities in greater detail, but I will preview them here. Pairing actualists think that we can give uniquely specifying pairing relations, either in the physical domain alone, or in both the solely physical and mental-physical arenas, which join causes and effects. These relations are not just brute causal relations—for my purposes positing that sort of relation counts as a version of Pairing Nihilism. Pairing nihilists deny that there are such pairing relations in either the solely physical or the mental-physical domains. They might argue that since we cannot identify relations pairing even physical causes with physical effects in a unique way, the demand that dualists have special obligations to provide mental-physical pairing
relations is similarly spurious. Pairing nihilists must then suggest why their position nonetheless fails to endanger our ordinary explanations in solely physical cases. Finally, interactionists must either explain the truth of Uniqueness and Privilege, given their own metaphysical commitments, or they must argue that they need not provide any such account.

Third, there must be a method for investigating psychophysical (i.e. mental-physical) causes and giving psychophysical explanations. Or, there must be some argument for why a method for investigation and explanation generation is not required at all (and thus not required either for purely physical or psychophysical causal relationships). This argument for No Methodology, as I have been repeatedly emphasizing this chapter, must be compatible with explaining the shared anti-skeptical position all parties we are considering would like to occupy. But it seems that the presence of some method is essential. Methods make the search for candidate causes relevantly disanalogous to a random search of the infinite modes of Thought and Extension. By ameliorating the difficulties associated with the either intra-attribute or inter-attribute mode infinitude, a method for determining particular causes and giving explanations of causal phenomena shifts our focus from the number of candidates to the way they are identified.

Note that the demand for some method for investigating psychophysical causation is consistent with our not employing it in all, or even very many, of our ordinary psychophysical causal ascriptions. We often ascribe mental causes to our own physical movements, of course, without consulting anything but our own thoughts. For example, I claim that I ambled towards the fridge because I had been thinking about the beer all day.
I raised my hand because I’d be thinking about a question to ask. We do this for others too, often operating with a folk psychological theory—discussed at length in the next chapter—that allows us to bypass a more methodological search for causes in favor of easy generalizations from past experience.

But recall that cases of pathological causation can in principle arise between any two modes of different attributes. So, especially because we must deal with fairly common human experiences, like depression and anxiety, as well as less common but still prevalent ones, like the Privilege or Uniqueness violations discussed earlier, we will often need to appeal to some method for clarifying these difficult cases. For instance, putative violations of Privilege, as occur in thought insertion for schizophrenics, will be explained by pointing to scientifically ascertained causes of the schizophrenic’s symptoms, in terms of unbalanced neurotransmitter levels. Such explanations are senseless unless we presume a method for arriving at them in a principled way, for they are far too complicated to be arrived at merely on the basis of common experience. This is true whether such experience is conceived either as the direct experience of the control stipulated by the principles or as providing material for an induction. Similarly, Uniqueness Violations are not science fiction, but fact, it seems. Still, the presence of a clear method for explaining why Uniqueness is rarely violated in the ordinary case is important. Further explaining what conditions make possible its violation, in the extraordinary ones, prohibits easy inferences from actual scientific advances to the prospects of much more far-fetched Uniqueness Violations, such as telekinesis. Thus, even if the method for investigating psychophysical explanations is only rarely employed, we seem to have compelling reasons for assuming that it is necessary to justify at least
some of our causal ascriptions and to prohibit other causal ascriptions that are impossible, given nature’s actual laws. 234

In summary, Spinoza thus needs to do one of two things. He must first show how his system can provide the necessary pairing relations, while the interactionist dualist 1) cannot give the pairing relations or 2) cannot defend their superfluity while retaining anti-skepticism compatible with their basic metaphysical commitments. Or, Spinoza must show why his alternative to interactionism can provide a method, which interactionists cannot, for generating reliable explanations of mental and physical phenomena. Spinoza’s methods must be compatible with ordinary, but not pathological, causation, and must be such that the dualist cannot endorse them, even though she offers no feasible alternative. Finally, if Spinoza can resist other alternatives to both interactionism and his view (i.e. physicalism as an alternative to both), as I’ll argue in chapter five, then he has crafted a powerful contender in the metaphysics of mind that deserves serious consideration.

3.4.4 The Canonical Regressive Version of Spinoza’s Argument for the EB

With these three metrics in mind—Uniqueness and Privilege in particular, the pairing relation problem in general, and the question of method—we can make several

234 Next chapter, I discuss some plausible interactionist dualist candidates for answering the method question: psychophysical laws and folk psychological explanations. One way of seeing how these candidates are generated is the following. Along one axis, it seems like the explanations called for here could be law-like, or explanatory but not sufficiently strong so as to be law-like (i.e. perhaps because they cannot be formulated in exception-less ways or cannot be given principled ceteris paribus clauses). Along another axis, it seems like the candidate explanations could be commonsensical, or non-commonsensical (i.e. because they are complex explanations given by the natural sciences). Psychophysical laws would then be non-commonsensical, strict laws, whereas folk psychological explanations would be non-lawlike, commonsensical explanations. Of course, in this taxonomy strictly lawlike commonsense explanations (law-governed folk psychology) or non-lawlike scientific explanations are both possible. But I don’t find either alternative even remotely plausible.
points. Spinoza’s disagreement with the physicalist differs greatly from his criticisms of the interactionist dualist. That’s because Spinoza can accept any answer the physicalist adopts on the question of pairing actualism vs. pairing nihilism, provided that the answer is compatible with the existence of a reliable method for generating the relevant observations. Moreover, in the wholly physical realm—which the physicalist simply asserts is exhaustive—both Spinoza and the physicalist can appeal to the same method, that of the natural sciences. This too obviates a possible conflict. The first fundamental point of tension between Spinoza and the physicalist concerns Spinoza’s assertion that there is more than a single fundamental way of being, or attribute, in reality, though we will see more areas of disagreement in chapter five.

With the dualists, by contrast, Spinoza has two points of argumentative entry. First, he can argue that Nihilism about the pairing relations is incompatible with the existence of a reliable method for generating psychophysical explanations. He can add that dualists cannot be pairing relation actualists because the relevant relations are unavailable. Or, second, Spinoza can argue directly that there is no reliable method for generating psychophysical explanations. I think this sort of claim will have two parts, the first of which denies the existence of psychophysical laws. The second claims that commonsense or folk psychology, which offers the nearest approximation to law-governed psychophysics, lacks features required of a theory with a reliable methodology for explanation generation.
The canonical, regressive version of Spinoza’s argument thus mostly leaves aside the physicalist and focuses its attention on the interactionist dualist. It is regressive in perhaps an even weaker sense than Ameriks’ characterization of Kant’s ambitions. By the conclusion of Spinoza’s regressive argument, I will not mean that EB is the uniquely necessary precondition for causation’s being ordinary, rather than pathological. Rather, given Spinoza’s other assumptions, or plausibly weakened versions of them, the EB provides the most plausible precondition, relative to its rivals, for avoiding pathological causation. This merely rephrases the point made several times before: the isolationist claims that his system can explain something important that the interactionist cannot. Arguably, Spinoza wants the stronger regressive claim—that the EB is an absolutely necessary precondition for the truth of causation’s being ordinary rather than pathological. But I’ll be satisfied here if it is substantially more plausible than its rivals.

1. **Interactionism, or Not EB**, for reductio: That is, it is not the case that every attribute and its modes is causally isolated from every other.

2. **PSR**: Every fact, state of affairs, event, or being has a cause or explanation for its truth or existence, if it obtains or exists, and a cause or explanation for its falsity or nonexistence, if it does not.

---

235 Recall, however, my earlier assertion that those physicalists who believe that mental and physical properties are distinct might be vulnerable to this canonical regressive formulation too, phrased in terms of properties rather than events. I discuss a version of this worry in chapter five, though in slightly different terms.

236 This is not to say, of course, that Spinoza’s own argument against other metaphysical systems was taken to be this weak in his mind. But since I do not aim to defend the PSR here, nor do I show how one might reliably evaluate whether some explanation is good enough to satisfy the demands of the PSR, I cannot defend the stronger version that plausibly convinced Spinoza. But again, one should not confuse the absence of explicit defense here with evidence that such a defense cannot be made.
3. **Plurality**: There are two attributes, Thought and Extension, with infinitely many modes of each.

4. **Naturalism**: Universal and exceptionless laws describe the nature and behavior of every mode.

5. **Anti-skepticism**: At least in many ordinary cases, we know the explanation the PSR guarantees.

6. Physical events like A occur, where A names some physical event like Adam’s arm rising at t1, and A has an immediate naturalistic sufficient reason for its occurrence that is a mode of Thought [Experience, Not EB, Plurality, Naturalism].

7. If Anti-Skepticism is true, when A occurs there is some explanation, which we can know, for why A occurs as in Ordinary Causation, rather than as in Pathological Causation [6, PSR, Anti-Skepticism].

8. So there is some explanation for why A occurs as in Ordinary rather than Pathological Causation [Anti-Skepticism, 7].

9. Necessarily, if there is some explanation for why A occurs as in Ordinary rather than Pathological Causation, there is a reliable method for

---

237 To refresh: **Ordinary**: Whenever neither the direct causal antecedents of tm1 nor those of tm2 are modes of some unknown third attribute, ‘ordinary’ causation occurs. This is the sort with which, we naturally say, we are most familiar. Adam wills to raise his arm at t0 (tm1 is located at t0), and Adam’s willing causes Adam’s arm (not Bob’s) to move at t1. Similarly, Bob wills to raise his arm at t0, and Bob’s willing causes Bob’s arm (not Adam’s) to move at t1.

The **Pathological Causation** case given previously is not exactly parallel, but would go something like this. When thinking modes of type X precede physical events like A, then A is caused by radically different thinking modes than expected. For instance, rather than tm1 (Adam’s willing) causing A (Adam’s arm rising), perhaps some other agent’s willing, like Bob’s willing, causes A at t1. Perhaps some extremely temporally or spatially distant agent causes A at t1, like George Washington or an angry practitioner of voodoo. And, of course, if we had focused on the mental event rather than the physical event A, we would be focused on pathological effects. That is, a willing to raise one’s arm might move a bicycle or start a tornado.
generating psychophysical explanations [Generality of A, No-Method inconsistent with Anti-Skepticism].

10. There is a reliable method of generating psychophysical explanations [8,9].


12. So interactionism is false; interactionism’s rivals offer substantially better explanations of the sort that Anti-Skepticism guarantees we know for events like A [1, 10, 11].

I will devote chapter four to pursuing (11) as a Spinozistic argument against the dualist. But it’s worth making a few final points here. First, the direct attack on the dualist’s ability to avoid the isolationist criticism in terms of method is relatively independent of the pairing problem considerations, as they are discussed in the secondary literature. So while I pursue both prongs of the disjunction, strictly speaking the dualist’s methodological issues are more fundamental. Moreover, it will be important to try to understand what the demand for a method for generating reliable explanations entails. My argument certainly doesn’t require me to reject the use of all commonsense or folk psychological explanations—that would be absurdly contrary to ordinary epistemic

238 As an aside: the difference between an argument with Attribute Plurality (two attributes) and Plentitude (infinitely many attributes) is that the argument with Plentitude is stronger, not simply because there are more possible candidate causes, but more relevantly because of Ignorance: humans lack knowledge of all attributes other than Thought and Extension. For if, given our knowledge of our own mental states and those of others, we still lack psychophysical laws or a method for generating reliable folk psychological explanations, then trying to provide a reliable method for generating explanations with causal chains involve attributes of which we lack any knowledge at all is substantially more likely to fail.
Neither does it require that the method I see as reliable in purely physical cases—the method of the natural sciences—be infallible.

Rather, Spinoza’s argument here could be read as an inference to the best explanation—that the explanatory barrier between the attributes provides the most compelling explanation for our ability to explain why A is ordinarily rather than pathologically caused. But it is more than that. As I argue next chapter, the project of constructing reliable psychophysical laws or relying on commonsense psychological explanations for prototypically cross-attribute causation (i.e. mental states influencing actions) is in much worse shape than it appears. It’s not just that the explanations offered via the method of the natural sciences are more compelling, but that in many important cases psychophysical explanations are systemically misleading or false. We have already gotten a taste of this in our consideration of cases like schizophrenic thought insertion, or mental control of distant physical objects, both of which commonsense psychology seems to preclude. Otherwise, these phenomena would not be nearly so surprising as they in fact are. So while the version of Spinoza’s insight that serves as a successful IBE would secure the majority of my aims, my arguments in chapter four are aimed at establishing stronger conclusions. The interactionist’s claims are not simply less plausible than the Spinozist’s or even the physicalist’s—they are substantially less so. This is not only because ordinary psychophysical explanations are often systematically misleading or false, but also because more sophisticated ones are not always compatible with the

239 Strictly speaking, of course, folk psychological explanations across attributes (i.e. he walked to the fridge because he wanted a beer) will all be false, but there are nearby claims that are true, and the cross-attribute explanations can be pragmatically used without issue in many ordinary circumstances. I discuss these issues more next chapter.
underlying metaphysics of the interactionist. So while the interactionist might wish to appeal to the more sophisticated explanations available for complicated cases, she cannot explain how such explanations are appropriate, if interactionism is true. More modestly, she cannot do so in nearly as convincing a fashion as the isolationist.

It’s also important to recognize that the historical Descartes accepted Attribute Plenitude not Attribute Plurality, and thus might seem to face the more serious form of the preceding argument, as I discussed in a footnote earlier this chapter. Moreover, Descartes specifically admits that we cannot understand the mind-body union. He sees something very puzzling about interactionism, especially in light of the lingering skepticism he evinces even late in the Sixth Meditation about our freedom from error in particular cases of interaction. He says: “It is quite clear from all this that, notwithstanding the immense goodness of God, the nature of man as a combination of mind and body is such that it is bound to mislead him from time to time.”240 This puzzlement makes it highly unlikely that Descartes would accept that *commonsense psychology* has a reliable method for generating psychophysical explanations. The only Cartesian bulwark against widespread error is the existence of the benevolent God of traditional theism, one whose existence cannot be established if Spinoza’s EB-based argument for Substance Monism succeeds. So I think Descartes himself would find appeals to psychophysical laws or folk psychological methods for generating explanations highly unpersuasive as a response to Spinozistic pressure. Whether contemporary dualists know something that Descartes did not is an open question, but the foregoing suggests that the dualist might have to attack Spinoza’s notion of God, rather

---

240 *AT 2: 61*
than simply directly evaluating his position on mind-body interaction, in order to
succeed.

In this chapter I have tried to show how Spinoza’s *Short Treatise* serves as a
transitional text, one where Spinoza grapples with the promise and the puzzles of his
Cartesian legacy. Spinoza’s deep and early commitments to the intelligibility of a natural
world whose development follows universal laws must be supplemented to achieve his
mature metaphysical system. Against a shared backdrop of commitments like the PSR,
Naturalism, Attribute Plentitude, and Anti-Skepticism, Spinoza tries to resolve
outstanding riddles in his interactionist theory of mind-body relations. I argued that
Spinoza’s attempt to answer the question—“Why could it [the soul which has nothing in
common with the body] not also move wherever it will all other bodies that already have
motion? (*KV* I/97)—can be seen as leading him to the isolationist position he argues for
in the *Ethics*. I provided one way of understanding how this unanswered objection from
the *Short Treatise* might be turned into an argument to pressure Spinoza’s opponents to
accept an isolationist system more similar to his own, by appealing to isolationism’s
significant explanatory advantages.
CHAPTER 4:
SPINOZA CONFRONTS THE INTERACTIONIST DUALIST

4.1 The Current State of Play

In previous chapters, I argued that Spinoza’s explanatory barrier between the attributes, which states that each attribute and its modes are causally and conceptually isolated from every other, plays an ineliminable role in Spinoza’s metaphysics. However, most interpreters have seen no good Spinozistic justification for believing that what the EB claims is true. In order to rectify this, I sought to develop an objection from Spinoza’s Short Treatise into an argument for the EB. Building on the claims of the second chapter, this argument presupposes that both Spinoza and his natural interlocutors, dualists and physicalists, are anti-skeptics. The task, as I outline it, is to explain why the mental-physical causation we observe occurs “ordinarily” rather than “pathologically.”

Spinoza’s argument takes different forms for dualists and physicalists, respectively. In part, there are historical reasons for this. After all, physicalism and the atheism with which it was so often conjoined were minority views in Spinoza’s day, unlike our current state of affairs. Cartesian dualism, by contrast, was increasingly dominant, serving at least as a starting point or critical lightning rod for other developing philosophical systems. The main reasons for the differing strategies, however, are theoretical. The explanatory strengths and weaknesses of dualism and physicalism are
located, naturally enough, in different places. In this chapter, I focus on two Spinozistic lines of attack against the dualist.

The final version of the argument I reconstructed from the *Short Treatise*, with which last chapter concluded, begins from Spinozistic principles amenable to being weakened without losing their role in the argument. That is, while Spinoza endorses the PSR, Naturalism, and Attribute Plentitude, these principles can be replaced by 1) some demand that ordinary cases of mental-physical causation have law-governed explanations 2) in a context that presupposes only two attributes, Thought and Extension, rather than infinitely many. The argument then suggests that, on supposition shared by Spinoza and his interlocutors, that skepticism is false, there must be some explanation for ordinary events, like (A) Adam’s arm rising at t₁, in terms of either thinking or extended modes. But I argued that any compelling explanation for why A occurs ordinarily rather than pathologically requires some method for explanation generation, especially in borderline or novel cases where common experience offers little explanatory guidance. This chapter’s central claim is that Spinoza offers compelling reasons for thinking that the interactionist dualist lacks a method for generating the requisite explanation for why A occurs ordinarily, not pathologically.

I argue for this claim in two parts. In the first, found in 4.2, I discuss the causal pairing objection to dualism found in the contemporary philosophy of mind literature. I suggest interpreting it as a limited case of the problem of pathological causation that concerns Spinoza, and I then examine the premises of the following **Pairing Argument:**

1. Interactionist dualism implies nihilism about the causal pairing relations.
2. Causal pairing nihilism implies that there is No Method for generating
reliable explanations of mental-physical causation.

3. No Method implies that there is no compelling explanation for why A occurs ordinarily, not pathologically.

4. **Conclusion:** If interactionist dualism is true, there is no compelling explanation for why A occurs ordinarily, not pathologically.\(^{241}\)

Section 4.2.1 discusses interactionist dualists who try to avoid Pairing Nihilism by providing uniquely specifying pairing relations joining causes and effects in mental-physical causation cases, while 4.2.2 discusses interactionist dualists who assert that no additional relation, *beyond the causal relation itself*, is required to or capable of pairing causes and effects. This group thus accepts Pairing Nihilism while arguing that, since only the causal relation can play the pairing role in wholly physical causal contexts, physicalism and interactionist dualism reach parity here. On this line, the pairing problem is thus a decisive objection to both views, or, as its defenders favor, to neither.

In 4.3, the second part of this chapter, I suggest that, regardless of one’s stance on the causal pairing relations objection, there is a more fundamental obstacle to the dualist providing the requisite explanations for why causation is ordinary rather than pathological. This obstacle can be encapsulated in the following **FP or Psychophysics Argument:**

4. If interactionist dualism is true, then either there are psychophysical laws, or commonsense, folk psychology can reliably generate explanations for why causation is ordinary, not pathological.

---

\(^{241}\) Note that in the body of the text to follow, when I refer to premise 4 or just (4), I will be referring to the first premise of the second argument, rather than to the conclusion of the first argument.
5. **5a:** There are no psychophysical laws; **5b:** Folk psychology lacks the features required of a theory for generating reliable explanations.

6. If folk psychology lacks the features required of a theory for generating reliable explanations, then No-Method—that there is no compelling explanation for why A occurs ordinarily, not pathologically—is true.

7. **Conclusion:** If interactionist dualism is true, there is no compelling explanation for why A occurs ordinarily, not pathologically.

Of course, dualists will want to claim that, even if both lines of attack sketched here succeed, I have not exhausted the options for generating the requisite explanations in cases of inter-attribute causation. Others will want to argue that one or another step of the preceding arguments can be plausibly resisted.

These responses are expected, insofar as few philosophical arguments are irresistible. I simply reemphasize two points made at the end of last chapter. First, my arguments by no means employ some unrealistically high, infallibilist standard for generating reliable explanations. Furthermore, they do not require me to reject all folk psychological explanations, which would run seriously counter to ordinary epistemic practices. My argument is simply that when hard cases arise, such as those of thought insertion, putative Uniqueness violations, or other examples I’ll discuss later, the dualist has no plausible method for resolving them. Neither does my argument require that the methods for generating purely physical explanations—those pursued in the natural sciences—be infallible. *Rather, I simply claim that there is a significant asymmetry*

---

242 Therefore, do not mistake the deductive form of the Pairing Argument or the FP or Psychphysics Argument, which I adopt for ease of exposition, for a defense of a deductive argument for my conclusions here. I defend only the particular conclusions actually stated in the body of the text.
between the reliability of the methods employed within a single attribute—for example, the methods of the natural sciences within Extension—and those employed across attributes. This asymmetry grounds something like an inference to the best explanation in favor of isolationism, since Spinoza’s isolationism remains much more capable of explaining why causation is ordinary rather than pathological, at least (for purposes of this chapter) with respect to dualism.

4.2 The Causal Pairing Objection to Dualism

John Foster reignited the contemporary discussion of what has come to be known as the problem of causal pairings, which he took to be a potentially serious objection to interactionist dualism. Physicalists like Jaegwon Kim and Karen Bennett, and defenders of dualism like Geoffrey Madell, Paul Audi, and Peter Unger, have discussed it in this vein. But even in the contemporary literature, tight formulations of the problem are tricky, and philosophers rely on importantly different versions of the argument, which are all nonetheless responding to roughly the same basic intuition. So I want first to discuss a few formulations of the causal pairing objection, and to suggest how it might be interpreted as a limited case of the problem of pathological causation that concerned us last chapter. This will set the stage for the two-part discussion advertised above.

Jaegwon Kim gives a version of the problem of causal pairings that relies on mental-physical synchronization between two pairs of minds and bodies. With the

problem stated this way, Kim tries to show that the physicalist can use *spatial relations* to pair causes and effects to produce explanations superior to the dualist’s. I think that pure spatial relations alone are insufficient for answering the pairing problem, but that some other wholly physical relations might do so, and that this is compatible with what little Spinoza says on the matter. But my primary aim is not to endorse Kim’s particular spatial relations proposal, but rather his basic thought that there is a significant asymmetry between dualist and physicalist/isolationist explanations, to the dualist’s detriment.Keeping this aim in mind helps contextualize the pairing problem as important but ultimately best interpreted as complementing the material in the chapter’s second part.

Kim introduces the causal pairing problem with the following case:

Suppose that two persons, Smith and Jones, are ‘psychophysically synchronized,’ as it were, in such a way that each time Smith’s mind wills to raise his hand, Jones’s mind also wills to raise his (Jones’s) hand, and every time they will to raise their hands, their hands rise. There is a constant conjunction between Smith’s mind willing to raise a hand and Smith’s hand’s rising, and, similarly, between Jones’s mind willing to raise a hand and Jones’s hand going up. If you are a constant conjunctionist about causation, this would suffice for saying that a given instance of Smith’s willing to raise a hand is a cause of the subsequent rising of his hand, and similarly in the case of Jones. But there is a problem. For we see that instances of Smith’s mind’s willing to raise a hand are constantly conjoined not only with his hand’s rising, but also with Jones’s hand’s rising, and similarly, instances of Jones’s willing to raise a hand are constantly conjoined with Smith’s hand’s rising. So why is it not the case that Smith’s volition causes Jones’s hand to go up, and that Jones’s volition causes Smith’s hand to go up?  

244 In 4.3, I discuss this basic thought in arguing that dualists have in-principle difficulties generating reliable explanations, which the physicalist avoids even if she cannot provide candidates for uniquely specifying pairing relations. However, this alliance of physicalist and Spinozist concerns only the cases here. I am not arguing for physicalism or suggesting that Spinozism is a version of physicalism. As we will see next chapter, physicalism faces other challenges that Spinozism more competently meets, which gives Spinozism an advantage over both views. But these challenges are different than those that concern me in this chapter.

245 Kim, *Physicalism or Something Near Enough*, 76
I underline part of Kim’s example to ward off my first red herring—dualists don’t dodge the problem by rejecting Humean explanations of causation as constant conjunction, or by adopting counterfactual dependence or any other account of non-primitive causality. Kim makes this point clear in his later discussion, but it’s worth stating up front. The causal pairing problem, in the contemporary discussion, is supposed to arise mostly independently of one’s favored view of causation.

Karen Bennett nicely emphasizes this point in her much cleaner formulation of the problem, which she states as follows: “The Cartesian dualist should think that it is metaphysically possible for there to be two intrinsically indiscernible Cartesian minds, such that a) they simultaneously perform indiscernible mental acts m1 and m2 and b) m1 causes some particular physical effect p that m2 does not cause.”

After explaining how spatial relations tidily pair m1 rather than m2 with p—throwing a rock in some spatial circumstances causes a window to break, while in others (i.e. on an empty beach) it does not—she notes the problem’s independence from specific views of causation. She says:

Notice that simply appealing to something like Humean constant conjunction, subsuming laws, or counterfactual dependence does not really answer the question…If m1 and p fall under types that are constantly conjoined, or if p counterfactually depends on m1, we can always ask why—and our answers are tightly constrained by the fact that m2 and p are not so connected…The point of the pairing problem is that the pairing relation does not supervene on intrinsic properties plus spatial relations.

Bennett’s observations are broadly correct, assuming one accepts the need for an additional relation, beyond the causal relation itself, to do the pairing. This assumption is

---

246 Bennett, “Mental Causation,” 320

247 Bennett, “Mental Causation,” 320
at least somewhat justified by the ascendancy of broadly Humean accounts of causation, but I’ll discuss those who reject it in 4.2.2. For now, I simply note that in disambiguating Kim’s initial formulation—by arguing that the problem is not restricted to Humean or any other specific causal views—Bennett introduces an unnecessary presupposition worth flagging.

Bennett supposes that the pairing problem concerns the inexplicable supervenience failure of the pairing relation on intrinsic properties plus spatial relations. This is a common assumption in the literature. Robb and Heil, for instance, suggest: “This is not the epistemological question of how we could know these are the pairings (although this is troublesome too). The question, rather, is metaphysical: in virtue of what are these the pairings…If minds are non-spatial souls, relative spatial location is unavailable to fill the pairing role. And since [mind 1] and [mind 2] are, by hypothesis, exactly similar, we cannot appeal to the different intrinsic properties that they might possess.”248 Similarly, Paul Audi remarks:

Two presuppositions of the argument require mention. The first is the possibility of qualitatively indiscernible things…Causes are always paired in a unique way with their effects, but it is only in cases of qualitatively indiscernible subjects that the demand for a non-qualitative basis for pairing is irresistible, simply because there is ex hypothesi no qualitative difference on which causal differences could rest…Second, the pairing problem arises only if causal relations cannot themselves be pairing relations. Another way of putting the point is that the pairing problem seems motivated by the presupposition that two things cannot brutally stand in a causal relation.249

---

248 Robb and Heil, “Mental Causation,” 8
249 Audi, “Primitive Causal Relations and the Pairing Problem,” 4-5

213
Leaving aside the second issue of brute or primitive causal pairings that I’ll return to shortly, there are two major issues with interpreting the pairing problem in this way, at least for my purposes. First, Spinoza, along with anyone else who accepts the PSR and the Principle of Identity of Indiscernibles (PII) that it entails, rejects the metaphysical possibility of qualitatively indiscernible things. This rules out, right at the beginning, any chance of finding a version of the pairing problem in Spinoza, if this assumption really is necessary. Crucially, Spinoza rejects this possibility independent of his commitment to EB—this is one of several places where it’s vital that the EB and PSR play complementary roles in his metaphysical system.

Second, qualitative indiscernibility represents an unnecessary addition to the intuitive presentation of the case that Kim, for example, gives us. Return to Kim’s statement of the problem: Smith and Jones are psychologically synchronized merely insofar as their acts of willing occur simultaneously, as do their arm raisings, making pairing difficult. But Kim did not say, nor does he need to say, that if Smith roots for the Cubs then Jones does too, or that Smith must be thinking of raising his arm to relieve an itch if and only if Jones is also. We have been no reason for thinking that Smith and Jones are mentally similar in any general sense, apart from their locally similar intentions vis a vis repeated arm raisings. So that Smith and Jones need to be qualitatively identical, mentally speaking, has played no part in Kim’s presentation of the case. The problem remains perfectly intelligible despite any such qualifications, precisely because, pace Robb and Heil, the ontological problem springs from the recognition of a serious explanatory problem.
Starting with the explanatory roots of the problem allows us to see how the causal pairing problem is a limited case of the threat of pathological causation. Return to Bennett’s formulation for a moment. For any two non-spatial minds, m1 and m2, which are similar in some relevant respect (i.e. both intending to raise an arm or perform some other action P), we can ask why m1 rather than m2 is causally related to P. Since both m1 and m2 are non-spatial, and thus (plausibly) do not stand in any spatial relations, we cannot use these relations to specify P’s unique causal antecedent. This might initially seem like a merely epistemic limitation, but upon reflection it’s difficult to see what sort of relation could explain the link between non-spatial mind and spatially located action, like an arm rising, while respecting in a principled way that m1 but not m2 causes p. This “how” question echoes Princess Elizabeth’s initial objections to Cartesian dualism.

But we can also ask a “when” question, by varying the temporal properties of each mental act too, especially since it seems natural to think of spatiotemporal, rather than solely spatial, properties as doing the pairing work, pace Kim’s focus on spatial relations alone. So m1 and m2 might be mental acts that occur at widely separated points in time, and we might ask whether one but not the other causes P. Only with the additional assumption that all causes are immediately temporally prior to their effects, which seems false for many reasons (i.e. smoking causes cancer but smoking is not immediately temporally prior to this effect, to mention an example that deserves more discussion), would we be able to rule out one of the candidate causes of P (i.e. m1 or m2).

Thus only several additional conservative assumptions, operating in the background of both Bennett and Kim’s arguments, restrict us to interpreting the pairing problem as a compelling but limited problem for dualism, as opposed to a general
explanatory puzzle. These conservative assumptions include at least the temporal proximity presupposition for causes and effects just mentioned, but also probably something like Uniqueness and Privilege. Those principles, as we saw last chapter, restrict each finite mind’s mental sphere of influence to just one body, and allow for only a single mind to interact with each particular body in the direct, immediate mental-physical way we are envisioning here.\textsuperscript{250} Noting these assumptions as such does not necessarily demean them. They are quite plausible, and Spinoza might endorse all three mentioned, at least in some form. The point is simply to notice that, until we make these additional conservative moves, we can put the pairing problem in far more dramatic ways, like those cases of pathological causation mentioned previously. For instance, we can ask what pairs m1 with actions of my body B, rather than with some distant motorcycle M—that is, why telekinesis does not occur. Or we can ask why it’s not the case that both m1 and m2 are paired with B—that is, why Privilege is true. Or we can ask why m1 but not m2 can cause further mental changes in the mental substance MS1—that is, why telepathy is non-actual. The basic questions of this section are: what sort of relation pairs various mental and physical causal relata, and in virtue of what does that relation, rather than some nearby alternative, obtain?

Once the causal pairing problem is seen in this light, as a limited case of pathological causation, we can glean insights from the two basic dualist responses in the literature to this more limited problem and apply them more widely. We can evaluate

\begin{footnotesize}
\textsuperscript{250} That is, of course my mind can interact with my own body, which can then interact with the body of my friend (to push him, say), but this action is indirect in a way that my interaction with my own body is not. Of course, since that more direct sort of interaction is fundamentally puzzling, the distinction here may be one of degree rather than kind, but regardless there seems to be a fairly basic difference that we routinely observe in thought and action.
\end{footnotesize}
their efficacy both with respect to the limited statement of the pairing objection found in the literature, and to the more general issue of pathological causation in Spinoza. I do conclude next section on an uncertain note—it is just not clear that premise 1 of the Pairing Argument is true. But recognizing the pairing problem as a more limited instance of Spinoza’s general problem of pathological causation clearly demonstrates the limits of the current debate as it stands. It also shows the way in which reflecting on Spinoza’s problem might give one better reasons than one currently has for favoring certain moves in the contemporary debate at the expense of others.

4.2.1 Premise 1 of the Pairing Argument: Does interactionist dualism imply nihilism about the causal pairing relations?

Now that we’ve seen the general form of the causal pairing objection to dualism, as given by Bennett and Kim, it is worth seeing how the physicalist proposes to answer the pairing question. This is important because if both physicalism and dualism must embrace pairing relation nihilism—the view that there are no uniquely specifying pairing relations between causes and effects (perhaps beyond the causal relation itself), either within or between attributes—the physicalist must press her point about the advantages of physicalist explanation in some other way. I think the physicalist is in a strong position to defend her relative advantages without relying on the pairing problem, as I discuss in 4.3. But here I want to describe the physicalist responses to the pairing problem in the

---

251 I will discuss this more in 4.2.2, when we see a set of parity arguments that explicitly embrace the conclusion that because both physicalists and dualists should be pairing nihilists, the pairing problem cannot possibly be a barrier to generating the sorts of explanations we are interested in. Here, though, Kim’s account of spatial relations as pairing relations provides a barometer against which to measure more positive dualist accounts, which 1) accept the pairing challenge and 2) argue that the dualist can meet it.
literature, and evaluate whether dualists who accept the pairing relation challenge can provide satisfactory answers to it.

Kim contends that only spatial relations can answer the pairing questions that necessarily arise in cases of solely physical causation, as well as in those mental-physical cases under discussion. Kim first gives the solely physical case: “two guns, A and B, are simultaneously fired, and this results in the simultaneous death of two persons, Adam and Bob. What makes it the case that the firing of A caused Adam’s death, and the firing of B caused Bob’s death, and not the other way around?” Two related strategies spring to mind. Either we could trace a continuous causal path from A to Adam and B to Bob, using some tracking device. Or we could look for a pairing relation that holds between A and Adam, B and Bob, but not between any other pairs of objects. Kim thinks: “In this particular case, when the two guns were fired, gun A, and not gun B, was located at an appropriate distance from Adam and pointed in his direction, and similarly with gun B and Bob.

It is these spatial relations (distance, orientation, etc.) that help pair the firing of A with Adam’s death and the firing of B with Bob’s death.” Again, notice that while A and B must be similar in some ways—they can’t fire differently shaped bullets, or else we could explain their effects without appeal to spatial relations—qualitative indiscernibility is unnecessary. All that’s required is enough similarity to get the problem going, plus a background assumption that some explanation is forthcoming and the

---

252 For now, we are assuming that solely physical cases of causation occur, which seems dialectically appropriate, since Descartes and other dualists no doubt agree.

253 Kim, Physicalism, or Something Near Enough, 78-79

254 Kim, Physicalism, or Something Near Enough, 79
further argument that spatial relations are all that’s available to play the pairing role. Kim further thinks that a similar story will hold for all cases of physical action between distinct non-mental objects, and I think parts of his story can plausibly be extended even for internal action with a complex object, in terms of some spatial relations among parts.255

Rather than evaluating Kim’s positive proposal, let’s consider his negative case against the dualist. Can the dualist tell a convincing story when we turn our attention to soul-body causation? Kim argues that she cannot:

There are two souls, A and B, and they perform an identical mental act at time $t$, as a result of which a change occurs in material substance M shortly after $t$. We may suppose that mental actions of the kind involved generally cause physical changes of the sort that happened in M, and moreover, that in the present case it is soul A’s action, not soul B’s, that caused the change in M. What relation might serve to pair soul A’s action with the change in M, a relation that is absent in the case of soul B’s action and the change in M?256

Kim then further argues that mind-to-mind causation cannot be explained any better. As he explains further, it’s not just that the dualist is hard pressed to describe the relevant pairing relation. Rather, Kim believes that are in principle difficulties with the

255 There will be complicated cases from quantum mechanics that might raise troubles for Kim’s strategy on the microscopic level, specifically as regards quantum entanglement. I’d bet that such cases could conclusively show that spatial relations alone are insufficient for answering the pairing challenge, though I won’t defend this assertion here. But remember, Spinoza’s problem is broader than Kim’s, and Spinoza need not, and indeed does not, share Kim’s physicalist commitments. In terms used earlier, while it may be that spatial relations will not due the trick in providing the necessary asymmetry between physicalist and dualist explanations, some purely physical relations, not reducible to merely spatial relations, might fare better. As Spinoza says in 3p2s: “For indeed, no one has yet determined what the body can do from the laws of Nature alone…for no one has yet come to know the structure of the body so accurately that he could explain all its functions.” Which is to say: Spinoza is interested in the broader asymmetry, not in Kim’s particular, spatial answer to the pairing question.

256 Kim, Physicalism, or Something Near Enough, 80

219
dualist’s identifying any relation capable of pairing causes and effects uniquely. He describes what we might label the *uniqueness objection*:

We seem to be in need of a certain kind of ‘space,’” Kim argues, “not physical space, of course, but some kind of a nonphysical coordinate system that gives every mental substance and every event involving a mental substance a unique location (at a time) and which yields for each pair of mental entities a determinate relationship defined by their ‘locations’ (analogous to the distance-orientation relationship between a pair of spatial objects).…But I don’t think we have any idea what such a framework might look like…I don’t think we have any idea where to begin.257

In addition to the fact that the dualist lacks a determinate framework for specifying unique pairing relations between mental substances and physical events, Kim identifies a second problem for dualist views. Kim’s second contention is that counterfactual statements about intentional relations will lack grounds—what we might call the *grounding objection*—if they cannot appeal to spatiotemporal relations. His defense of this claim begins with an example of two mental substances, B and C, having all the same intrinsic properties and, at the same time, acquiring a further property G. He then says: “I think we must countenance the following to be a possible situation: A’s having F at t causes B to have G at t*, but it does not cause C to have G at t*.”258

If the dualist posits a counterfactual that holds for one mental substance but not another, Kim thinks, he must provide a general reason why this is so. “There must be an intelligible and principled account of why [a] first counterfactual is true and [a] second is false. I do not believe we could simply assert this as a brute fact for which no explanation

---

257 Kim, *Physicalism, or Something Near Enough*, 81-82

258 Kim, *Physicalism, or Something Near Enough*, 83
is possible or needed.”259 One might wonder what Kim’s general reasons are for rejecting brute counterfactual dependence, outside a more broadly rationalist framework perhaps committed to the PSR, but regardless, his response shows how our problems interrelate. The ungroundedness of counterfactuals, of course, stems from the lack of a determinate relation-specifying framework analogous to the spatial one. And both appeals, at bottom, suggest the unintelligibility of causation of any sort, grounded intermediately by any other sort of relations, that doesn’t ultimately rest on spatiotemporal foundations.

Some dualists simply reject Kim’s pairing challenge, but those who do not should shoulder the burden of identifying plausible pairing relation candidates and arguing that they perform as well or better than the spatial relations Kim defends. Specifically, they must argue, at least, that the dualist’s candidates avoid the uniqueness and grounding objections. Geoffrey Madell takes up this burden in his thorough discussion of Kim’s version of the causal pairing problem. He argues that the dualist can provide the requested pairing relations—intentionality relations—once she has undermined and inverted the physicalist’s supposition about the relative priority of intentional and causal relations, respectively. He says: “[the] problem for Kim is what seems to me to be his extraordinary misconception as to the nature of intentionality. Kim claims that ‘we need causal relations to understand intentionality.’ If this were the case, then there might indeed be the problem that Kim supposes that he has highlighted…But the claim that intentionality rests on causation is simply false.”260

259 Kim, Physicalism, or Something Near Enough, 84

260 Madell, “The Road to Substance Dualism,” 58
Madell goes on to say that he can solve the causal pairings problem by accepting A. J. Ayer’s three criteria for determining bodily ownership. These are: my body is uniquely under the control of my will, delineated by my sensations, and provides my point of view on the world. Madell thinks Kim errs because he supposes that the first, causal criterion is exhaustive, and he also seems to allow that an agent might have a sort of ‘objectless willing’, in Madell’s words. For Madell, Kim erroneously believes that intentionality is parasitic on causation. Madell thus interprets Kim as saying that we cannot tell towards what body an agent’s willing is directed towards until it enters into uniquely specifying causal relations with it.

But Madell thinks this gets matters exactly backwards. Intentionality is “not a matter of a causal process running from an object in the world to the subject, but of the directedness of thought from the subject to an object in the world. [So] surely there can be no obstacle to accepting that, as I said above, I think volitionally of that right arm which is part of the body delineated by my sensations, and which provides the center from which I view the world.” And since intentionality precedes causation, causal uniqueness is secured by the fact that my intentional relations are uniquely directed at one body and no other. This inversion of the priority of intentionality and causation thus dispenses with the uniqueness objection.

Madell’s view bears strong affinities to Daniel Garber’s fascinating discussion of ‘what Descartes should have told Elizabeth,’ and faces similar problems. Elizabeth,

\[\text{261}\] Madell, “The Road to Substance Dualism,” 59

\[\text{262}\] Garber, “Understanding Interaction: What Descartes Should Have Told Elizabeth,” 168-188, in *Descartes Embodied*
notably, requested an explanation for the soul’s ability to determine bodily motion. She says:

I ask you to tell me how man’s soul, being only a thinking substance, can determine animal spirits so as to cause voluntary actions. For every determination of movement seems to come about either by propelling of the thing moved, by the manner in which it is propelled by that which moves it, or else by the quality and shape of the surface of the latter. Now contact is required for the first two conditions, and extension for the third.²⁶³

On a simplified version of Garber’s intriguing reading, what Descartes should have told Elizabeth is that we can only understand physical causation (in terms of bodily impact) by reference to mental causation (in terms of intentionality). Moreover, certain bodies (i.e. those tightly connected with minds) can only be individuated by first delineating the scope of their corresponding mind’s intentional action. The question is not, therefore, how the dualist can explain mind-body causation without courting unintelligibility, but rather how we can understand the physical world without presupposing such causality. Similarly, Madell thinks we can understand this mind-body causation by first understanding where my willing is intentionally directed (towards whichever body delineates my sensations and provides my point of view). We can then explain the unique causal connection between my mind and body in terms of this intentional connectedness.

Spinoza and contemporary physicalists have different responses to Madell’s strategy of inverting the priority of intentional and causal relations, and then arguing that the uniqueness intrinsic to the intentional secures the uniqueness of causal pairing. A physicalist like Kim simply complains that this strategy makes all causation

²⁶³ PE, 22 (Letter of May 6, 1643)
unintelligible. In order to understand causation we need to understand the mind-body union, which is either primitive or grounded in some complex, mentality-laden intentionalist account. But we cannot do this because we do not understand what this union is. He says: “If God ‘united’ my mind and my body to make a person, there must be a relationship R such that my mind stands in relation R if and only if that mind and that body constitute a unitary person. In uniting my mind and my body, God related the two with R. Unless we know what R is, we do not know what God wrought.” And, we might continue, if Madell (and Garber) are right, then in failing to understand R, we also fail to truly understand any cases of physical causation. But, Kim might continue, at least sometimes we do understand physical causation despite remaining ignorant of R. So Madell’s strategy of explaining the causal in terms of the intentional cannot be the right account.

By contrast with contemporary physicalists, Spinoza thinks Madell is mostly right. After all, it’s axiomatic for Spinoza, as he claims in 2ax4, “[that] we feel a certain body is affected in many ways.” Later, he argues, “the object of the idea constituting the human mind is the body, or a certain mode of Extension which actually exists, and nothing else” (2p13). Neither Spinoza nor Madell faces any particular mystery about identifying which minds and bodies are connected, at least after Spinoza incorporates the rest of his metaphysical system, in particular the necessity of mind-body identity. But, of course, the differences between them are enormous. For one, for Spinoza this relationship between an idea (the mind) and its object (the associated body) isn’t causal, or even explanatory at all! It’s not that causality fails to exhaust, as Madell claims, the mind-body

Kim, Physicalism, or Something Near Enough, 78
relation, but that causality or other conceptual connections between the two play no role at all. That’s all thanks to Spinoza’s idiosyncratic mind-body identity, of course, which relies on the explanatory barrier whose justification we are currently interested in. So Spinoza’s response must be conditional: Madell’s strategy of inverting causal and intentional relations might succeed, if this aspect of the pairing problem were the only issue facing the dualist. But I think Spinoza believes the problems for dualist explanation run much deeper, and thus that the dualist cannot, ultimately, avoid the criticism. It will take more work to see just how this is so, however.265

Meanwhile, we can ask how Madell’s account fares with respect to Kim’s second objection: that counterfactual statements about intentional relations will lack grounds if they cannot appeal to spatial relations. Recall that his defense began with an example of two mental substances, B and C, having all the same intrinsic properties and, at the same time, acquiring a further property G. Kim suggests: “I think we must countenance the following to be a possible situation: A’s having F at t causes B to have G at t*, but it does not cause C to have G at t*.”266 But, Kim thinks, the dualist cannot give a general reason why this counterfactual holds in one case but not another. “There must be an intelligible and principled account of why [a] first counterfactual is true and [a] second is false. I do

---

265 For Spinoza, all this is compatible with accepting something like Madell’s claim, which is that causal relations might fundamentally be explicable only in intentional, or formal causal, or broadly teleological terms. In fact, I think this reading, if correct, would significantly strengthen Spinoza’s position, a point I return to at the end of the present work. For a powerful defense of a teleological understanding of causation in the service of panpsychism, see Morch, Panpsychism and Causation, and for a similar argument for understanding all causation as formal causation in Spinoza specifically, see Hubner, “On the Significance of Formal Causation in Spinoza’s Metaphysics.” Nonetheless, it bears noting that most discussions of Spinoza seem to presuppose that all causation in Spinoza is efficient causation, and with good reason, at least as concerns several major texts from the Ethics.

266 Kim, Physicalism, or Something Near Enough, 83
not believe we could simply assert this as a brute fact for which no explanation is possible or needed.”

Kim’s request for explanation here has a decidedly Spinozistic flavor, in several respects. Like Spinoza, he seeks to rule out brute relationships in this domain, and like Spinoza he sees the dualist as incapable of doing so. But the discussion here seems to reveal either a third criterion beyond uniqueness and grounding, or at least a slightly different emphasis, which is that the pairing relations between causes and effects not be haecceistic. O’Connor states what I take to be Kim’s presupposition here admirably clearly, suggesting:

Causality, we might say, is non-haecceistic: objects do not have a primitive disposition to act on certain other individual objects; they are instead disposed to act on any objects have the right general characteristics. At different times, the same object will achieve much the same effect on different tokens of the same general type.

That’s why, in the example given above, the difference in counterfactuals between the intrinsically identical B and C looks so difficult to understand. What could make it the case that intrinsically identical substances had different properties, starting from the same intrinsic base set, if not some difference in their spatial relations?

But while I think that Spinoza accepts Kim’s assumption that causal relations be non-haecceistic, he could not accept Kim’s description of the case in which both the grounding and non-haecceistic objections make their appearance. The problem for Spinoza is simple. Kim sets up the problem as if there could be such intrinsically

267 Kim, Physicalism, or Something Near Enough, 84
268 O’Connor, “Causality, Mind, and Free Will,” 106
identical substances, and then finds unintelligibility downstream, in virtue of the dualist’s failure to explain the relevant counterfactuals of causality without appeal to spatial relations. But Spinoza just rejects the setup. The PSR entails the Principle of the Identity of Indiscernibles, which makes multiple intrinsically identical substances a genuine impossibility. We already saw this point earlier, when I argued that in order to have any hope of finding a version of the pairing problem in Spinoza, we could not assume that qualitative indiscernibility was one of its starting assumptions. I further argued that the problem could be intuitively developed without this assumption.

Leibniz gives a nice statement of the rationalist reasons for rejecting this setup, reasons that Spinoza would presumably share. Indeed, Leibniz supposes that genuine substances that are spatially separated must express their spatial separation intrinsically. He writes, in his letters to de Volder: “Hence, in nature, there cannot be two bodies at the same time perfectly similar and equal. Also, things that differ in place must express their place, that is, they must express the things surrounding, and thus they must be distinguished not only by place, that is, not by an extrinsic denomination alone, as is commonly thought.269 Leibniz’s view, in slogan form, is this: no difference without an intrinsic difference.

By contrast, Kim stipulates that two substances might be intrinsically identical (yet distinct), differing only in their spatial relations, and then tries to argue that the dualist cannot explain or ground the relevant counterfactuals concerning their respective intentionality relations. But for Spinoza, Kim has introduced incoherence right from the start, with the assumption that such situations are genuinely possible. So it is difficult,

269 Leibniz, Essays, 175
from Spinoza’s perspective, to disentangle Kim’s grounding objection from the impossible case with which he motivates the objection.

I thus do not feel confident evaluating whether Kim’s grounding objection to dualism succeeds, since on Spinoza’s behalf I reject the starting points from which it proceeds. This leaves two objections remaining: the absence of a uniquely specifying framework for mental-physical causal relations, and the general requirement that the pairing relations not be haecceistic.

Insights from two other dualist philosophers, Peter Unger and Timothy O’Connor, can be combined with Madell’s earlier inversion strategy, in order to begin allowing us to sum up the results of this section. First, Peter Unger has insightfully developed several clever haecceistic accounts of causation, within both dualist and purely physicalist frameworks.²⁷⁰ Now making room for dualistic haecceistic causation may well be Unger’s ultimate goal in developing such frameworks. But it’s by no means clear to me that Unger’s account of physicalist haecceistic causation clashes with any necessary tenets of broadly physicalist dogma, even if none of that dogma’s defenders accept it. That means, however, that Kim’s argument can only show that the conjunction of physicalism and the requirement on non-haecceistic causation can undermine dualism. But if the dualist can show that, once the requirement on non-haecceistic causation is dropped, the dualist and the physicalist are both on equal footing, then Kim has not achieved his goals. He has not shown that there is no defensible dualist position on this question, which is what the causal pairing objection is designed to do.

²⁷⁰ Unger, Empty Ideas, 68-76 and elsewhere in the same volume, as well as in All the Power in the World, in several places, describes these frameworks.
Second, philosophers like O’Connor, among others, have answered Kim’s challenge to provide at least the rough outlines of a non-spatial framework that might secure the unique relations between souls required for mental causation. They have denied, that is, that we have no idea what such a framework would look like, contrary to Kim’s assertion otherwise. For instance, O’Connor sketches the following possibility:

Suppose that God, in generating a series of souls, ordains that in their initial state they conform to a mathematically describable array, with each soul carrying the information of its present ‘location’ in the array as a primitive intentional state. Souls ‘move’ through the array over time by forming intentions to occupy a specified location. Among the basic laws of this souler system is a dynamical one that governs the actual re-arrangements as a function of all such intentions. (Perhaps, analogous to a time-sharing condominium arrangement, souls continually form ordered preferences as to their subsequent location. The dynamical law might factor in previous success in obtaining highly-ranked preferences, give a certain weight to preferences to remain in one's present ‘neighborhood', and so forth.271

Spatial metaphors pervade O’Connor’s non-spatial alternative—souls ‘move’ from ‘location’ to ‘location’ within an array. Still, as he suggests, this should not itself be taken as a conclusive objection—spatial talk within one’s system does not automatically make its existence parasitic on spatial relations.272 Dualists, then, have begun to provide somewhat detailed accounts responsive to the features of the pairing problem—uniqueness, grounding, and non-haecceity—that Kim doubts they can address even in principle.273

271 O’Connor, “Causality, Mind, and Free Will,” 107

272 See O’Connor, “Causality, Mind, and Free Will,” 107, especially his remarks on phase space in quantum mechanics.

273 It’s a further question, of course, whether Unger’s haecceistic causation adequately describes the causal behavior we actually observe, or whether the dualist can motivate restricting this brand of causation to mind-body interactions. O’Connor’s mathematically ordered array of souls still presents a rather abstract alternative to purely physical pairing relations. Plus, he does appeal to ‘primitive intentional
Finally, Madell’s attempt to invert the priority of intentionality and causation, as well as Garber’s similar attempt on Descartes’ behalf, is clever but not completely convincing *in isolation*. Their accounts threaten to render both mental-physical and purely physical causation unintelligible, as Kim claims, at least without placing the accounts within a more comprehensive theory that supports the inversion strategy on broader grounds. Minimally, Madell and Garber’s Descartes seem to require that their interlocutor grant them that this widespread unintelligibility is inevitable on competing views. Moreover, we must be able to make enough sense of the mind-body union to fruitfully use the intentionality-causality inversion in productive explanations. This is an obstacle some more comprehensive theory would presumably surmount, but I see no way to do so with the material Garber and Madell have presented thus far.

I have serious worries about the viability of all the dualist proposals considered thus far, but I also recognize the murkiness of the dialectical situation at this point. Given that, I think the appropriate conclusion to draw is that it is not clear from the foregoing arguments whether (1) interactionist dualism entails nihilism about the pairing relations. Of course, this does not mean that the dualist can successfully generate the explanations that allow her to explain why causation is ordinary, rather than pathological. For all that’s been shown thus far, it may simply be that the pairing relation problem is not the best way of getting at the central differences between the physicalist, dualist, and Spinozist states’, which suggests that he might have trouble with Kim’s grounding objection. On the other hand, O’Connor might instead appeal to God’s intentions to ground each soul’s initial primitive intentional states. After that point, the souls simply develop dynamically in response to the system of laws coordinating each soul’s intentional interactions with every other. Moreover, it’s not clear that Kim’s vision of physicalist relations avoids an equivalent commitment to primitive states. This parity is important, since Kim does seem to think of spatial location as primitive in some sense, not supervening on the intrinsic properties of a thing.
positions. Indeed, as we will see next section, Audi and others have argued that neither the dualist nor the physicalist can successfully answer the pairing problem, which Audi takes to undermine the force of the pairing challenge to the explanation of causation.

4.2.2 Premises 2-3 of the Pairing Argument: Causal Pairing Nihilism Undermines Mental-Physical Explanations

Paul Audi, Karen Bennett, Michael Tooley and others have suggested that the pairing problem may arise from accepting an incorrect picture of causation. Audi’s explicit strategy in defending dualists against the pairing problem is to use several of Tooley’s arguments to suggest that \textit{causal relations are primitive}.\textsuperscript{274} This, Audi thinks, undermines one of the two presuppositions required for the pairing problem to get off the ground. The first required assumption grants the possibility of qualitatively indistinguishable entities, both mental and physical. The second assumption is that “causal relations cannot themselves be pairing relations. Another way of putting the point is that the pairing problem seems motivated by the presupposition that two things cannot \textit{brutely} stand in a causal relation.”\textsuperscript{275} Audi’s claim is that both physicalists and dualists should deny this second assumption.

Similarly, Karen Bennett suggests that the dualist’s best response to the pairing problem is to argue that it poses a challenge physicalism itself also fails to answer. Recall Bennett’s succinct formulation: “The Cartesian dualist should think that it is

\textsuperscript{274} In other words, according to Audi and others, we have Pairing Primitivism rather than Pairing Nihilism, but the connotations of the terminological point are important. For Spinoza, of course, armed with the PSR, a primitive relation just is one that does not really obtain. Audi, of course, thinks some primitive understanding of causation is required here.

\textsuperscript{275} Audi, “Primitive Causal Relations and the Pairing Problem,” 4-5
metaphysically possible for there to be two intrinsically indiscernible Cartesian minds, such that a) they simultaneously perform indiscernible mental acts m1 and m2 and b) m1 causes some particular physical effect p that m2 does not cause. Following her description of the worry, Bennett mentions two possible ways of generalizing the pairing problem, so that the dualist faces no special burden to respond to it.

First, anyone who holds in statue-lump cases that there are two objects, and that they can be involved in distinct events, faces the problem of say, for instance, whether the statue or the lump causes one’s toe to break when it falls on it. I won’t explore this much here, since I think the right response to those cases is to argue, regardless of one’s favored relation between statues and lumps, that they share all their causal powers. The second, more interesting case comes from the physicalist Michael Tooley:

Imagine a world containing only two intrinsically indiscernible particles, each of which undergoes an intrinsically indiscernible process that has a 50% chance of generating a daughter particle. Imagine that one daughter particle is generated equidistant between the two originals. What makes it the effect of one of the original particles rather than the other?

This strategy of pointing out that your opponent has the same problem that you do doesn’t yet make it go away, but it comes close. It comes even closer if the shared problem lets the interactionist dualist reasonably assert that the pairing relation is primitive just because causation is primitive. As Bennett notes, the dualist’s appeal to a primitive relation is significantly more plausible when non-dualists like Tooley are drawn to the view for reasons wholly independent of their desire to defend interactionism. Audi

276 Bennett, “Mental Causation,” 320

277 This statement occurs in Bennett, “Mental Causation,” 321, but the original example is due to Tooley, “Causation: Reductionism vs. Realism.”
embraces this strategy, relying heavily on Tooley’s cases to stake his case for primitive causal relations as primitive pairing relations.

Of course, any argument that encourages inventing a primitive relation and running with it is bound to seem suspicious to Spinoza. Indeed, Tooley’s original paper is, for those inclined to rationalism of some sort, a model example of *how not to argue,* since his main arguments appeal to the possibility of 1) indeterministic causal laws; 2) uncaused events; 3) probabilistic laws; and 4) exact replicas of causal situations in deterministic worlds.278

In general, the strategy of arguing that the pairing problem is unanswerable for everyone, and thus probably irrelevant to our ability to provide compelling explanations for why causation is ordinary and not pathological, mixes even worse with Spinozism than with contemporary physicalism. Of course, all involved parties admit that skepticism is false and thus that some such compelling explanations exist.

But Audi and Tooley’s arguments are difficult to accept because they begin from such deeply anti-Spinozist foundations. Specifically, several brute facts are stipulated. First, both mental and physical entities are imagined to be capable of differing while not differing qualitatively in any way. Second and relatedly, these qualitatively indiscernible physical entities are used to construct cases that cast doubt on the coherence of the demand for causal pairing relations that are not primitive. This then absolves the dualist of her obligation to respond to the problem. Finally, a primitive causal relation is brought

---

278 The relevant assumptions are stated explicitly in Tooley, “Causation: Reductionism vs. Realism,” 225, 230. Of course, the assumption of indeterministic laws is significantly more plausible in light of contemporary physics—Spinoza would clearly have favored the deterministic versions of quantum mechanics, perhaps Bohmian mechanics, for example—but as far as I can tell the assumptions of exact replication of causal situations or uncaused events are no more plausible now than they were to Spinoza.
in to solve the problems generated by the assumption of qualitatively indistinguishable but primitively differing particles, in both the mental and physical realms. Spinoza, of course, has general metaphysical reasons, stemming from the PSR and the PII, for finding each of these moves unacceptable. But, as each move finds some support in the literature—especially assuming qualitatively indistinguishable physical particles—it’s hard to evaluate whether Spinoza’s position connects with the extant arguments.

Fortunately, however, there might be another level at which these questions concerning primitivity and explanation can be adjudicated. Audi, at least, seems sensitive to the worry that there might be a more serious problem lurking in the neighborhood of the pairing problem as it’s typically discussed. Audi notes that dualists must make sense of the fact that mental substances interact with particular bodies in, as far as we know, only one way, described by the conjunction of the following two principles, which we have encountered several times before:

Uniqueness: A given finite mind, throughout its career, causally interacts with just one body.
Privilege: A given finite mind, throughout its career, is the only mind that can directly causally influence a given body.279

It’s important to keep track of why the demand—that the dualist make sense of Uniqueness and Privilege—is appropriate, in order to appreciate the force of my criticisms of Audi’s attempt to do so on the dualist’s behalf. The central challenge of this chapter is, given that causation is ordinary rather than pathological, how best to explain that fact. Physicalists like Kim propose the pairing problem as part of their argument

279 Audi, “Primitive Causal Relations and the Pairing Problem,” 13; as I did in chapter 3, here I modify the statements of Uniqueness and Privilege slightly, removing the possibility claims from within the principles themselves, so that the modal status of the principles can be evaluated more clearly.
against dualism. Spinoza defends the problem of pathological causation, of which the pairing problem is one instance, as an objection against all non-Isolationist views. And Audi’s strategy in responding to both groups is *tu quoque*—the pairing problem afflicts physicalists or Spinozists too, and so cannot be a special problem for the dualist.

For Spinoza, this response is unattractive for three reasons. First, the setup of most initial statements of the pairing objection relies on the possibility of two things being qualitatively indistinguishable, which Spinoza rejects as impossible. So Spinoza requires a version of the objection that sticks with the intuitive presentation and avoids this requirement. I have argued that this Spinozistic version captures the problem’s key features. Second, Spinoza rejects a dualist strategy that uses wholly physical causal cases—like the Tooley case of two indistinguishable particles given above—to try to enforce parity between mental-physical and physical-physical causation. By Spinoza’s lights, such cases don’t establish parity because the original pairing objection does not require, and indeed positively excludes, indistinguishable possibilities. Finally, the dualist uses parity between mental-physical and purely physical causation to argue for a view of causation as primitive. For the physicalist, primitive causation might be intuitively unacceptable, but hard to resist if the original formulation of the pairing problem, retaining qualitative indistinguishability, is used. For then the parity objection appears to work. But for Spinoza, making causation primitive, especially given Spinoza’s close association of causation and explanation, is unattractive. And crucially, the dualist actually requires a second primitive here, namely the truth of Uniqueness and Privilege. This second primitive creates the asymmetry between physicalists who accepts qualitative indistinguishability and Spinoza, on one hand, and dualists, on the other.
Uniqueness and Privilege are both plausible principles—indeed, Spinoza probably accepts a version of each. So why are Uniqueness and Privilege primitive for the dualist, but acceptable, at least on some interpretations, for Spinoza? It’s worth beginning with Spinoza here, since his explanation is much clearer. Immediately, however, one encounters an ambiguity in the principles, since neither ‘mind’ nor ‘body’ is precisely defined. Spinoza cannot help himself to the natural dualist reading—a finite immaterial soul substance distinct from anything physical interacts with one and only one physically extended body—because of both his Substance Monism and Mind-Body Identity.

But this is not Spinoza’s problem, but Audi’s dualist’s, since Audi means for Uniqueness and Privilege to be self-evident truths of our experience. He thus seems to owe us an ontologically neutral account of the terms that figure in these self-evident truths. Audi says: “It seems a plain fact that I have a single body that I alone control.” As I discussed last chapter, the sense in which these principles are ‘plain facts of experience’ is also ambiguous. It might mean that we directly experience controlling one and only one body, in a direct realist way. Or it might be shorthand for the claim that experiences of exercising control over a single body (our own), of trying and failing to exercise control over other spatially disconnected bodies and failing, and of never directly experiencing another mind’s control over our own body, are collectively a good inductive basis for motivating the two principles. This inductive reading would not make

---

280 For instance, in 2ax4, Spinoza claims: “We feel that a certain body is affected in many ways,” and in 2p13, he says, “The object of the idea constituting the human Mind is the body, or a certain mode of Extension that actually exists, and nothing else.” 2p13 appears to assert the only this mind interacts with this body (and vice versa), while 2ax4 seems to claim that the mind interacts with just one body. But, as we will see, this cannot be the whole story.

281 Audi, “Primitive Causal Relations and the Pairing Problem,” 12
violations of the two principles metaphysically impossible, but might justify the claim
that cases of pathological causation are non-actual, given the actual laws of our world.

Suppose we take this latter, inductive reading of the principles and try to combine
it with Spinoza’s metaphysics. Then, as 2p13 has it, “the object of the idea constituting
the human mind is the body, or a certain mode of extension which actually exists, and
nothing else.” For Spinoza, the ‘mind’ is just the idea that has the whole human body,
and nothing else, as its object, and this claim follows from earlier premises in Part II of
the Ethics, which ultimately (i.e. 2p13 cites 2p9c, which itself cites 2p6, 2p7, and 1p28)
depend both on the EB, and more particularly on Spinoza’s Monism and Identity Theory.
On this reading, Spinoza endorses Uniqueness and Privilege as necessary consequences
of his mind-body identity theory, which is itself a necessary truth entailed by Substance
Monism. But his notion of ‘mind’ ends up being much broader than we typically think,
since it 1) is not an immaterial soul but is also 2) not just the brain or central nervous
system, but an idea identical with whatever whole body is its object.

Alternatively, we can take a more commonsense reading of ‘mind’ and ‘body’,
and combine this interpretation with Spinoza’s definition of one body or individual in
2p13def: “When a number of bodies…communicate their motions to each other in a
certain fixed manner, we shall say that these bodies are united with one another and that
they all together compose one body or individual.” On this alternative, the Uniqueness
Violations cases of the last chapter, in which the mind exerts direct control over a bionic
limb, for instance, could be taken as instances where a finite mind, conceived now in a
commonsensical rather than Spinozistic way, does interact with more than one body (i.e.
the human body plus the bionic limb), also conceived commonsensically. But on a more
Spinozistic reading we should simply say that the bionic limb now communicates its motions in an appropriately fixed manner that it now constitutes a single body with the human body to which it is attached. Relatedly, the human mind identical to this body is now even more extended, since it includes not only the ideas of the brain and central nervous system, but also the ideas of the remainder of the organic human body and the ideas of the bionic limb as well. Whichever interpretation one favors, Spinoza’s metaphysics of mind-body identity ensures that the connections that exist are non-brute, necessary and explicable consequences of his identity theory and his theory of the individuation of bodies. There is nothing contingent or brute about whether and in what manner Spinoza accepts the two principles.

The contrast with Audi’s dualist could not be more stark, and the contingency and bruteness of the dualist’s explanation motivates the charge that the Spinozist has a significant explanatory advantage here. As Audi himself makes clear, the dualist must accept these principles in a somewhat odd way, since his justification is from experience—“it is a plain fact that I have a single body that I alone control.” On either

\[282\] Audi, “Primitive Causal Relations and the Pairing Problem,” 12; I should note that Lycan (“Giving Dualism Its Due,” 561-562) considers how the dualist might try to explain Uniqueness and Privilege, and he has a rosier take on the dualist’s prospects (at least in some moods). He says: “But then what explains the unique relationship between mind and body? This is indeed an embarrassing question, but the answer is to be found in whatever would explain the appearance of minds in the evolutionary process.” Lycan’s thought is that the dualist should suggest that Cartesian egos evolved, and are spatial (but still immaterial), and that once we realize that 1) not every trait is an adaptation—some are byproducts of adaptations, and 2) since the Cartesian soul interacts with the body, it may well be adaptive, the dualist has a plausible story about how it might evolve. But then, whatever \textit{that} story is can be used to explain the unique relationship between mind and body. I have several responses. First, Lycan argues that the most compelling form of dualism will accept the evolution of souls, rather than their eternity, as Descartes does. This is much different than the dualism considered thus far. Still, his points on the evolutionary process, from the dualist’s perspective, are well taken. The problem, it seems to me, is that, if some evolutionary story explains Uniqueness and Privilege—presumably Lycan thinks this introduces parity between the physicalist and the dualist on the question—\textit{then I think that Lycan’s dualist should think that Uniqueness and Privilege, as stated, are false}. That is, Lycan’s dualist should accept the possibility of the same sorts of Uniqueness violations—bionic limbs, electromagnetic communication with bodies at a distance—that the
disambiguation of this claim discussed above, Uniqueness and Privilege come out as only contingently true, such that violations of each are metaphysically possible but just not, according to the dualist, actual in our world. Why does the dualist need to go this route? Well, if the dualist takes the principles to be necessary, then the pairing argument—the qualitatively indistinguishable minds of Smith and Jones are psychologically synchronized and arm raisings S1 and J1 occur—suggests that “necessarily, if x and y have different causes or effects, they differ either in their qualitative properties or in some relevant circumstance.” Since by Audi’s hypothesis the minds of Smith and Jones are presumed to be qualitatively identical, but to have different effects, the two minds must differ in some relevant circumstances. For the physicalist, spatial relations provide the differing circumstances that explain the differential effects of two otherwise qualitatively identical minds.

Audi’s thought is then that the dualist seems hard-pressed to say what circumstance could account for this difference. Specifically, what could make it impossible that either (a) ‘both Smith and Jones cause S1 but neither causes J1’ is true or (b) ‘neither Smith nor Jones causes either S1 or J1’ is true. If Uniqueness and Privilege are necessary truths, then ‘(a) both Smith and Jones’ and ‘(b) neither Smith nor Jones’ are both impossible answers to the question of who causes the arms to rise. And it doesn’t physicalist should accept. But even disregarding how distant this brand of dualism is from present concerns, there is a deeper issue. Spinoza’s explanation seems superior to both the physicalist’s and Lycan’s dualist’s. So even if we accept Lycan’s questionable assumptions—that our lack of a model for interactionism might be seen as a mere consequence of our lack of a model for causation, and that with interactionism and evolution the dualist just about reaches parity with the physicalist—Lycan’s dualist still has to drop Uniqueness and Privilege, and his explanation is still inferior to Spinoza’s. And, as will become clear next chapter, I think that the physicalist might have trouble explaining why certain skeptical scenarios do not obtain, given that they do not, so this just puts Lycan’s dualist in the same boat as the physicalist, both (I will continue to argue) inferior to Spinoza.

Audi, “Primitive Causal Relations and the Pairing Problem,” 13

283
seem like the dualist can make sense of this necessity, at least without appeal to some spatial relation to ground the pairings. As Audi puts it:

This is why it is important to note that Uniqueness and Privilege are contingent truths. If they were necessary, (a) and (b) would be impossible. What is more, if (a) and (b) were impossible, then treating singular causal facts as primitive would be no help to the dualist. For (a) and (b) simply represent two different ways that the causal facts could lie with respect to four particulars. If such facts can hold primitively [as the causal primitivist thinks they must], then surely both (a) and (b) are possible.\textsuperscript{284}

Audi’s thought seems to be that the defender of Uniqueness and Privilege, since she appeals to primitive relations not plausibly viewed as necessary, should also think that other causal states of affairs should at least be possible. For what could explain why the actual primitive state of causal affairs—according to which Smith’s mind causes S1 and Jones’s mind causes J1—holds necessarily? Moreover, for the interactionist dualist who is also a theist, God’s omnipotence seems to allow for His causing some other finite mind, but neither Smith’s nor Jones’s, to cause S1 and J1, or for some other finite extended thing to do so.\textsuperscript{285} Of course, the interactionist theist will add, God does not deceive us in this way, but there is no in-principle impossibility in his doing so. So, Audi thinks, the dualists should just rely on our shared experience that each principle is at least contingently true, even while admitting that pathological causal states of affairs are possible. That is, we should accept that perhaps both synchronized agents, or neither synchronized agent, or some other agent entirely, could possibly exercise control over the

\textsuperscript{284} Audi, “Primitive Causal Relations and the Pairing Problem,” 14

\textsuperscript{285} Audi, “Primitive Causal Relations and the Pairing Problem,” 14, fn. 21 explicitly discusses the relationship between the two principles and divine omnipotence, and makes basically this point, although there he speaks of God’s omnipotence licensing God’s directly causing S1 and J1. My versions of Uniqueness and Privilege are restricted to finite minds, however, so what Audi actually discusses would not directly counterexample them. However, the related examples discussed in the main text do.
bodies of the two agents in the various pairing cases considered thus far. We should just simply hasten to add that such possibilities do not, in fact, obtain.

Near the very end of his paper, Audi suggests:

One might worry that the very possibility of such cases would give way to rampant skepticism, for one might think we are in no position to verify that there are no pairs of qualitatively indiscernible minds... While it may be true that things would seem to me just as they are if there were another mind qualitatively indiscernible from mine (systematically overdetermining my movements), this does not constitute a reason for skepticism. I cannot offer an account why not, but notice that much more than dualism is at stake if we hold that we cannot be justified in believing something when things would seem just the same if it were false. Familiarly, that would impale, at least, all our perceptual knowledge.\(^\text{286}\)

Audi’s imagined qualitatively identical mind systematically overdetermining one’s movements is recognizable, of course, as a version of the sorts of pathological causation scenarios considered thus far.

We might put the argument expressed by the paragraph above as follows:

1. Things could seem exactly as they now seem while causation was pathological rather than ordinary.
2. We cannot be justified in believing that X is the case if things would seem the same regardless of whether X were true.
3. **Sub-Conclusion:** We cannot be justified in believing that causation is ordinary.
4. If the principle in (2) is correct, we also cannot be justified in believing that any of our perceptions give us knowledge.
5. But we do have perceptual knowledge, contrary to (4).

\(^{286}\) Audi, “Primitive Causal Relations and the Pairing Problem,” 15
6. So even though we seem to be able to infer that we are not justified in believing that causation is ordinary, we cannot infer this since the principle in (2) is false.

This would not be a bad argument, if either the physicalist or Spinoza required the principle given in (2) for their argument. But neither requires (2) or anywhere suggests that he is committed to it.

Audi’s argument here is thus irrelevant, directed at a controversial epistemological premise, (2), that none of his opponents need to endorse or anywhere explicitly commit themselves to. At minimum, then, Audi requires more evidence to justify attributing this argument to his opponents. Rather, the basic charge these opponents make against the dualist is simply that the dualist’s brute adherence to Uniqueness and Privilege puts him at a serious explanatory disadvantage. By contrast, Spinoza can offer an explanation either for the truth of Uniqueness and Privilege (on one interpretation given above), or for the very limited cases of their falsity, in which it is not the case that a given finite mind interacts with one and only one body (on another interpretation given above). Each explanation arises as a necessary consequence within Spinoza’s broader framework of mind-body identity, and no mind-body connections are brute, for Spinoza. Every unique and privileged mind-body correlation that obtains is explicable in light of more fundamental facts about the monistic structure of God or Nature.

Moreover, it’s not even clear that the physicalist’s (as opposed to Spinoza’s) explanation for Uniqueness and Privilege is brute in quite the same way. For the physicalist can tell a complex story, grounded in physics, biology, chemistry, and
neurophysiology, about the conditions required for the two principles to hold true. Actions result from neural communication between the spinal chord, brain, and muscles, and were enough of these physical connections severed (as after a paralyzing car accident), the mind’s unique connection to the rest of the body would be severed too. Strictly speaking, like Spinoza the physicalist will want to say that Uniqueness and Privilege are probably false, on a commonsense reading of ‘mind’ and ‘body’, even though as a matter of fact most minds control one and only one body. However, on a more expansive interpretation of mind and body each principle might come out true for the physicalist as well. But for the physicalist, the conditions under which this typical one-one association might be violated—bionic limb control, another’s direct neural control over my limbs—however fanciful, will themselves be explained with reference to more fundamental physical facts. Such situations are possible, if they are, because an atypical electrical or chemical route of communication that didn’t exist previously has been established. By contrast, it is not obvious why the presence of some additional electromagnetic interaction between my brain and some distantly located body should have any effect on my non-spatial soul’s ability to communicate with that body. It is thus unclear how the dualist should incorporate these cases into her explanatory framework. So the foregoing too seems to advantage the physicalist in favor of the dualist.

In conclusion, in this section I have argued that many of the arguments designed to establish parity between the physicalist and the dualist on the question of pairing relations rely on premises unacceptable to Spinoza. In particular, these parity arguments require PSR and PII violation for their formulation. Of course, I argued earlier that the pairing problem can and should be stated without these assumptions, so perhaps the
parity arguments can be reformulated along similar lines. At the moment, however, I do not see how the dualist or the causal primitivist physicalist might do this.

Finally, I suggested that, even bracketing these most basic concerns about qualitative indistinguishability, the explanatory difficulties reemerge when trying to account for Uniqueness and Privilege. So even if the pairing relations are primitive because causation itself is primitive, as Audi has it, there is still an explanatory asymmetry between the dualist and her opponents. Spinoza has the resources to explain either in what sense Uniqueness and Privilege hold, or under what precise conditions they are false, if they are. Whatever correlations exist for Spinoza are necessary consequences of more fundamental features of God or Nature and his general theory of mind-body identity, and are explicable in light of these more general laws. The dualist, by contrast, treats Uniqueness and Privilege as actually true but possibly false, but then proceeds to offer no general justification for thinking that they are actually true, beyond ‘plain facts of experience’ called into question by the counterexamples of the last chapter. Moreover, the dualist cannot explain why certain physical facts—the presence of additional electromagnetic radiation between my brain and a bionic limb—should affect my non-spatial soul’s ability to communicate with additional extended bodies.

In 4.2.1, I argued that while I believe that extant dualist proposals for answering the pairing problem directly—from Madell, Unger, and O’Connor—have serious issues, it was not clear that Kim’s objections to them were conclusive. Here in 4.2.2, I suggested that there is similar unclarity about the status of Audi and Bennett’s arguments for nihilism about the pairing relations for both the dualist and the physicalist. This stems, in part, from the assumption of necessarily false principles, according to the Spinozist
perspective I am defending. Nonetheless, I concluded by pointing to a remaining
explanatory asymmetry that I argued the Spinozist (and perhaps the physicalist) was
better positioned to meet, regardless of one’s view on the success of the arguments for
causal pairing nihilism. Combined, sections 4.2.1 and 4.2.2 offer strong evidence that the
extant dualist responses to the pairing problem are unlikely to succeed. They thus serve
as an important plank in my argument against the dualist alternative on Spinoza’s behalf.

We can strengthen the foregoing results by examining more closely what dualist
explanation looks like in ordinary cases. In what follows, I will argue that, independent of
one’s view on the pairing problem, we have serious reason to doubt the success of dualist
explanation in ordinary cases. The nonexistence of law-governed psychophysics, plus the
fact that folk psychology is not a theory capable of generating reliable explanations on
the dualist’s behalf, makes it hard to see how the dualist will be able to explain why
causation is ordinary rather than pathological. Thus, while I have given Kim’s particular
proposal—that only spatial relations can serve as pairing relations—an inconclusive
evaluation, I want to vindicate his more basic insight that, on these questions at least,
there appears to be a stark explanatory asymmetry between the dualist and her opponents.
Said simply, the FP or Psychophysics Argument better illustrates, with respect to more
fundamental questions, the explanatory asymmetry that most interests me on Spinoza’s
behalf.

\[287\] However, this response is moderated by two factors. First, as I noted at the beginning of this
chapter, I am not defending a deductive conclusion—interactionist dualism is false—but rather the claim
that we have almost no reason at all to believe or find it at all plausible that interactionist dualism is true.
Second, as I discuss immediately below, I think the FP or Psychophysics Argument gets at more
fundamental issues than the Pairing Argument.
4.3 Folk Psychology and Nonexistent Psychophysics: Premises 5a-5b of the FP

Psychophysics Argument

To avoid the charge that dualism faces asymmetrical difficulties explaining why causation is ordinary rather than pathological, dualists can pursue further strategies, two of which I find more plausible, besides those of 1) providing causal pairing relations or 2) arguing for parity on the pairing problem, both of which I considered above.

Interactionist dualist strategies for explaining why causation is ordinary rather than pathological can be seen as lying along two independent axes—the positions discussed below, psychophysical laws and folk psychological explanations, are the two most plausible combinations of the responses one can give along each axis. So first, the explanations provided could be strictly law-like, or they could be weaker than strictly law-like (i.e. perhaps because laws are exceptionless and no such explanations are exceptionless in the requisite ways, or because laws require principled ceteris paribus clauses and no such explanations have these clauses). Second and independently, the explanations might come from commonsense, ordinary experience, or they might not (i.e. most likely because they are complex explanations from the natural sciences).

With this taxonomy in mind, we can see that the interactionist dualist might argue that there are, or could be, psychophysical laws, irreducible to physical laws, which govern specifically dualist (i.e. cross-attribute) interactions. Second, the dualist might argue that folk psychology (henceforth: FP) and folk psychological explanations mutually acceptable to dualists and physicalists/Spinozists can explain why causation is ordinary
not pathological, \emph{regardless of underlying differences in ontology}. The thought here is that if all disputing parties require higher-level folk psychological explanations, the acceptance of which explains why causation is ordinary not pathological, then the problem of pathological causation floats free of disputes about underlying ontology. This would be good for the dualist, since then the problem cannot be used as an objection to his metaphysical views in particular. But strictly speaking, the dualist might also defend non-law-like explanations from the natural sciences, or strictly law-like explanations from commonsense, as those options exhaust the remaining space in the taxonomy given above. I simply don’t think it helps the dualist to go in either of these two directions, and moreover as a matter of fact dualists have tended to defend their position in psychophysical or folk psychological terms.

In response to the defender of psychophysical laws, I argue that we have no good reason to believe in the existence of such psychophysical laws or any corresponding explanations, nevermind any good reason to believe that anyone currently has any implicit or explicit knowledge of these laws. In response to the advocate of folk psychological explanations, I argue that many folk psychological explanations are false or misleading, and that dualism offers basically zero non-
emph{ad-hoc} guidance for determining when this is so. Unlike competing forms of physicalism or Spinozistic Isolationism, which confine strict explanation to the domain of a single attribute and thus do not require folk psychological explanations within their ultimate ontological

\footnote{I give arguments in favor of these three considerations—requiring pairing relations and/or psychophysical laws and/or minimally, some method for explaining why causation is ordinary, not pathological—on pp. 193-199 of the present work. The structure in this chapter thus far is then: having already discussed and evaluated arguments in favor of the requirement that there be causal pairing relations, we can now turn our attention to the latter two requirements.}
framework, dualists can offer no principled way to determine when folk psychological explanations have, or are likely to, fail. When physical causes not plausibly construed as a belief or desire (i.e. a simple serotonin deficit) compete with beliefs or desires (construed as mental by the dualist and as some physical state by the physicalist) as causes of some further belief or action, Spinozists and physicalists can explain this fairly well, in principle if not always in practice. That is, Spinozists and physicalists can better explain which explanations are preferable when neurobiological and folk psychological explanatory schemes, and not just particular explanations, conflict. Dualists, by contrast, offer no principled explanatory framework or methodology for doing so, rendering their reliance on folk psychology independent of an underlying single-attribute strict explanatory framework unreliable and hence insufficiently explanatory. After discussing psychophysical laws, I turn to a more concrete examination of folk psychological explanation’s problems and the challenge they pose to the dualist.

David Chalmers, in his 1996 defense of the possibility of a highly naturalistic form of dualism in *The Conscious Mind*, suggests establishing a dualist explanatory framework from unknown, but in principle possible, psychophysical laws:

> Whenever we find an information space realized phenomenally, we find the same information space realized physically. And when an experience realizes an information state, the same information state is realized in the experience’s

---

289. This claim does not suppose that actual physicalists or Spinozists will eschew folk psychological explanations in ordinary cases.

290. In the present chapter, I often speak in terms of particular examples (i.e. an agent’s low serotonin level causes depression), but I mean for the examples to generalize. That is, while there is no deep problem with, in some cases, there being a dual explanation for an agent’s depression (i.e. some neurological explanation plus some broadly mental explanation), there is a problem when mental and neurobiological explanations conflict systematically, that is, when we ascribe mental causes to what are basically entirely neurobiological phenomena. So when I speak of ‘conflicting explanations’ in what follows, I usually have a more general conflict, rather than a single case where there is both a neurological and mental cause of a given state, in mind.
physical substrate. ... Principles concerning the double realization of information could be fleshed out into a system of basic laws connecting the physical and phenomenal domains.291

Atmanspacher, commenting on this same passage, adds: “Truly psychophysical phenomena in this spirit are neither physical nor psychological, and they are subject to psychophysical laws, neither to physical laws nor to psychological laws…[But] needless to say, these basic laws are yet to be discovered.”292 This last point encapsulates the first serious problem with this dualist strategy. Despite millennia of mental-physical interaction—most of human history, according to the dualist—and stunning recent advances in our understanding of the physical, we have made no progress in discovering any distinctively psychophysical laws. That is, we have discovered no laws connecting irreducibly mental phenomena to irreducibly physical ones—indeed, we have found no uncontroversial examples of any psychophysical phenomena at all!293 Moreover, there seems to be at least three problems for this sort of view, which Lycan elegantly lays out. He is worth quoting at length, even though I do not thereby endorse all facets of his view:

First, there is a problem about disciplinary authority. How might microphysicists be moved to posit Sellarsian sensa or Chalmers’ ‘what it’s like’ properties? Microphysicists do not study human behavior, or neural processing…The proposed revision of physics is not motivated by the physical data that are the microphysicist’s evidence base…Second, it seems that in any case the details of

291 Chalmers, *The Conscious Mind*, 284/286

292 Atmanspacher, “20th Century Variants of Dual Aspect Thinking,” 260, and see also Atmanspacher, “Notes on psychophysical phenomena.” Davidson, “Mental Events,” also denies the existence of psychophysical laws in defending a version of physicalism he calls anomalous monism.

293 Of course, there are plenty of laws connecting higher-level physical states, perhaps in-principle irreducible to lower-level physical states, to some other physical states. But this is not psychophysics—it’s just physics (or biology or chemistry—my argument does not depend on explicitly reductionist premises, though I am sympathetic with the program as confined to Extension). What we are after is genuine psychophysical laws, and those we do not have.
microphysics should be irrelevant…Mental properties are determined by
neuroanatomical properties, regardless of what constitutes the latter themselves.
Changes in the physics underlying biology and chemistry should not matter in any
way to the mind. Third, the Strong Naturalist [like Chalmers] faces a dilemma. If
any reduction of mind to the natural order requires a re-conception and expansion
of physics to incorporate novel entities and principles not motivated by the
physical data themselves, then either those entities and principles will be localized
where we now take minds to be (viz. in the central nervous system), or like other
entities and principles of fundamental physics they will pervade nature.294

Lycan’s second point—about the irrelevance of microphysics—is hotly debated
even among non-dualists, but his first point seems powerful and the dilemma in his third
point seems difficult to resist. For Lycan, the localized view of mental entities is
“coherent but unacceptable,” because it represents an “imaginable but grotesque”
arbitrariness in their distribution, given that we are assuming that they have the status of
new fundamental entities. And the second, ubiquitous view of mental entities as
fundamental is panpsychism, not dualism, which while interesting is certainly not the
view we are attempting to defend here.295

Given our current lack of good reasons for believing in the existence of
psychophysical laws, nevermind our current total ignorance regarding what they are if
they do exist, as well as the substantive theoretical difficulties Lycan identifies with there
ever being any such laws, I think there are good reasons for the dualist to prefer the folk

294 Lycan, “Recent Naturalistic Dualisms,” 356-357

295 The remaining quotes are from Lycan, “Recent Naturalistic Dualisms,” 358. It will turn out that
I think Spinoza should perhaps be interpreted as a sort of panpsychist, which means that he may be able to
take advantage of the position Lycan outlines here, albeit in a way that is even more compelling than the
versions of panpsychism Lycan considers. I discuss these issues in chapter 6, where additional arguments
against localized mentality are considered.
psychological route. That is, the dualist should argue that she could adequately explain away the threat of pathological causation and generate reliable mental-physical explanations *regardless of her underlying ontological disagreement* with Spinoza and with the physicalist. She should suggest that FP explanations mutually acceptable to her and to her opponents form a reliable basis for giving explanations of the relevant phenomena. If this were the case, the most serious issues with pathological causation would be resolved without entering into the underlying ontological dispute, thus making the problem of pathological causation a poor weapon against the dualist’s ontology in particular.

To make matters more concrete, consider two folk psychological explanations, one patently compelling, the other much less so.

**Beer**: When [A] someone *desires* a beer, *believes* that beer is in the fridge, *sees* the fridge in the next room, and does not observe any obstacles in her path or does not have any stronger beliefs dictating another course of action, [B] she tends with high probability to *walk towards* the fridge to retrieve a beer and to *become happier* in expectation of having a cold one. Often, [B] occurs *because of* [A].

**Depression**: When [A] someone *desires* to be happy, *believes* that having certain positive experiences and avoiding certain negative experiences is conducive to this happiness, and is mostly happy at time $t$, and does not *have too many or the most serious sort of negative experiences or interactions* (i.e. conversations with others, interactions inducing physiological stress, etc.) *after this time*, [B] she tends with high probability to *remain mostly happy* after time $t$, assuming the

---

296 I should say, however, that if some version of quantum mechanics that depends essentially on mentality were to become dominant in fundamental physics, this might change my opinion of the status of Lycan’s objections (though of course we’d still need to make some connection between these fundamental mental processes of wave function collapse and the mental phenomena we started off interested in explaining). Unfortunately, as Lewis, *Quantum Ontology*, makes pretty clear, collapse views are 1) pretty clearly disfavored by working physicists and 2) do not depend, in any case, on mentality, but only some measurement process. Of course, one might extend mentality to any measurement process to escape this latter problem, but that just runs the dualist into panpsychism, probably, contrary to her intentions as we interpret them here. Some interesting proposals for avoiding this consequence, rebutting the foregoing, and remaining within dualism, rather than panpsychism, while using the insights of quantum mechanics, occur in Corradini and Meixner (eds), *Quantum Physics Meets the Philosophy of Mind*. 251
overall balance of experiences remains roughly as before. Often, [B] occurs because of [A].

Examples of this sort roughly cohere with a model of folk psychology as a collection or systematization of platitudes proposed by David Lewis. He suggests that such generalizations have something like the following general form: “When someone is in so-and-so combination of mental states and receives sensory stimuli of so-and-so kind, he tends with so-and-so probability to be caused thereby to go into so-and-so mental states and produce so-and-so motor responses.” Presumably, these generalizations are useful not only for making predictions, but also as explanations for observed behavior in many cases. But this platitude conception of folk psychology is not the only one in the literature, as I’ll soon discuss.

But for now what’s important to note is the following. Beer seems like a perfectly good explanation of why Sheila walked to the fridge with a smile on her face. But Depression is a terrible explanation, even ignoring Spinoza’s view for the moment, for either Sheila’s current happy, or current depressed, state. Imbalances, or lack thereof, in her neurotransmitter levels are far more explanatory of her condition than what’s given in Depression, at least in a large percentage of cases. This appears to be a case where a neurobiological explanation—neurotransmitter imbalances explain depression—competes with a folk psychological one—certain events, feelings, or life experiences generally explain depression. It seems reasonable to wonder when explanations of the second kind are preferable to those of the first, and vice versa. Of course, there may be a third class of cases where both neurobiological and folk psychological causes are

---

297 Lewis, “Psychophysical and Theoretical Identifications,” 256
operative—nothing in my argument suggests otherwise. But occasionally folk psychological explanations are wholly inappropriate, and I actually think that these occasions are far more frequent than we might initially think. But determining which cases these are is a problem that dualism offers almost no assistance in ameliorating in any way.

A more fundamental issue, beyond these conflict cases, is that folk psychological explanations almost always include, or directly entail, cross-attribute explanations (mental states explaining actions, or physical states explaining mental states). Spinoza’s Isolationist, therefore, regards all such explanations as false, though there may something in the neighborhood of such explanations he can accept as true. If the Isolationist can show that the dualist has systematic difficulties with describing appropriate causes in cases like Depression, therefore, mostly as a result of the cross-attribute, often conflicting character of folk psychological explanations (i.e. conflicting with neurobiological explanations), this lends further credence to the Isolationist’s blanket rejection of such cross-attribute explanations. Of course, the inference cannot be that folk psychological explanations are never legitimate, from the fact that they sometimes fail. Rather, the thought is that folk psychology fails fairly spectacularly, and fairly systematically, so that it offers little help to the dualist who wants a response to the problem of pathological causation. But before showing that the dualist does have these systematic difficulties and explaining how this jeopardizes her explanatory framework, I want to briefly sketch how my focus on pathological causation connects with broader discussions of folk psychology.
4.3.1 The Literature on Folk Psychology and the Mind-Body Problem

Most philosophical reflection on folk psychology is not explicitly concerned with the question of pathological causation that, I have been arguing, interests Spinoza in his argument for the explanatory barrier. That the FP literature is recent (i.e. 20th and 21st century primarily), and thus takes dualism much less seriously as an option in the metaphysics of mind, probably partially explains this neglect. Still, several attempts to explain folk psychology’s relationship with the mind-body problem, broadly construed, do exist, and I think they can helpfully illuminate my concerns.

As Stich and Nichols tell the story, two problems constitute Cartesian dualism’s legacy in philosophy: the interaction problem and the problem of other minds. The interaction problem is extremely puzzling in its own right, and if one accepts dualism, they claim, the problem of other minds is intractable. In Stich and Nichols’ version of early 20th century philosophy, verificationists argued that meaningful empirical claims must be verifiable by actual or possible public observation statements. Therefore, a Cartesian non-spatial substance could not have beliefs, desires, or pains, insofar as those terms retained their meanings. For no actual or possible public observations could verify them. This, according to Stich and Nichols, shifted the mind-body debate towards questions of meaning, and when verificationism started to look hopelessly circular, other theories of meaning attained greater prominence, including, perhaps most importantly, David Lewis’s functionalism. Stich and Nichols then draw their first connection to folk psychology:

298 What follow summarizes Stich and Nichols description in “Folk Psychology,” 235-240. When I use exact quotes, I cite directly.
Since the account maintains that mental state terms are given by functional definitions, the view is often known as functionalism. We can now see one reason why philosophers of mind have been concerned to understand the exact nature of commonsense (or folk) psychology. According to functionalism, folk psychology is the theory that gives ordinary mental state terms their meaning.²⁹⁹

The second, related reason they identify follows somewhat from the first. Since functionalism is taken seriously as a theory of the meanings of mental state terms, and since functionalism appears to treat FP as an empirical theory, the theory might turn out to be erroneous.

Paul Churchland is perhaps the most rabid defender of the claim that FP is not merely possibly, but is actually, a false empirical theory. And he too has a story to tell about how FP relates to the questions we are concerned with here, as well as some of the risks of reliance on FP. He says:

Seeing our commonsense conceptual framework [FP] for mental phenomena as a theory brings a simple and unifying organization to most of the major topics in the philosophy of mind, including the explanation and prediction of behavior, the semantics of mental predicates, the other-minds problem, the intentionality of mental states, the nature of introspection, and the mind-body problem.³⁰⁰

Churchland’s thought is that, once we see FP as a theory, we can use it to provide the commonsense laws linking behavior and our explanations of it (note the direct connection to ordinary vs. pathological causation). At first, his characterizations sound largely positive. For instance, we can provide a functionalist account of mental state terms using FP, as Stich and Nichols note. Moreover, FP provides a solution to the problem of other minds, since “the conviction that another individual is the subject of

²⁹⁹ Stich and Nichols, “Folk Psychology,” 237
³⁰⁰ Churchland, “Eliminative Materialism and the Propositional Attitudes,” 68
certain mental states…is a *singular explanatory hypothesis* of a perfectly straightforward kind. Its function, in conjunction with the background laws of folk psychology, is to provide explanations/predictions/understanding of the individual’s continuing behavior, and is credible to the degree that it is successful in this regard over competing hypotheses.”  

Churchland introduces and seems to defend something very similar to what I suggest on the dualist’s behalf: FP explanations operate on a higher level, independent of the underlying ontology, to secure some desired anti-skeptical conclusions, including realism about other minds and reliable explanations for behavior caused ordinarily, rather than pathologically. In the terms of an example just given, the explanation in Beer works regardless of whether beliefs and desires should be characterized in dualist or physicalist/Spinozist terms. The explanation floats free of ontological disagreement.

Finally, Churchland suggests that the major positions in the metaphysics of mind can be classed according to what they predict about the ontology of FP relative to that of completed neuroscience, or completed physics. The dualist, according to Churchland, will think that FP’s ontology will not reduce smoothly to that of completed neuroscience or physics, while identity theorists predict the reduction and functionalists deny it on the grounds that reduction is incompatible with the multiple realizability of mental state terms in different physical substrates. From several perspectives, then, I hope to have shown how folk psychology relates to the issues we are concerned with here. When FP is interpreted as a theory, it provides an alternative to the positing of robust but wholly

---

301 Churchland, “Eliminative Materialism and the Propositional Attitudes,” 69

302 Churchland, “Eliminative Materialism and the Propositional Attitudes,” 72
unknown psychophysical laws for the dualist to generate reliable explanations of behavior and of the ordinariness of causation. Interpreting FP as a theory in this way also unifies several important questions in the philosophy of mind, and, insofar as FP’s explanations are reliable, it allows the dualist to separate her ontological commitments (whether FP reduces to completed neuroscience) from her explanatory ones. This all initially looks quite positive for the dualist. Of course, Churchland eventually pulls the rug out from under folk psychology, which significantly weakens the utility of the foregoing for the dualist’s purposes. In the next section, I want to offer qualified endorsement of some of Churchland’s arguments, and to situate this endorsement in the context of this chapter’s overall arguments.

4.3.2 Folk Psychology and Explanation: Premises 5b-6 of the FP Psychophysics Argument

In the previous sections, I built on my arguments from chapter three and tried to say a bit more in favor of what I called, at the beginning of this chapter, (4) interactionist dualism implies either that there are psychophysical laws, or that commonsense or folk psychology can reliably generate explanations for why causation is ordinary, not pathological. I suggested these possibilities as alternatives to providing causal pairing relations or to arguing that no one—physicalists, Spinozists, or dualists—can answer the pairing challenge. I also said a bit about why I believe 5a—there are no psychophysical laws, and 5b—commonsense psychology lacks the features required of a theory for

---

303 Those arguments occurred on pp. 193-199, and the previous sections’ material elaborates on or simply reiterates some of the conclusions pressed earlier.
generating reliable explanations. I now want to defend 5b at greater length—FP is not a
theory capable of generating sufficiently reliable explanations—with the help of
arguments stated most forcefully, perhaps, by Churchland. Still, others who lack
Churchland’s distinctive commitments endorse them as well. I further suggest that we
might find similar reasoning in Spinoza’s Ethics.

In endorsing Churchland’s arguments, I do not thereby endorse his most well-
known conclusion, that of eliminative materialism, which claims that FP is such a
radically false theory that both its explanations and its ontology of beliefs and desires
require replacement.\textsuperscript{304} I think we have good independent reasons for being realists about
beliefs, desires, emotions, etc. and moreover, I am here defending Spinoza, who is clearly
not an eliminative materialist. For example, notice that Ethics 2p1 \textit{first} seeks to establish
that Thought is an attribute of God, and only in 2p2 does Spinoza defend the same claim

\textsuperscript{304} Ravenscroft, “Folk Psychology as a Theory,” for example, has interpreted Churchland as employing the following, probably invalid, argument. 1: Beliefs and desires are the posits of folk psychology. 2: Folk psychology is false. 3: The posits of false theories do not exist. 4: Therefore, beliefs and desires do not exist. As many, for example, Kitcher, “In Defense of Intentional Psychology,” and Von Eckhardt, “Folk Psychology,” 300-307, as well as Ravenscroft, “Folk Psychology,” have noted, this argument might be invalid if there are FP-independent reasons for positing beliefs and desires, or said differently, if beliefs and desires are the posits of two different theories on independent grounds. From another direction, Stich and Nichols, “Folk Psychology,” 360-365, have argued that FP can be interpreted in two different ways. On the \textit{mindreading} sense, FP consists of a set of capacities for predicting, explaining, and attributing mental states and behavior on the basis of such mental states, while on the \textit{platitude} sense FP is a set of generalizations that systematize commonsense psychological platitudes in a perspicuous way. Moreover, they claim, the mindreading sense admits of two interpretations. Churchland and others have interpreted our mindreading capacities as the application of general psychological laws to particular circumstances. But recent simulation theorists have argued for a different interpretation (Stich and Nichols, “Folk Psychology,” 363): “Mindreading can be explained as a kind of mental simulation that requires little or no information about how the mind works (Gordon, 1986; Heal, 1986; Goldman, 1989; Harris, 1992). If these simulation theorists are right, and if we accept the mindreading account of folk psychology, then \textit{there is no such thing as folk psychology}…This would be bad news for eliminativists, since if there is no such thing as folk psychology then their core argument—which claims that folk psychology is a seriously mistaken theory—has gone seriously amiss.”
for Extension. Indeed, as I have suggested and will argue for at great length in the following chapters, one significant advantage of Spinoza’s position is his ability to offer compelling arguments against physicalist as well as dualist positions. However, Spinoza would, I think, be highly receptive to certain aspects of Churchland’s arguments, as applied to dualism specifically, rather than to realism about beliefs and desires in general. I already argued in 4.2 that an explanatory asymmetry in favor of Spinozist/physicalist explanations remains, as regards the cluster of issues surrounding the pairing problem and Uniqueness and Privilege. If we are not justified in believing in irreducible psychophysical laws, then the best remaining dualist strategy consists in arguing that folk psychological explanations employed by both dualists and their opponents can adequately explain why causation is ordinary not pathological.

But the dualist is wrong—her ontological and explanatory commitments are not so easily separable because some, indeed many, folk psychological explanations are false or misleading. Cases like Depression are just the tip of the iceberg. As we saw before, underlying ontological disagreements come to the fore when adjudicating conflicts, either between competing folk psychological explanations or between FP explanations and alternate physical causes. In such conflicting cases, the Spinozist and the physicalist are once again much better equipped than the dualist is to explain the phenomena.

---

305 For this reason I don’t have to answer counterarguments to eliminativism that claim, for example, that FP cannot be eliminated because no purely descriptive theory, of the sort Churchland favors, can account for FP’s normative character. And many other such counterarguments are ineffective when Churchland’s arguments are marshaled in favor of a thesis much less radical than eliminativism.
What then are Churchland’s central arguments for the falsity of FP, and does Spinoza share similar concerns? Churchland highlights three central challenges for FP, the first and third of which will be most important for my argument. He says:

Even so, the presumption in FP's favor is spurious, born of innocence and tunnel vision. A more searching examination reveals a different picture. First, we must reckon not only with FP's successes, but with its explanatory failures, and with their extent and seriousness. Second, we must consider the long-term history of FP, its growth, fertility, and current promise of future development. And third, we must consider what sorts of theories are likely to be true of the etiology of our behavior, given what else we have learned about ourselves in recent history. That is, we must evaluate FP with regard to its coherence and continuity with fertile and well-established theories in adjacent and overlapping domains—with evolutionary theory, biology, and neuroscience, for example—because active coherence with the rest of what we presume to know is perhaps the final measure of any hypothesis.306

According to both Churchland and Spinoza, I argue, tunnel vision prevents us from seeing subtler factors at work that undermine FP’s explanatory relevance in a vast majority of cases. Moreover, it’s not just that FP fails to provide the whole explanatory story. Many times FP explanations actively compete with, and perform substantially worse than, the stories told by, for example, biology and neuroscience. For example, think once again about purely physical explanations of depression, solely in terms of neurobiology, as contrasted with explanations of depression that appeal to the beliefs, desires, and life experiences of the person who is depressed. But the explanations in biology and neuroscience do not appeal to any irreducibly mental phenomena, but only to physically characterizable states, as causes and explanations.307 When folk psychological

306 Churchland, “Eliminative Materialism and the Propositional Attitudes,” 73

307 Again, this is not to say that biology and neuroscience only appeal to states characterized in terms of fundamental physics—reductionism is not presupposed by the argument here. They simply eschew explanations in terms of immaterial causes or states, as a methodological presupposition (though of course it is no part of biology or neuroscience to state that such immaterial causes are impossible).
explanations compete with neurobiological ones, I suggest, the dualist is hard-pressed to
determine in a principled way why and in which circumstances one sort of explanation
takes precedence over the other.

To see the scope of FP’s predictive and explanatory limitations, first consider two
two examples, one of which we have seen before.

**Depression:** When [A] someone desires to be happy, believes that having certain
positive experiences and avoiding certain negative experiences is conducive to
this happiness, and is mostly happy at time $t$, and does not have too many or the
most serious sort of negative experiences or interactions, [B] she tends with high
probability to remain mostly happy after time $t$, assuming the overall balance of
experiences remains roughly as before. Often, [B] occurs because of [A].

**Sleep:** Someone will [B] sleep, with high probability, mostly when [A] they
desire to rest, believe that resting will contribute to improving their overall
functioning, and receive sensory stimuli compatible with their successfully
sleeping (i.e. dark room, a bed, etc.). Often, [B] occurs because of [A].

Both **Depression** and **Sleep** are wildly misleading or straightforwardly false explanations
for someone’s state of happiness or propensity to fall asleep in many cases. Undoubtedly,
further refinements of these folk psychological explanations could make them more
relevant to a greater number of cases. But that is irrelevant, for I do not argue that further
Chisholming cannot improve similar folk psychological explanations, perhaps a great
deal. Rather, in many ordinary cases, the folk psychological explanation will be wholly
inappropriate. For many individuals, their mood depends primarily on the balance of
neurotransmitters in their brain, nearly independently of external circumstances or any of
their implicit or explicit beliefs or desires. For narcoleptics, who are found to be 25 to 50
of every 100,000 individuals according to a 2007 study, the propensity to fall asleep is
nearly independent of any external stimuli, or of any beliefs or desires concerning a need for rest that the narcoleptic may have.\textsuperscript{308}

**Depression** and **Sleep** are just two of very many mental phenomena poorly explained with the resources of folk psychology. So, for example, folk psychology has trouble accounting for the nature of mental illness and its likely behavioral manifestations, creativity, the functions of sleep, the ability to represent complex 3D spaces based on 2D retinal projections, memory, perceptual illusions, and learning itself.\textsuperscript{309} To Churchland’s list we might add FP’s inability to explain the subtle factors that influence our desires, emotions, and choices, as well as those factors that dictate our ability to explain our own or others’ future preferences.\textsuperscript{310}

Spinoza, while focusing on the mind in Part II of the *Ethics*, mentions several of these issues himself, including those underlined above. For instance, in explaining how error consists in the privation of knowledge, he gives two examples (2p35s):

Men are deceived in that they think themselves free…an opinion which consists only in this, that they are conscious of their actions and ignorant of the causes by which they are determined…Similarly, when we look at the sun, we imagine it as about two hundred feet away from us, an error which does not consist simply in this imagining, but in the fact that while we imagine it in this way, we are ignorant of its true distance and of the cause of this imagining.

Perceptual illusions, as well as the subtle factors affecting our choices, are taken not only to deceive us when we make judgments or attribute motives to ourselves or to

\textsuperscript{308} See, for example, Longstreth et. al., “The Epidemiology of Narcolepsy.”

\textsuperscript{309} The examples are due to Churchland, “Eliminative Materialism and the Propositional Attitudes,” 73.

\textsuperscript{310} Stich and Nichols, “Folk Psychology,” 245-255 discuss several of these problems for slightly different purposes. There is a particularly extensive literature on the way that subtle ordering and presentation effects influence our choices that seems quite poorly captured in terms of FP.
others. They are also quite resistant to correction (2p35s): “If later we come to know that it is more than six hundred diameters of the earth away from us, we nevertheless imagine it as near.” This is true not only for perceptual illusions, but more fundamentally for our knowledge of our own motives, beliefs, and desires (2p23, 2p28): “The mind does not know itself, except insofar as it perceives the ideas of the affections of the body…the ideas of the affections of the human body, insofar as they are related only to the human mind, are not clear and distinct, but confused.”

An extreme case of this wholly general inability to adequately discern the causes of our actions, or those of others, arises with quite common forms of mental illness. To take just two of myriad examples in more detail, consider disassociative identity disorder, on one hand, and depression, on the other. The DSM 5 defines the first criterion of disassociative identity disorder as follows: “Two or more distinct identity or personality states are present, each with its own relatively enduring pattern of perceiving, relating to, and thinking about the environment and self.”

Siri Hustvedt focuses on both these examples in her discussion of the mind-body problem in light of contemporary neuroscience. She says:

The fact that this disorder [DID] turned into an epidemic in the 1980s has often been read as testimony to its ‘fake’, ‘manufactured’, or ‘unreal’ character…[But] research has confirmed that there are physiological differences between one personality and another in people with disassociative identity disorder (DID). They include allergic sensitivities—one personality has hay fever; another

---

311 I might as well note here that I am not the first person to see a serious challenge to distinctively Cartesian views of dualism in the phenomena of disassociative identity disorder, as well as the split-brain cases more famous in the literature on ethics and personal identity. Hasker, “Persons and the Unity of Consciousness,” develops split-brain and DID cases into an argument against Cartesian forms of dualism and in favor of his own emergent dualist view. His argument differs in most of its particulars from mine, and his ultimate position is one I am not congenial to, but the basic insight is right on.

312 “Disassociative Identity Disorder,” in Diagnostic and Statistical Manual of Mental Disorders 5
doesn’t—differences in endocrine function, skin arousal, color vision, and varied responses to the same medicine.\textsuperscript{313}

The conflict between folk psychological and neurobiological explanations is particularly sharp with DID, for the dualist. Suppose, with Descartes, that interactionism requires, rather than precludes, tight relationships between the immaterial and material, between mind and brain. Still, in some cases, we are justified in privileging explanations of disassociative identity disorder in terms of beliefs and desires—as when someone fakes the syndrome to receive attention or special treatment. Other times a physiological explanation seems absolutely necessary, as in ‘more genuine’ cases of the disorder. There, radical differences in allergic sensitivity, endocrine function, and medication responsiveness seem to demand some biological explanation. For the Spinozist or physicalist, beliefs and desires are just one subset of physical causes among many. They can be explored, of course not infallibly, with the resources of psychology and neuroscience, and they can be trumped in the order of explanation when other physical causes—those arising in other areas of the brain not responsible for beliefs and desires, for example—are more relevant to making accurate predictions of behavior. By contrast, the dualist will sometimes want to give a folk psychological explanation of DID in terms of beliefs and desires. That is, sometimes the dualist will say:

\textbf{Fake DID}: When someone [A] \textit{desires} attention/achievement of other goals, and \textit{believes} that acting as if two or more distinct identity and personality states are present in the same individual, and receives sensory stimuli suggesting that these actions will contribute to achieving her goals in the given circumstances, then [B] she will act in so-and-so ways. Often, [B] occurs because of [A].

\textsuperscript{313}Hustvedt, \textit{A Woman Looking at Men Looking at Women}, 205; her citations for such phenomena come from Coons, “Psychophysiological Aspects of Multiple Personality Disorder,” 47-53 and Barlow and Durand (eds), \textit{Abnormal Psychology}, 201-207.
Other times, of course, the dualist will want to concur with the physicalist/Spinozist in giving a more straightforwardly physiological explanation, and the FP-explanation of **Fake DID** will be inappropriate. But folk psychology itself does not indicate which circumstances are which. But whereas for the Spinozist/physicalist the explanations are of the same, physical, sort in either case (though they are not, of course, *the same* physical explanations), the dualist must explain why two radically different sorts of explanation are used at different times. Ideally, she must also justify her reliance on one sort or another in different circumstances. But I do not think dualists can give any principled reason for preferring FP or physiological explanations in the cases where they do, in fact, prefer one or the other.

Moreover, it becomes clear in such conflict cases that physicalists/Spinozists do not actually need to use FP at all—on their view, there is an explanation that employs only beliefs and desires characterized wholly physically, even if in practice they cannot yet make good on this promissory note as an explanation we could actually use. This is okay, though, since they are *defending the existence of* some wholly physical explanation, rather than its epistemic accessibility or utility for our purposes. Dualists, by contrast, include as part of their philosophical position that two sorts of explanation—the folk psychological and the merely physical—are appropriate at different times. But the dualist’s promissory note—that there is some non-ad-hoc justification for preferring one to another *that is explanatorily motivated on the assumption that dualism is true*—is far less likely to be cashed. Again, the question is not whether the dualist can pragmatically use FP explanations in some cases and wholly physiological explanations in others.
Rather, the challenge for the dualist is to explain why they favor, and *more importantly why they are justified in favoring*, certain explanations in some circumstances but not others. The dualist must respond to this challenge in light of her inability to provide the causal pairing relations, or the psychophysical laws, that offered alternative possibilities to relying on folk psychology within her explanatory framework.

Depression offers related challenges. The DSM 5 states: “The common feature of all these disorders is the presence of sad, empty, or irritable mood, accompanied by somatic and cognitive changes that significantly affect the individual’s capacity to function. What differs among them are issues of duration, timing, or presumed etiology.”314 The presence of both somatic and cognitive changes, to reiterate, is of course formally compatible with the interactionist dualist’s picture. But it is far less clear that FP can give explanations that respect the latter bit of the DSM’s definition: divergent etiologies for various depressive disorders. Again, some cases of depression will seem much more amenable to folk psychological explanations, where events and experiences in the patient’s life will have specific consequences that result in the patient’s depression. For others, however, an almost purely physiological explanation in terms of neurotransmitter dysregulation makes more sense. Again, FP itself gives zero guidance as to which is which. Moreover, giving FP explanations as if the first sort of depression obtains, when it is really the second, purely physiological sort, leads to misunderstandings, actively bad advice, and a fundamental disrespect for the suffering person. “Just think positively, try to make changes in your life to create conditions X, Y, Z” is horrible advice if what one needs is an SSRI.

314 “Depressive Disorders,” in *Diagnostic and Statistical Manual of Mental Disorders 5*
Once again, the physicalist and the Spinozist, when confronted by such cases, can with justification drop folk psychology from their explanations entirely, in favor of wholly physical explanations in both cases. The dualist, by contrast, must keep folk psychology in some cases, but not others, and there must be some justification for her doing so, again in light of the dualist’s failure to answer the pairing challenge or to develop a psychophysical science. As Churchland concludes: “FP is at best a highly superficial theory, a partial and unpenetrating gloss on a deeper and more complex reality.”\textsuperscript{315} We might add: if this partial and unpenetrating gloss on a deeper reality is an essential part of the dualist’s explanatory framework—as it seems to be in light of the pairing challenge and the absence of a developed psychophysics—then the dualist remains at a serious explanatory disadvantage relative to her physicalist and Spinozist opponents. The dualist’s need to rely on folk psychology at crucial explanatory junctures, plus her inability to explain when and why this occurs in any principled way, is the crucial asymmetry between her position and her opponent’s.

Folk psychological explanations are often sufficient for large swaths of everyday life, when behavior seems predictable and causes seem ordinary. But FP’s wholesale irrelevance to a whole host of mental phenomena, as well as its active obscuring of adequate explanations when it conflicts with more straightforwardly physical explanations of phenomena like sleep, mental illness, and others, is costly for the dualist. The dualist wanted to use the mutual acceptability of some folk psychological explanations to make irrelevant the underlying ontological disagreements with the physicalist or Spinozist in explaining why causation is ordinary not pathological. But

\textsuperscript{315} Churchland, “Eliminative Materialism and the Propositional Attitudes,” 74

267
such FP explanations, it turns out, are dispensable from her opponents’ explanatory frameworks but not from the dualist’s. I think this constitutes a serious challenge for the dualist who relies on folk psychological explanations to justify her confidence in ordinary causation.\(^{316}\)

In this chapter, I pursued two lines of attack against the dualist on Spinoza’s behalf. The basic argument, as laid out last chapter with reference to Spinoza’s objections to his own position in the *Short Treatise*, is that interactionists cannot account for our knowledge that causation is ordinary, rather than pathological. That is, on the assumption that dualism is true, they cannot explain causation’s being ordinary rather than pathological. In pursuing the first line, I sought to develop and evaluate the causal pairing problem as a limited instance of the threat of pathological causation. The pairing problem is not a conclusive objection to dualism. But it is unclear that it affects physicalists and Spinozists to the same extent that it affects dualists, and it is certainly true, as I argued, that Spinoza explains Uniqueness and Privilege better than his dualist opponents. In developing the second line, I sought to show that the absence of reliable psychophysical laws (and concomitant methods for discovering them), as well as the explanatory failures of folk psychology, offer the dualist little assistance in her attempt to generate reliable explanations for ordinarily caused, as opposed to pathologically caused, phenomena.

The cumulative effect of these two lines of inquiry, I suggest, is that Spinozists and perhaps physicalists have a significant advantage in explanatory capacity with respect to the question of whether causation is ordinary, rather than pathological.

\(^{316}\) Not to mention, as Churchland does, that dualists might also want to use folk psychology to justify a non-skeptical solution to the problem of other minds.
Nonetheless, I have been narrowly focused on problems that concern Spinoza and that motivate his introduction of the explanatory barrier between the attributes in the *Ethics*. I have not evaluated any of the most powerful positive arguments in dualism’s favor, though some of these will be discussed in the following chapters. But, as I mentioned earlier, in spite of Spinoza’s alignment with the physicalist *on the issues of this chapter*, Spinoza nevertheless has compelling reasons for siding with the dualist and against physicalists on a whole host of other issues to be discussed shortly. Spinozism is a middle path between dualism and physicalism, and in that spirit I now turn to Spinoza’s engagement with the latter.
CHAPTER 5:
SPINOZA AND PHYSICALISM

5.1 Spinoza the Physicalist?

In the *Short Treatise*, Spinoza describes an objection from what I have been calling the threat of pathological causation. I have suggested that Spinoza believed, upon reflection, that he could not answer this objection without introducing the explanatory barrier between the attributes, which he then does at the beginning of the *Ethics*. More carefully, independent of Spinoza’s own motivations, his texts can plausibly be interpreted as following this progression. Then, after trying to show how this tersely worded thought from the *Short Treatise* might be turned into an argument for the EB, I developed and evaluated this argument as an attack on interactionist dualism from two distinct but complementary directions. I concluded that, while neither the causal pairing problem nor the explanatory inadequacies of folk psychology or psychophysics are by themselves fatal, these problems conjoined seriously inhibit the dualist’s ability to explain why causation is ordinary, not pathological.

In introducing my reconstruction of Spinoza’s argument, I emphasized that Spinoza ultimately seeks to criticize both dualism and physicalism, but that his difficulties with each take drastically different forms. One problem with interpreting Spinoza as arguing against the latter group is that commentators have often read Spinoza,
sometimes explicitly but often implicitly, as at least some brand of physicalist-materialist. So, for instance, Jacobi fires the opening salvo in the ‘pantheism controversy’ in German idealism by arguing: “Consistent philosophy is Spinozist, hence pantheist, fatalist, and atheist.” Materialism, of course, does not appear on Jacobi’s list. But the presence of prima facie conflicting designations—pantheist and atheist—suggests that he might have welcomed the inference, since one of Jacobi’s reasons for the atheism charge is that God is identified with the extended whole of Nature. Several Soviet thinkers saw in Spinoza an essential precursor to Soviet scientific materialism: “[Spinoza is] essentially a great atheist and materialist.” Atheism and materialism, of course, are popularly (though of course not conceptually necessarily) conjoined, so we might add Steven Nadler’s defense of reading Spinoza as an atheist, rather than a pantheist, to the list of those who might implicitly accept the materialist charge.

In the Anglo-American tradition, Stuart Hampshire at least occasionally represents Spinoza as a kind of materialist. And Jonathan Bennett, while insisting that Spinoza is in fact a panpsychist, clearly appears pretty wistful about this fact, and often seems to suggest that Spinoza has the philosophical resources to defend a version of materialism.

---

317 His claim is cited in di Giovanni, "Friedrich Heinrich Jacobi,” Introduction. Jacobi’s claim appears in The Spinoza Conversations Between Lessing and Jacobi.


319 Nadler, “Spinoza the Atheist,”

320 See, for example and fittingly enough, Hampshire, “A Kind of Materialism.”
If Spinoza did not have that doctrine of panpsychism, his naturalism would commit him to some kind of materialism…Unlike Leibniz, who was catapulted across much of the philosophy of mind by his doctrine that all substances are comparable to human minds, Spinoza keeps his feet on the ground, despite his panpsychism: he travels overland, the hard way, and we can learn a lot by following his steps.321

Bennett’s second sentence is instructive, for he explicitly contrasts two notable idealist/panpsychists of the early modern period, Leibniz and Spinoza, with reference to the relative groundedness of their respective systems. This groundedness, in turn, seems tacitly identified with physicalist naturalism. So while Bennett does interpret Spinoza as confronting many of the central problems of the philosophy of mind head-on, Spinoza nonetheless “lets himself off from that task [of showing how a mostly non-mental world could generate small pockets of mentality through the same physical laws that rein throughout the rest of Nature] by introducing mentality as a basic feature of the entire universe.” Bennett’s interpretation of Spinoza’s panpsychism as dereliction of materialist duty should be no surprise, however. Bennett, remember, forcefully defends a number of the central challenges to the tenability and justifiability of the explanatory barrier, several of which were discussed in the first chapter.

There is certainly something to this trend in Spinoza interpretation. Spinoza cannot, however, be even a (mere) panpsychist, much less a physicalist, since God is the being with infinitely many attributes (1d6), and since each mode can be conceived in each of these infinitely many fundamental ways. Not only does mentality pervade nature at the most basic level, but also there are many basic ways of being, unknown to human beings, which are similarly pervasive. Still, despite the overwhelming textual evidence

321 Bennett, *A Study of Spinoza’s Ethics*, 38
against the physicalist interpretation, Spinoza undoubtedly harbors a far greater spiritual affinity with physicalism than dualism. Their shared capacity for monistic, unified explanations undoubtedly contributes to this affinity, for within a given domain (the physical, say), the EB prohibits any appeals to mental causes or effects. So, in particular, all bodily actions, which constitute the vast majority of our engagement with other human beings, are to be explained in purely physical terms.

Consequently, my arguments in this chapter reflect Spinoza’s closeness to physicalism in several respects, and thus take a different form than those in the last. In the first part of this chapter (5.2 & 5.3), I seek to accommodate Spinozism to physicalism. That is, I argue that Spinoza can accept a version—which often captures the core insight—of the strongest arguments for physicalism. Undoubtedly, the physicalist can exclude Spinoza’s position by definition, and so formulate her position independent from his, but the extant arguments do little to distinguish between them, or so I argue here.

Then, in the chapter’s second part (5.4), I mention a host of areas where Spinoza seems to have an explanatory advantage over the physicalist, either generally or with reference to the sorts of physicalists to whom Spinoza is most frequently compared. This section comes closest to pursuing the same strategy as the previous chapter does against the dualist, but the differences are substantial. At least in this chapter, I will mostly point out typical objections to physicalism that Spinoza can avoid or mitigate more successfully than the physicalist herself can.\(^{322}\) After briefly mentioning these other problems, I then focus my attention on the physicalist’s prospects for providing an acceptable account of

\(^{322}\) Note: I will not be discussing, in this chapter, the most well-known arguments recent dualists have offered against physicalism, such as the knowledge and conceivability arguments. However, I will discuss the conceivability argument briefly in the last chapter.
intentional content. I argue that this is one challenge Spinoza does not have to meet, which may be a significant advantage if the physicalist’s difficulties remain as intractable as they currently seem. The final chapter then continues this trend and combines these insights. There I show how Spinozistic principles and insights can provide a powerful foundation for the currently popular research program examining the viability of various forms of so-called ‘Russellian panpsychist monism.’

5.2 Spinoza and Three (Not So Good) Arguments for Physicalism

Physicalism has been the dominant worldview in Anglo-American philosophy for much of the past seventy years, but actual arguments for the position, though perhaps powerful, are few.\textsuperscript{323} It’s tempting to think that widespread acceptance of physicalism follows primarily from overwhelming dissatisfaction with interactionist dualism, faith in or “worship of the natural sciences,” particularly physics, and a certain lack of imagination in envisioning alternatives.\textsuperscript{324} Nevertheless, we can identify several (specifically five, discussed in this chapter overall) arguments for physicalism, as well as a sixth, related issue (supervenience) that can probably be turned into an argument without too much trouble. Of course, there may be other arguments that I have missed.

\textsuperscript{323} According to the 2009 PhilPapers survey, which was taken after approximately 15 years of concerted attacks, by Chalmers, Jackson, and others, on physicalism as the dominant worldview in the metaphysics of mind, 64.2\% of philosophers still identify as physicalists. This suggests, to my mind at least, that the numbers would have been even higher beforehand, though of course this is interesting sociology, rather than evidence for any argument I am giving. See the results of the survey here: https://philpapers.org/surveys/metaresults.pl

\textsuperscript{324} See Burge, “Mind–Body Causation and Explanatory Practice,” for the claim that physicalism is based on science worship. Lycan, \textit{Consciousness}, gives some standard reasons for embracing physicalism because of problems with dualism, but Lycan, “Giving Dualism Its Due,” admits that science worship rather than these arguments probably better explains his position. For some inkling that lack of imagination on these questions isn’t universal, see Gibb, “Mental Causation as Double Prevention,” and Lowe, “Substance Causation, Powers, and Human Agency,” as well as Wilson, “Supervenience Based Formulations of Physicalism,” and others.
Still, I think that these represent the most popular and actually motivational arguments for physicalism in the literature. In this section and the next (5.2 and 5.3), I want to suggest simply either that the advertised argument for physicalism is not successful, or that Spinoza can happily accept it in some form, or (sometimes) both.

Mid-century physicalists like J.J.C. Smart or Ernst Nagel gave an inference to the best explanation argument for physicalism (IBE Physicalism). Lycan describes the argument thus: “[1] It is reasonable to think that every mental state or event at least has a corresponding type of brain state or event. [2] In general: When Xs are invariably accompanied by Ys and you can find nothing to distinguish Xs from Ys, the best, because most parsimonious, explanation is that the Xs just are the Ys.] [C] The best, because most parsimonious, explanation of those correlations is that the mental states/events just are the ‘corresponding’ brain states/events.” There are two devastating problems with this move though. First, considerations of parsimony seem to play a role only very late in the game, once all other things are equal, and it seems unlikely that non-physicalists will agree, or that physicalists have shown, that this is the case here. Second, it seems like the mental type-physical type identities whose materialization Smart appears to be counting on (in the first premise) never appeared. Lycan continues: “Whatever is in common as between all human beings who believe that a Frenchman has been assassinated in

---

325 I will not be discussing Davidson’s, “Mental Events,” arguments for anomalous monism, but just note that Davidson is admittedly influenced by Spinoza in formulating these arguments, and thus that there is liable to be a great deal of overlap between his position and Spinoza’s (which I will nonetheless not argue for explicitly here). Koistinen, “Causality, Intensionality, and Identity,” and Della Rocca, Representation and the Mind-Body Problem both take similar stances on the issue of overlap between Davidson and Spinoza, however.

326 My discussion of Smart/Nagel and Armstrong/Lewis follows Lycan, “Giving Dualism Its Due,” 552-555 closely.
Trafalgar Square (to take an old example of Dennett’s), that feature could not possibly be characterized in neuroscientific terms.”  

Finally, and most relevantly for my purposes, Spinoza can accept the argument, and indeed gives a version of it, in 1p10, which we saw in chapter one. There he says: “It is evident that although two attributes may be conceived to be really distinct... we still cannot infer from that that they constitute two beings, or two different substances.” Later, in 2p7, he explicitly draws Smart’s conclusion, that the mental events are the physical events (he speaks in terms of modes, not events, however): “So also a mode of extension and the idea of that mode are one and the same thing, but expressed in two different ways.” IBE Physicalism thus poses no threat to Spinoza. The implied further claim, that the mental is derivative or that the physical is somehow more fundamental, remains absent from at least this argument, and thus Spinoza faces no pressure to accept it here.

One might think, however, that Spinoza has multiplied something here without necessity, and thus faces pressure to explain why he posits distinct mental and physical causal orders, even if individual nodes in the causal matrices are claimed to be identical. Three short responses will have to be sufficient for our purposes here. First, the above argument, against employing simplicity considerations prematurely, remains in full force. The question is not whether causal orders have been multiplied, but whether they have been multiplied without necessity. Second, Spinoza’s reasons for seeking multiplicity here are far more general than buttressing his particular solution to the mind-body

327 Lycan, “Giving Dualism Its Due,” 553

328 In discussing a more compelling argument for physicalism, which I label the Causal Closure Argument, I will have to speak more precisely about the relationship between modes, events, properties, etc. in Spinoza’s ontology. For now, though, I assume that we can translate Spinoza’s position to contemporary terminology without much difficulty.
problem, and concern the nature of God as the being with infinitely many attributes, each of which expresses an eternal and infinite essence. Taking issue with his solution thus requires questioning his general framework, rather than combatting some ad hoc response developed solely for purposes of this mind-body debate. Third and perhaps most importantly, even identity theorists like Smart and Nagel will need to preserve a certain multiplicity, if not of entities that of something else—properties, state types, modes of presentation, etc. In the specific version of IBE Physicalism given above, for example, Lycan formulates the premises so that a dualism of state types is retained while event identity is asserted. It is a tricky question, given Spinoza’s nominalism and his minimalist ontology of substance and modes, just how best to formulate the corresponding claim, but it is clear that Spinoza wants to affirm the conclusion he shares with Nagel and Smart: the mode of Extension (brain state or event) and the mode of Thought (mental state or event) are one and the same thing.

Armstrong and Lewis give a second, broadly functionalist argument for physicalism (Functionalist Physicalism). “[1] Pain=Whatever state of a person plays the role P (being typically caused by tissue damage, and in turn causing wincing, crying out, favoring, etc.) [We know this a priori; we have all got the concept of pain]...[2] the

329 See Newlands, Reconceiving Spinoza, especially the introduction and chapter 1, for additional discussion of the problem of the One and the Many in Spinoza.

330 This offers an interesting way of categorizing positions on the mind-body problem, for these observations suggest that nearly any view will have to go dualistic somewhere—in affirming a plurality of properties, state types, modes of presentation, etc. Moreover, most (though not all) views will want to lend this plurality some degree of objectivity—that is, they will want to claim that these different properties or modes of presentation are more than just reflections of our linguistic structure. Spinoza’s innovation is to affirm an incredibly strong version of monism, for both Substance and modes, while embracing infinite plurality in the realms of causation and explanation, across God’s infinitely many attributes.
occupied of role P=the firing of c fibres…[C] pain=the firing of c fibres QED.”^331 Once again, there appear to be at least two potentially fatal problems here. Regarding the first premise, the causal concept of pain does not appear a priori—pain’s qualitative aspects are essential to the concept, but pain’s typical causes and effects are not. Writhing, for instance, is not an a priori effect of pain. Additionally, it seems like ordinary objects could be rigged to play the typical pain roles without having any of the qualitative aspects of pain as experienced.^332 Lycan is particularly excellent on the second premise. He says: “[Premise 2] seems fine until one realizes that its first word is ‘the’. 2 begs the question against the dualist view [held by some, but by no means all, dualists] that role P is causally over-determined. The typical causes cause both neural events and immaterial pain events, and pain behavior is doubly caused by the neural events aforementioned and immaterial pain events.”^333

Now Spinoza probably cannot accept premise 1 as stated, since it seems to pretty clearly violate the EB. As 2p6d states: “So the modes of each attribute involve the concept of their own attribute, but not of another one; and so (by IA4) they have God for their cause only insofar as he is considered under the attribute of which they are modes.” But, since the argument appears to have two false and/or question begging premises, this shouldn’t matter too much. Moreover, of course, Spinoza again accepts the conclusion (recall 2p7).


^332 Campbell, Body and Mind, 100 as well as Block, “Troubles with Functionalism,” give versions of this objection. Again, thanks to Lycan for the citations.

^333 Lycan, “Giving Dualism Its Due,” 555
Turning in a more scientifically minded direction, many philosophers have argued that dualist views violate the conservation laws of energy and momentum.\(^{334}\) This objection closely relates to the Closure-principle arguments, discussed below, but is independent of them. Here, the thought is that interactionist dualism is strictly incompatible with specific conservation principles found in physics, rather than with plausible philosophical principles (like Closure) that perhaps receive empirical support from the physical sciences. Montero gives an initial characterization of this sort argument (ACE):

1. Energy is conserved in any closed system.
2. The universe is a closed system.
3. There are causal relations between the mental and physical.
4. Causation involves the transference of energy.
5. Thus, the mental is the physical.\(^{335}\)

This argument, too, appears wracked with rather serious difficulties. Versions of premise 2 that refer to the physical world, rather than the universe, as a closed system, one that neither affects nor is affected by anything outside of it, would give an argument for physicalism, Montero thinks. However, to the extent that we have evidence for anything’s being a closed system, she suggests, we have evidence that the whole universe is, rather than the physical world. The physicalist cannot simply straightforwardly assume


\(^{335}\) Montero, “What Does the Conservation of Energy Have to Do With Physicalism,” 385
the two coincide. Numerous counterexamples to causation as energy transference have been proposed, in addition, that render premise 4 unsatisfactory.  

But even if we suppose that causation is energy transference, Montero thinks, if we did witness apparent violations of the conservation of energy due to mental entities or events, it is unclear whether the scientific community would accept or be justified in accepting apparent violations as genuine ones. Inductive evidence that they might not be comes from the physicist Wolfgang Pauli’s hypothesis, in 1930, of the existence of neutrino particles to account for apparent momentum/energy conservation violations in certain radioactive processes. Like Pauli, who postulated the neutrino not to save the phenomena but to preserve the conservation laws, in the presence of a putative violation mental energy, vital energy, or some other form of new energy would simply be postulated instead. The stronger move the physicalist requires, Montero thinks, involves claiming that whatever has energy is physical. If this additional premise is added, however, we get the following (VACE):

1. Energy is conserved in any closed system.
2. The universe is a closed system.
3. There are causal relations between the mental and physical.
4. Causation [always and only] involves the transference of energy.

336 Montero, “What Does the Conservation of Energy Have to Do With Physicalism,” 390-91 discusses several of these counterexamples.

337 Both Montero and Wilson, “Supervenience Based Formulations of Physicalism,” make much of this example.

338 Papineau, “The Rise of Physicalism,” 24-25 discusses historical instances where this was actually done, although in those contexts the commitment to the conservation laws among the members of the scientific community was less robust than it is now.
5. Anything with energy is [only] physical.

6. Thus: the mental is physical.\textsuperscript{339}

Montero’s diagnosis of the remaining problem with this argument should be fairly obvious: premises 3-5 are jointly sufficient to establish the conclusion without the first two premises, where conservation of energy is explicitly invoked. So while the latter three premises might indeed constitute an interesting argument for physicalism, the argument should not be interpreted as a showdown between non-physicalists and physics. Instead, tendentious philosophical claims about causation (as exclusively energy transfer) and energy bearers (as exclusively physical) do all the argumentative work, while conservation law violation becomes a red herring. And, for the record, Spinoza can accept that anything with energy is physical; this just doesn’t exclude that same thing’s also being mental. So Spinoza’s version of mind-body identity is compatible with the argument’s conclusion. Additionally, while conservation of energy does not appear to be Spinoza’s most fundamental reason for the complete causal closure of the physical domain, he appears to affirm energy conservation nonetheless, at least on some interpretations. Stuart Hampshire, for instance, suggests:

It seems natural to translate the now unfamiliar phrase ‘Motion and Rest’ as ‘energy’; one can then represent Spinoza as in effect saying that the extended world is to be conceived as a self-contained, and all-inclusive, system of interactions in which the total amount of energy is constant; and, secondly, he is in effect saying that all the changing qualities and configurations of extended bodies can be adequately represented solely as transmissions or exchanges of energy within this single system.\textsuperscript{340}

\textsuperscript{339} I add [always] and [only] to Montero’s argument to make it more plausibly secure its desired conclusions.

\textsuperscript{340} Hampshire, \textit{Spinoza and Spinozism}, 63
On some interpretations Spinoza thus affirms the principle of causation as energy transference (localized only to Extension), and energy conservation (more generally, for Extension), which jointly play an important role in the argument Montero evaluates. Despite this similarity, I want to join Montero and others who have argued that only debatable, oftentimes dialectically difficult to establish supplementary premises can bridge the gap between conservation laws and the thesis of physicalism.341

The argument that the conservation principles directly entail the truth of physicalism is thus not a good one, I think. But that does not mean that advances in the physical sciences are irrelevant to the questions considered here. The trick is to figure out exactly how. For instance, Papineau has shown that conservation laws have an uneven reception in the history of philosophy. So while I am fairly confident that Spinoza would have embraced some conservation principles wholeheartedly, I doubt they have the appropriate explanatory status to serve as fundamental arguments for Spinozism. More basic metaphysical principles, rather than ones specific to the realm of Extension, must play that role.

Still, as Papineau tells it, Leibniz did build one of his central objections to Cartesianism around Descartes’s misunderstanding of the applicable conservation laws.

341 Averill and Keating, “Does Interactionism Violate A Law of Classical Physics,” gives an interesting reading, according to which the conservation laws come in weak and strong versions, only the weak versions of which are actually required for the practice of physics. Gibb, “Closure Principles and the Laws of Conservation of Energy and Momentum,” discusses the connection between conservation principles and the closure principles to be discussed shortly. She argues quite compellingly that 1) the conservation laws alone are certainly not sufficient to provide an argument for physicalism and 2) conservation principles might not even really support Closure principles. She thinks that further claims are required: that the only way something non-physical could affect a physical system is by supplying or redistributing energy or momentum, and that redistributing momentum or energy cannot be done without supplying some of one or both. And she argues that each is difficult to defend in the context of the argument.
Leibniz, however, formulates and employs his updated (and broadly accurate) conservation principles as part of an argument for his system of pre-established harmony (occasionally interpreted as a sort of epiphenomenalism, though I think this misses the point—the fundamentality of the mental). So the extant formulations of the conservation principles available during Spinoza’s time would have served equally well, it seems, for defending the sort of mind-body parallelism he (and Leibniz) endorses, rather than any sort of physicalism. That philosophers who know this history still take the conservation principles to provide an argument for physicalism thus seems simply to express displeasure with metaphysical systems employing pre-established harmony or broad mind-body parallelism. There’s no new argument for physicalism that excludes parallelism here.

These three arguments—IBE Physicalism, Functionalist Physicalism, and VACE—fail. Moreover, Spinoza may be able to accept a version of the first, as well as (perhaps) a version of the third whose premises are sufficiently disambiguated, which makes these arguments blunt weapons against his views in any case.\footnote{The disambiguation I have in mind involves accounting for the thicket of intensionality present in Spinoza’s system. So, for instance, Premise 3 of VACE is true on some readings: there are causal relations between mode 1, conceived as physical, and mode 2, conceived as physical, where mode 1 can also be conceived as mental. Premise 4 is perhaps true, depending on Spinoza’s views of prevailing counterexamples to energy transference theories of causation, when confined to Extension: Causation in Extension always and only involves transference of energy. And Premise 5 is true too: anything with energy is physical, but that same thing is also mental, by Spinoza’s lights.} Some VACE-like argument undoubtedly explains some of physicalism’s stunning rise in popularity in the latter half of the 20th century. But I do not think that a direct entailment from the conservation laws represents the best way of locating physicalism’s attraction. Two additional arguments, both suggested by Papineau and later refined and discussed by...
many others, do better. Indeed, I think they represent by far the strongest considerations in favor of physicalism, and I think Spinoza can happily accept both of them, yet another sign that something has gone amiss with the physicalist’s argumentative project.

5.3 Spinoza and Two Better Arguments for Physicalism

5.3.1 Papineau: Completeness, a First Pass

David Papineau, in his essay “The Rise of Physicalism,” links the eponymous rise to a growing belief in the completeness of physics. The completeness of physics serves as the crucial premise in Papineau’s argument for physicalism, and two further empirical arguments undergird acceptance of this completeness principle, in his telling. I agree that this intuition of completeness, as opposed to the conservation principles more specifically, in large part explains physicalism’s overwhelming acceptance. I want to show here that if one takes this completeness intuition seriously, as I do, then Spinoza can honor it. So if the completeness intuition motivates one’s physicalism, as I suspect it does for many, Spinoza offers an exciting alternative worth considering. Indeed, Spinoza provides an eloquent justification of this intuition on purely philosophical grounds several centuries before it becomes generally acclaimed. So I first discuss Papineau’s sketch of the first argument from the completeness of physics to physicalism, whose central premise receives support from two further empirical arguments, before finally returning to consider a second, more carefully formulated version of what I label the Causal Closure Argument for physicalism.343

343 In the more carefully formulated version given below, I will distinguish, following Bennett, “Exclusion Again,” between the various uses to which one might put an argument of this sort, and comment on each in relation to Spinoza’s views. But for now just note that the argument is sometimes
Papineau’s basic thesis is that empirical developments in the 19th and 20th centuries made the following argument’s central premise seem increasingly justified to both practicing scientists and to philosophers.

1. **Completeness of Physics**: All physical effects are fully determined by law by prior physical occurrences.
2. **Causal Influence**: All mental occurrences have physical effects.
3. **No Universal Overdetermination**: The physical effects of mental occurrences are not all overdetermined.
4. **Conclusion**: Mental occurrences must be identical with physical occurrences.\(^{344}\)

The relevant empirical developments, which make the completeness of physics increasingly persuasive, come in two flavors, one rather abstract and grounded in theoretical physics, the other a basically negative finding in physiology. Both seek to motivate the thought that there are no special mental forces affecting the physical world.

**Fundamental Forces Argument**: There is reason to believe in the conservation laws (of energy and momentum). All fundamental forces are conservative forces (i.e. they respect the conservation of energy and momentum). Either all special forces reduce to fundamental forces, or they do not. We have good reason, given recent developments referred to as a Causal or Closure argument for physicalism (and against dualistic or emergentist positions), and sometimes as the Causal Exclusion argument (as against non-reductive physicalism in favor of reductive physicalism). The first use is more relevant for my purposes, but I will say a bit about each below.

\(^{344}\) Papineau, “The Rise of Physicalism,” 9; note again that this is not the canonical formulation of the argument I will be discussing, so worries about the strength of some of the premises (i.e. the use of ‘all’ in the Causal Influence premise, or the deterministic statement of the Completeness of Physics in premise 1) should be put on hold for the moment.
in physics, for supposing that all special forces do so reduce. *Moreover, if some special mental force does not reduce to some more fundamental conservative force, then we have no reason to believe this special mental force respects conservation of energy and momentum.* But since we do have reason to believe in the conservation laws, there must be no such special force.\(^{345}\)

My presentation of the argument from fundamental forces differs from Papineau’s only in emphasis, but I think the shift in emphasis is worth noting, since it draws attention to a connection between this argument and the arguments against dualism I presented in the previous chapter. As we have already seen above, Papineau does not think, unlike the (perhaps confused) proponent of VACE considered earlier, that the conservation laws directly entail physicalism—he admits that epiphenomenalism, pre-established harmony, and other non-interactive dualisms escape any such argument.\(^{346}\) But he does think three things. The first is that we do have general, compelling reasons, based in the theory of physics, for thinking that all the special forces will ultimately reduce to the fundamental, conservative forces. Second, as we will see in the physiological argument, he thinks that *we have no evidence* for any special mental forces or occurrences not explicable in physical terms.

But this is not the end of the argument—the central explanatory defect still awaits description. For suppose these general theoretical reasons—favoring unification of all forces to the fundamental ones—are inconclusive. Then, even if the non-physicalist

\(^{345}\) My presentation follows Papineau, “The Rise of Physicalism,” 27-29, but it differs in several respects to be emphasized shortly.

\(^{346}\) Papineau, “The Rise of Physicalism,” 10

286
responds by noting that absence of evidence (of some special mental force) is not
evidence of absence (of that force), she has a further problem: she must still explain,
upon positing this special mental force, why she is justified in believing that the special
force will be conservative.

Of course, Papineau thinks, she will want to assume that it will be conservative,
since the conservation laws are widely accepted principles of natural science, and bare
consistency with the rest of what we know, at least, seems desirable on a theoretical level.
Compare this with last chapter, where I supposed that all parties involved—the
interactionist dualist, the physicalist, the emergentist (we might add), and Spinoza—want
to avoid skepticism. It is not enough, however, to state one’s opposition to the
undesirable position, or to insist that one’s views do not commit one to it. Papineau
challenges the non-physicalist to explain why, if her special mental force does not reduce
to the other fundamental, conservative forces, she expects it to be conservative
nonetheless. As Papineau explains:

Now we need independent evidence for supposing they [special mental forces] are
conservative, and it is not clear where it is to come from. In effect, then, positing
sui generis vital or mental forces threatens to undermine the inductive grounds for
upholding the conservation of energy in the first place. For it makes the
assumption of their conservativeness an independent assumption, an assumption
for which we lack any independent evidence.347

Providing some explanation for one’s continued commitment to the conservation
laws, given one’s rejection of physicalism, is seen in this formulation of the argument as a
central and perhaps insuperable challenge.

Of course, Papineau cannot logically preclude the non-physicalist from making a commitment to respecting the conservation laws an independent desideratum of her theory. But he can press her to explain what possible grounds, independent of the reduction of all special forces to fundamental, conservative ones, she might have for this desideratum. This problem becomes especially acute because special and fundamental forces can be acting on the same system at the same time, which puts significant pressure on the non-physicalist to keep her special mental forces conservative. (The sum of a conservative and a non-conservative force is non-conservative—conservation failures are infectious). Similarly, in the last chapter I tried to argue that the possible explanatory routes the dualist might try to take to explain why causation is ordinary, rather than pathological, are unconvincing. This failure does not logically prohibit the dualist from professing allegiance to anti-skepticism and interactionism and building an *ad hoc* explanatory model on which these two positions cohere. But it does make it seem like wanting to avoid philosophically undesirable consequences, rather than any resources within her theory, best explain these commitments.

The physiological argument then supplements this more abstract, theoretical argument in a negative way, simply by calling our attention to the absence of evidence for special mental forces as our understanding of the brain and body have increased.

**Physiology Argument:** There is no evidence for special mental forces irreducible to or inexplicable in terms of ordinary physical forces, either special or fundamental. The former ordinary physical forces reduce ultimately, in our most complete theory of physics, to the fundamental, conservative forces.
I do not have much to say about this argument. On one hand, any inductively justified claims will, at some point, have to allow inferences from the absence of evidence to the evidence of absence (i.e. from the non-observance of black swans to the claim that all swans are white). The fact that many such inferences fail does not mean that some do not succeed. On the other hand, the non-physicalist opponent will surely contest the success of the Physiology argument, especially given the relatively nascent state of neurophysiological research.\footnote{Indeed, there seems to be an interesting tension here for some physicalists. Consider the physicalist who endorses both the Physiology Argument and some hope of giving a naturalized account of intentional content. As I discuss in section 5.4.2, eliminative physicalists like Rosenberg have criticized their physicalist colleagues using what I call the Pessimistic Meta-Induction on naturalized theories of content. The basic charge here is that absence of evidence of success in this project of naturalizing the mind is evidence of absence of possible success. Suppose, as a physicalist, you’d like to reject Rosenberg’s argument. You then face the burden of suggesting why you endorse a similarly structured argument against the non-physicalist in the Physiology Argument case. Of course, this tension is not insoluble—you could reject all arguments of this form. But given the centrality of some such argument for inductive inference, it’s unclear how successful this resolution actually is.}

Before turning to Bennett’s second formulation, which I will call the Causal Closure Argument, I want to reiterate the importance of Papineau’s explanatory challenge and its connection to my overall project here on Spinoza’s behalf. Papineau’s explanatory challenge is far less frequently discussed in the literature on physicalism’s relationship to its rivals, but I think this is a serious mistake. \textit{Whenever} one is presented with a deductive argument whose conclusion one rejects, one has the option of simply denying a premise. This is a trivial observation, of course. But what it shows is that the non-physicalist’s strategy, when presented either with Papineau’s argument or Bennett’s Causal Closure argument below, will be simply to identify a premise for which the physicalist has provided insufficient evidence. The non-physicalist can then argue that we have insufficient reason to believe the premise, and that the argument requires it.
Furthermore, if the non-physicalist can then provide a model for how the offending premise might be false, a model compatible with her preferred view (a form of dualism or emergentism, say), then she can justifiably claim, dialectically at least, to have responded to the deductive argument. What Papineau’s (and Spinoza’s) explanatory challenges adds is some pressure to explain why and how one can maintain other natural commitments (viz. to anti-skepticism, to ordinary rather than pathological causation, to the conservation laws) in light of one’s alternative (viz. interactionist, irreducible special mental forces, emergentist) model. Spinozism allies itself well with Papineau’s physicalism, therefore, not because Spinoza was some kind of materialist, but because he highlights similar explanatory deficiencies in opposing views. Papineau admits that the non-physicalist can reject his empirical arguments—they are not deductive, after all. But the non-physicalist then accepts a special explanatory burden to explain either why the special mental forces are conservative, or, if they are not, how the interaction of special, non-conservative forces with fundamental conservative ones does not undermine the conservation laws quite generally. Similarly, Spinoza admits that interactionism is not inconceivable. But, on the assumption that interactionism is true, Spinoza argues that there are probably insuperable difficulties explaining why causation is ordinary, rather than pathological. These explanatory difficulties, for both Papineau and Spinoza, give good (though not deductive) reasons for thinking that the opposing views are false.
5.3.2 Bennett: the Causal Closure Argument for Physicalism

5.3.2a: Causal Closure, Wilson’s Objection, and Spinoza’s Agreement with Wilson

Papineau elegantly describes the deep explanatory worries for non-physicalists, echoing Spinoza’s challenge to interactionists in a way that demands further reflection, but his statement of the argument for physicalism is not the most precise one. Karen Bennett’s expanded, more careful version of the Causal Closure argument allows us to evaluate both the various uses to which an argument of this sort may be put, and the ways in which Spinoza’s position differs from other attempts to avoid the physicalist force of the argument. I’ll thus spend some time discussing the target of the argument, and then mention some other ways of evading its conclusion that may be fruitfully compared to Spinoza’s own strategy. But first, we need to see the argument, which I’ll call Causal Closure, presented by Bennett as a group of five inconsistent claims, one of which must be rejected:

1. Distinctness: Mental properties (and perhaps events) are distinct from physical properties (or events).
2. Completeness: Every physical occurrence that has a cause has an immediate, complete, wholly physical cause.\(^{349}\)
3. Efficacy: Mental events sometimes cause physical ones, and sometimes do so in virtue of their mental properties.

\(^{349}\) Bennett’s, “Exclusion Again,” 280-281 version reads instead: Every physical occurrence has a sufficient physical cause. I replace her version with one due to Sophie Gibb, “Defending Dualism,” for several reasons. First, Gibb qualifies the claim by arguing only that for every physical occurrence that has a cause, it is of a certain sort. This allows one to interpret indeterministic events as uncaused without trouble, but it reduces to Bennett’s version if one thinks this interpretation is incorrect. Second, Gibb makes it harder on the physicalist’s opponent by requiring that the physical cause not merely be sufficient, but also immediate and complete and wholly physical.
4. Nonoverdetermination: The effects of mental causes are not systematically overdetermined; they are not on a par with the deaths of firing squad victims.

5. Exclusion: No effect has more than one sufficient cause unless it is overdetermined.

Bennett notes that her presentation differs from several of those in the literature, emphasizing the various aims to which the argument may be put. Philosophers like Crane and Kim have presented this set of claims as an argument against Distinctness in particular, with the goal of arguing against non-reductive physicalism in favor of reductive versions. The problem with this strategy, Bennett thinks (and I’ll shortly agree), is that the involved parties (dualists, emergentists, non-reductive and reductive physicalists) all mean somewhat different things by Distinctness. Spinoza, I think, can make the same complaint about several of the crucial premises. But while I will address several of Bennett’s remarks relating to this aspect of the debate for my own purposes, I mostly want to avoid taking a stand here on whether reductive or non-reductive versions of physicalism are preferable. Since I am not arguing for a physicalist interpretation of Spinoza, my engagement with this aspect of the problem will be quite limited.

According to Bennett, all this is as it should be, since instead of focusing on internecine warfare among physicalists, Bennett aims to show “[that] all physicalists have

---

350 Crane, *Elements of Mind*; Kim, *Supervenience and Mind*

351 Specifically, I will discuss in a later section whether Spinoza can handle the objection from multiple realizability. Since Spinoza, were he any sort of physicalist, would be a reductive physicalist, it is important to see that this objection does not affect him. But overall, given that I do not read Spinoza as a physicalist at all, the majority of my discussion will focus on problems for physicalism generally, as opposed to some brand of physicalism in particular.
a well-motivated solution to the exclusion problem that no dualist has. Physicalists’ best option is to deny Exclusion, and thereby endorse a strategy that I [she] call[s] ‘compatibilism.’ This is strictly weaker than an argument against dualism as such, primarily because Bennett is unsure whether the dualist has sufficient justification for accepting Completeness. But since our primary purpose here is to evaluate whether Spinoza can accept the Causal Closure argument—his primary complaints with the dualist lie elsewhere, as presented in the previous chapter—we can ignore this complication.

What is this strategy that all physicalists, but no dualists, can employ to deny Exclusion? While the details are complicated, several of them are relevant for sorting out dualist, physicalist, and Spinozist readings of the argument, so I’ll briefly point to three aspects of the strategy. To get clear on why physicalists deny Exclusion, Bennett thinks, it’s important to note that physicalists and dualists typically accept another premise, Distinctness, for very different reasons. While dualists typically do so, on her reading, because of worries about explaining consciousness physically, “non-reductive physicalists, in contrast, typically endorse Distinctness for a combination of reasons having to do with the purported multiple realizability of mental-state types, and with the semantics of mental state terms.” This difference is worth keeping in mind because Bennett thinks all sorts of physicalist can affirm that while there is some sense in which

---

352 Bennett, “Exclusion Again,” 283

353 Bennett, “Exclusion Again,” 285; I want to advertise Spinoza’s response to the problem of multiple realizability very briefly, as it comes out in Bennett’s formulation. Spinoza is a token, not type, identity theorist. He is, moreover, a nominalist, and would thus likely deny the existence of robust mental state types. Of course, he thinks that many mental states are similar across people—otherwise the latter 3 books of the Ethics are superfluous—but I think he would deny the presupposition of mental state types (in a strict sense) that underlies the multiple realizability problem.
mental events and properties are distinct from physical ones, there is another perfectly
good sense in which they are not. Whether mental events or properties are identical to or
realized in physical events or properties is far less important, according to Bennett, than
the gulf that divides the physicalists from the dualist who believes conscious events or
properties cannot be explained physically at all!

The final two aspects of Bennett’s strategy to unify physicalists around a denial of
Exclusion concern differences in the modal strength of dualist and physicalist views.
Bennett’s second point is that both reductive and non-reductive physicalists 1) deny the
existence of special psychophysical laws while 2) endorsing global supervenience of the
mental on the physical with metaphysical, as opposed to merely nomological, necessity.
On Bennett’s reading, all physicalists, and no dualists, can endorse something like
Jackson’s ‘minimal’ physicalism:

Any world which is a minimal physical duplicate of our world is a duplicate
\textit{simpliciter}.\textsuperscript{354}

Bennett’s syncretic physicalist takes physicalism to be a contingent truth about
the actual world, but nonetheless endorses metaphysically necessary supervenience of the
mental on the physical. Bennett’s third claim is that all physicalists, but no dualists, can
deny one of the following two counterfactuals (listed below), which together are
necessary and sufficient, Bennett thinks, for the truth of Overdetermination. And since
the physicalist can deny one of them (O2) that the dualist cannot, Bennett thinks, she can
accept Distinctness and Non-Overdetermination. This completes the unified physicalist’s
strategy, since physicalists cannot relinquish Completeness or Efficacy. Given that the

\textsuperscript{354} Jackson, \textit{From Metaphysics to Ethics}, 12
five premises form an inconsistent set, according to Bennett, all physicalists must 
therefore deny Exclusion—so the argument serves not, as Kim thought, to put pressure 
on non-reductive physicalists, but as a way of unifying physicalists against dualists.

I want to focus briefly on Bennett’s discussion of how the physicalist can justify 
Non-overdetermination by rejecting one of the following two counterfactuals. I am less 
interested in the particulars of her case here than in a general set of presuppositions that 
come to light in her defense of it. Bennett thinks that, to avoid overdetermination, both 
the physicalist and the dualist must be able to reject one of the following counterfactuals.

O1: If m had happened without p, e would still have happened.
O2: If p had happened without m, e would still have happened.355

Let’s ignore for a minute Bennett’s case for the physicalist’s ability to reject O2, 
and briefly focus instead on why the dualist can supposedly deny neither. For Bennett, 
“she [the dualist] will not claim that it [O1] is vacuous, because she thinks that m can 
indeed happen without p…And the dualist will not want to say that (O1) is false, either. 
To say that (O1) is false is to say that if m were to happen without p, e might not occur. 
But that suggests that p is required, that m is not in fact good enough to do the work.”356

Now I’m not actually sure that all interactionist dualists would be persuaded by 
Bennett’s case here, since a natural component of interactionist dualism seems to be 
commitment to the thought that the physical substrate plays an integral role in causing 
certain effects. But nevermind that. There’s a more fundamental issue here, one that

355 Bennett, “Exclusion Again,” 288; m and p are the mental and physical events, respectively, and 
e is the effect caused.

356 Bennett, “Exclusion Again,” 289
comes out in Bennett’s discussion of the dualist’s ability to reject O2, which we’ll look at while again ignoring whether the physicalist can successfully deny either counterfactual.

Bennett is worth quoting at length on this point:

Can she [the dualist] adopt this strategy, and say with me that if the physical cause had occurred without the mental one, it would not have caused the same effects? Not without abandoning standard ways of evaluating counterfactuals. For the dualist, the closest world in which the C-fibers fire without pain is not a world in which various surrounding physical facts go differently. It is not a world in which the C-fiber stimulation takes place in a petri dish, or otherwise without crucial background conditions that actually obtain. It is instead a world in which the psychophysical law that links appropriately situated patterns of C-fiber stimulation to pains is violated. It is not a full-blown zombie world, mind you—that would clearly involve the kinds of “big, widespread, diverse violations of law” that Lewis says it is of the first importance to avoid (1979, 47). It is instead simply a world in which just that particular physical occurrence fails to give rise to the sort of mental one that usually accompanies it. That is merely a “small, localized, simple violation of law,” that allows us to “maximize the spatio-temporal region throughout which perfect match of particular fact prevails” (47-48). This one tiny little violation of psychophysical law is a lot easier to accomplish—if it can be accomplished at all—than a big sweeping change in circumstances.357

If one steps back and considers both modal aspects of Bennett’s strategy, I think her argument looks downright irrelevant to the issues at hand. To see this requires gaining a bit of distance from specifically late 20th and 21st century concerns in metaphysics, but the added perspective is enlightening. In order for Bennett’s strategy to work as an argument against the dualist, it must be the case that 1) there is a clear sense in which metaphysical and nomological modality differ, and the dualist must be able to affirm only nomological but not metaphysical supervenience of the mental on the physical. For Bennett’s overall unification project, it’s crucial that only the physicalist, and not the dualist, can affirm metaphysically necessary supervenience. The second key

357 Bennett, “Exclusion Again,” 291-292
plank of Bennett’s argument requires that 2) the dualist must accept the Lewis analysis of counterfactuals!

To her credit, Bennett anticipates resistance on both fronts. On the Lewis analysis of counterfactuals point, however, she provides no additional argument, merely doubling down. Preserving similarity in matters of particular fact is more important than preserving similarity in laws. She ignores issues about differing intuitions of modal similarity, and tacitly presupposes that such similarity judgments will be more or less uncontroversial. That is, she assumes first that telling whether another possible world preserves more similarity in facts or laws is straightforward, and second that we should prefer the former to the latter. By contrast, I think that the former project of determining relative world similarity in particular matters of fact is hopeless, so that the second—the dualist’s preference for preserving similarity in laws—never becomes relevant. More importantly, as we will see below, some conceptions of natural laws make the Lewis requirement—identifying differences in law as distinct from differences in matters of particular fact—impossible to fulfill.

Bennett does better in discussing the question of metaphysical versus nomological supervenience, though she does not notice the central connection between this point and the earlier disagreement over properly evaluating counterfactuals. Bennett discusses Jessica Wilson’s provocatively argued defense of the double thesis that 1) physicalists should be necessitarians about the laws of nature and 2) that (1) collapses the distinction between metaphysical and nomological possibility and necessity. Bennett responds as follows:

It is not clear why it would follow that supervenience with metaphysical necessity is not sufficient for physicalism. The mere claim that there is no real distinction
between nomological and metaphysical necessity can only show that there cannot be any nomologically-but-not-metaphysically-necessary supervenience relations—and thus that Chalmers’ version of property dualism (1996) is not coherent. It cannot itself show that a position that endorses a nomological-and-metaphysically-necessary supervenience claim can legitimately count as dualist.\textsuperscript{358}

But I think Bennett still misses the point of, and thus the force behind, Wilson’s central criticism. Wilson believes that there is a substantial conflict between physicalism’s naturalistic commitments and the assumption that metaphysical and nomological possibility do not coincide. But Wilson’s related point is that the actual laws of nature are plausibly \textit{holistic}—in the sense that any world with \textit{any} current law of nature has \textit{all the same} laws of nature.\textsuperscript{359} But if this is true, evaluating any of the counterfactuals in the Lewis style will be deeply problematic—there is no way to keep some of the actual laws while swapping out others. If the actual laws are holistic, and if there are any emergentist/dualist psychophysical laws, then, if nomological and metaphysical modality coincide, every world will have such psychophysical laws. Emergent laws will be metaphysically necessary! And if laws necessitate their instances, something Wilson may or may not be committed to, then every difference in law is a difference in matters of particular fact, so that the Lewisian distinction between the two cannot be maintained. Even apart from the question of whether the claims just made are true, it surely seems correct to say that Bennett’s physicalist cannot simply proceed \textit{on the assumption that emergent laws are not metaphysically necessary}. That is, she cannot blithely evaluate counterfactuals where some hold and others do not, and argue that the

\textsuperscript{358} Bennett, “Exclusion Again,” 297
\textsuperscript{359} Wilson, “Supervenience Based Formulations of Physicalism,” 453-454
dualist/emergentist must accept them, in a dialectical context where the modal status of emergent laws is precisely what’s at issue.

This is, I think, a huge problem for any physicalist argument against emergentist dualism along the lines of Causal Closure, absent supplementation from more general considerations of the sort I consider in chapter three. I thus think Wilson’s position is a real contender that Bennett’s argument does little to dislodge, especially when the relationship between both modal commitments of Bennett’s view (to different modalities and to Lewis’s analysis of counterfactuals) are fully appreciated. Bennett would need to point to some further explanatory disadvantage in Wilson’s position, in addition to the deductive argument given here. But she cannot simply assume several controversial modal principles and consider her conclusion established.

This disagreement between Wilson and Bennett over their respective modal commitments may seem orthogonal to my main concern in defending Spinoza here, but it is not. In fact, it provides a first reason, which supplements the discussion that immediately follows, why Spinoza might evade Bennett’s argument for physicalism. The key connection is that Spinoza shares Jessica Wilson’s view of the laws of nature as metaphysically necessary, and thus follows her in collapsing the distinction between metaphysical and nomological modalities. They disagree, of course, on several other points—whether the actual laws are or could be emergent, and whether Closure is true—but they are united in rejecting the implicit view of modality, and the accompanying evaluation of counterfactuals, that Bennett admits is central to her argument. But this is

---

360 As I’ll mention below, I think the biggest issue is that Wilson, “Supervenience Based Formulations of Physicalism,” 433 accepts that her version of emergentism violates the causal closure of the physical. If she’s right, then I think her position is inferior to both Spinoza’s and the physicalist’s. But I
still an important preliminary point, since it shows that the Causal Closure argument does not succeed even against emergentist dualist views without additional controversial presuppositions not stated in the argument itself. This increases the likelihood that Spinoza’s position, which is nevertheless closer to physicalism than emergentism in several ways, will be able to evade the Closure argument in a satisfying way.

5.3.2b The More Spinozistic Response

However, Spinoza’s best response to the Causal Closure argument is to 1) point out that many of the premises are systematically ambiguous and that 2) once we disambiguate them, Spinoza can accept them while continuing to deny physicalism. So recall again the five claims of the Causal Closure argument:

1. Distinctness: Mental properties (and perhaps events) are distinct from physical properties (or events).
2. Completeness: Every physical occurrence that has a cause has an immediate, complete, wholly physical cause.\(^{361}\)

think she’s wrong, and I think that Gibb, “Defending Dualism,” has convincingly shown why, in her defense of emergentist dualism. Thus I think that a complete Spinozistic argument against these emergentist forms of dualism would have to reckon instead with 1) the coherence of emergentism at all, 2) the possibility of brute psychophysical laws, and 3) violation of closure as violation of the PSR. It seems like Wilson and Gibb represent two different strands, in that Wilson argues that closure must be violated (and that it’s not a problem) but that the psychophysical laws aren’t brute, while Gibb argues that the emergentist need not violate closure, but that the laws must be brute. If this dilemma holds, Spinoza has a plausible argument against both forms.

\(^{361}\) Bennett’s, “Exclusion Again,” 280-281 version reads instead: Every physical occurrence has a sufficient physical cause. I replace her version with one due to Sophie Gibb, “Defending Dualism,” for several reasons. First, Gibb qualifies the claim by arguing only that for every physical occurrence that has a cause, it is of a certain sort. This allows one to interpret indeterministic events as uncaused without trouble, but it reduces to Bennett’s version if one thinks this interpretation is incorrect. Second, Gibb makes it harder on the physicalist’s opponent by requiring that the physical cause not merely be sufficient, but also immediate and complete and wholly physical.
3. Efficacy: Mental events sometimes cause physical ones, and sometimes do so in virtue of their mental properties.

4. Nonoverdetermination: The effects of mental causes are not systematically overdetermined; they are not on a par with the deaths of firing squad victims.

5. Exclusion: No effect has more than one sufficient cause unless it is overdetermined.

First, note the systematic ambiguities, nearly all of which stem from two distinctive Spinozistic commitments: Spinoza’s nominalism, and his treatment of all mental/physical terms and contexts as referentially opaque. I have mostly ignored Spinoza’s nominalism, and will continue to do so here, but Spinoza’s commitment to the referential opacity of attribute-specific (i.e. Thinking or Extended) contexts and terms deserves more careful treatment.\textsuperscript{362}

Spinoza accepts that mental properties or the causal relations within Thought are distinct from physical properties or the causal relations within Extension—each set of properties and causal chains belongs to a fundamentally different attribute or aspect of God or Nature’s way of being. Suppose for the sake of argument that an event is just a

\textsuperscript{362} There is a third issue, related to but not identical with the question of Spinoza’s nominalism, which concerns his extremely sparse ontology of only (one) Substance and (infinitely many) modes. I suggested in chapter 1 that we should think of Spinoza’s picture of attributes and modes along determinable/determinant lines. So, in the attribute of Thought, my mind is a particular determinate way of being Thinking. But we still require gradations along this determinable/determinant axis nonetheless, since ‘my mental state of thinking about lunch later today’ is even more determinate a way of being Thinking than my mind, in total, is. In what follows, I will try to speak in terms of states, events, causal relations, things, and properties, following contemporary usage, while mostly avoiding ‘mode’ talk, which I think unduly complicates the statement of Spinoza’s position.
thing’s bearing a property at a time.\textsuperscript{363} Then, while mental events considered as mental (i.e. a thing’s bearing a Thinking property at a time) are distinct from physical events (i.e. a thing’s bearing an Extended property at a time), the mental thing just is the physical thing. In other Spinozistic terms, the thing that is the mind is identical to the thing that is the body, even though the mind, considered as mental, is distinct from the body, considered as physical. As we saw in chapter one, this results from the fact that minds and bodies, mental and physical things, share all their attribute-neutral or purely extensional properties, and Spinoza thinks that only these properties are relevant for identity claims. As I also emphasized in that chapter, the ‘considered as’ locutions here are not mind-dependent designations, on pains of idealism, but objective designations independent of anyone’s mental activity.

For Spinoza, then, Distinctness captures something deep and important about the world’s organization. Spinoza affirms a real and important distinction between the attributes of Thought and Extension, and the corresponding differences in the properties and causal relations that fall under each. But identity theory captures the opposite side of matters. The mind and the body share all of the properties relevant for preserving identity, and thus the thing, which bears distinct Thinking and Extended properties, is ultimately one. 2p7s: “So also a mode of extension and the idea of that mode are one and the same thing, but expressed in two different ways.”

\textsuperscript{363} Obviously, the literature surrounding each of these categories is vast and complicated. What I aim to do here is to simply state Spinoza’s view reasonably precisely as an alternative to physicalism, in a way that captures the important role played by referential opacity. See Kim, “Events as Property Exemplifications,” for a standard view of this sort.
Moving on from Distinctness, Spinoza also accepts a version of Completeness that respects the opacity of causal contexts: every physical occurrence considered as physical has a complete, immediate, wholly physical cause. It’s just that every occurrence or physical event can be conceived in at least two (by the human intellect) and indeed infinitely many (by God’s intellect) ways. Non-Overdetermination is completely unproblematic as stated, and Exclusion is easy too, with the addition of referential opacity—no effects have multiple causes considered within a certain explanatory framework. Of course, the same effects can be considered both physically and mentally, but again the postulated cause will be one and the same, albeit considered in two different ways.

Efficacy is the most difficult to interpret, since on the one hand Spinoza clearly affirms that mental events sometimes, indeed always, cause physical events, as long as each thing/property/time combination is considered in just one attribute. However, the claim—mental events cause physical events ‘in virtue of their mental properties’—is difficult to square with Spinoza’s metaphysics. So I think that Spinoza must thus deny one conjunct of Efficacy.

The standard view is that denying Efficacy commits one to epiphenomenalism. Spinoza’s view is not, however, epiphenomenalist in any objectionable way. A typical way of characterizing epiphenomenalism is as the view: “Mental events are caused by physical events in the brain, but have no effects on any physical events.” What’s important here is an asymmetry in causal efficacy between the physical and the mental, whereby physical events cause mental events, but not vice versa. The first conjunct is

364 Robinson, “Epiphenomenalism,” Intro
false though, for Spinoza—mental events are not caused by physical events, considered as such, any more than physical events are caused by mental events, considered as such. Conceived neutrally, however, the thing or event conceived mentally is of course relevant to the thing or event conceived physically. But causal explanations of the physical effect must include the prior cause considered physically, rather than mentally, so as to respect Spinoza’s deep division between the causal and explanatory realms of Thought and Extension. Still, Spinoza is ultimately an identity theorist, and each thing is as genuine a cause within the order of every attribute as any other thing is. Unlike the epiphenomenalist who posits that mental events are causal danglers, Spinoza allows and indeed requires robust mental causation, albeit confined solely to the realm of Thought. Spinoza lacks the objectionable asymmetry between mental and physical causation characteristic of mainstream epiphenomenalist positions.

One might think that the position sketched on Spinoza’s behalf—identity theory that denies the most natural reading of Efficacy but does not thereby embrace epiphenomenalism—works better for some mental states than others. Robinson describes one traditional criticism of such views by noting that this identification strategy might be plausible for certain mental states, like beliefs and desires, but that certain other mental states or properties, such as qualia, pose a problem. He says: “Absent insight into the necessity of the connection between neural properties and qualitative properties, we are arguably in an explanatory position similar to traditional epiphenomenalism.”365 Now in some sense this remark clearly does not apply to Spinoza’s position. Unlike others who have pursued similar identifications, such as Donald Davidson, Spinoza both endorses the

365 Robinson, “Epiphenomenalism”, section 3.3
referential opacity of both explanatory and causal contexts (something Davidson denies). He also has an easy explanation for the necessity of all connections, and *a fortiori* for the connections among intentional, qualitative, and physical states, grounded in fundamental metaphysics. That is, the necessity of the connection is no more surprising than the necessity of any other (indeed all other) connections within Spinoza’s metaphysics; they stand or fall together.

In another sense, of course, Spinoza has clearly not answered the explanatory demand that proponents of qualia might push here. But I do not thereby think that he is in the position of having 1) to bifurcate his view of the mental to accommodate epiphenomenal qualia or 2) to leave qualia unexplained. Some have thought that Spinoza was consigned to (2): Lia Levy, for instance, suggests, “His works also seem to fail to provide arguments and theses that plainly articulate a theoretical alternative to the Cartesian notion of consciousness he so often opposes,” before attempting a reconstruction of such a theory on Spinoza’s behalf. In the final chapter, however, I’ll provide a way of reading Spinoza, somewhat consonant with a form of Russellian monism or panpsychism, which I think evades these worries, even if the theory is not explicit in Spinoza’s published works. For now, I think it’s enough to note that

---

366 Davidson, “Spinoza’s Causal Theory of the Affects,” fn. 13, 110-111, denies that referential opacity for causation as opposed to causal explanation makes sense to him. For an example of how this view would go in Spinoza’s case, see Della Rocca, *Spinoza*, who argues that Spinoza first makes all qualitative properties representational properties, and then explains the connection between representational and physical properties as I have above: namely, the thing that bears representational property R is identical to the thing that bears physical property P, even though R does not equal P.

367 Miller, “The Status of Consciousness in Spinoza’s Theory of Mind,” 203, Wilson, “Objects, Ideas, and Minds,” 126, and Bennett, “A Study of Spinoza’s Ethics,” 184-191 offer further pessimistic conclusions along these lines. Marshall, *The Spiritual Automaton*, by contrast, makes Spinoza’s theory of consciousness the centerpiece of his theory of mind, contrary to interpreters who find that Spinoza has nothing to say on the subject. Of course, the view of consciousness that results differs in many respects from the concerns believers in qualia deem most important.
epiphenomenalism in its traditional form cannot possibly be Spinoza’s considered position, despite his denial of one conjunct of Efficacy.

Given these reflections on the ambiguities here, we can sum up Spinoza’s response to the Causal Closure Argument as follows. On the most plausible reading, Spinoza simply denies Efficacy, while further denying, as we saw immediately above, that this has any epiphenomenalist consequences. Certainly, he denies epiphenomenalism in one common understanding of the term—physical events do not cause or give rise to mental events, considered as such, and more fundamentally, there is no causal asymmetry between the mental and the physical, no sense for Spinoza in which the mental is a causal dangler. Whether Spinoza can integrate some of the key insights of the recent literature on consciousness, without making at least these qualitative aspects epiphenomenal, is a question I leave for next chapter.

Regardless, for both Bennett’s physicalist and for Spinoza, the crucial Completeness premise remains secure. This echoes Hampshire’s interpretation of Spinoza’s view, according to which “the physical world must be conceived as complete in itself, self-generating and self-maintaining.” Other interpreters have also found this aspect of Spinoza’s view compelling. William Seager puts the point this way:

The significance of Spinoza’s position is that it can both (1) endorse the causal closure of the physical, which many take to be a regulative principle of, and perhaps even entailed by the structure of our most fundamental physical theories, and (2) deny that mentality must in some way be reducible to the physical.369

368 Hampshire, *Spinoza and Spinozism*, 63

369 Seager responds to Atmanspacher, “Some 20th Century Variants of Dual Aspect Thinking,” 274; Atmanspacher’s denial of closure on Spinoza’s behalf occurs at “Some 20th Century Variants of Dual Aspect Thinking,” 247, where he writes: “Causal closure in Spinoza is violated in a subtler way. Since the modes, which do not interact directly, derive from the one substance, this substance may inject effects, intrusions as it were, into the modes so that they cannot be causally closed in principle.” Atmanspacher gives very little reason for thinking this is the correct interpretation of Spinoza. His basic thought (284-286)
Seager precisely expresses the strength of Spinoza’s position as I interpret him here. This point is crucial: the physicalist’s task, we might say, is to push her opponent either towards dualism or epiphenomenalism, and to argue that both of these options are unacceptable. But the argument from the Completeness of physics, which I have called Causal Closure, does not accomplish this aim, for two reasons. First, Spinoza faces no pressure to accept any sort of ontological dualism; he is its most ardent critic. Second, Spinoza has good reasons to deny epiphenomenalism, and moreover the outstanding questions about his ability to integrate qualia are ones that similarly afflict physicalism, and thus cannot be used to give the physicalist an advantage here. It is no accident that Kim calls one of his masterful treatments of these problems *Physicalism or Something Near Enough*—this latter concession points to his preference for acknowledging that qualia, if they are to remain real, must be interpreted epiphenomenally.

As I’ll argue next chapter, Spinoza need not accept this unhappy resolution of things. For now it’s simply enough to note that the Causal Closure argument is one whose most important premises, specifically the Completeness premise, Spinoza can happily accept. Even more, Papineau’s explanatory challenge to non-physicalists—to explain why conservation laws should be respected by their models, given that these laws now represent independent assumptions rather than laws receiving inductive support from the

_____________________________

seems to be that if Spinoza is interpreted as panentheist, rather than pantheist, then we need to see the causal relations between God as Substance and individual modes as formal, rather than efficient, causes, and further that causal closure should be conceived as closure with respect to efficient causes. First, I do not think Seager or I need to read Spinoza as pantheist rather than panentheist to claim that he affirms Closure. The relevant point is the disagreement about formal rather than efficient causes, and here I will simply note that the assumption that efficient causation exhausts causation, for Spinoza, is standard in the literature. See Renz, “Explicable Explainers,” for a representative case of this assumption.
completeness of the system as a whole—poses no threat to Spinoza. For the
Completeness of physics is one of the guiding principles of Spinoza’s own system, a
straightforward metaphysical consequence of his determinism, his monism, and the
explanatory barrier between the attributes.

5.3.3 Brief Interlude on Supervenience

Before moving on to discuss Spinozism’s advantages over physicalism, I want to
briefly discuss supervenience and Spinoza. One might wonder whether Spinoza can
accept one or all of the various supervenience theses that many contemporary
philosophers have found especially plausible for characterizing the relations between the
mental and the physical. Of course, the expected answer might seem obvious: Spinoza
can certainly accept any supervenience claim one might formulate, given my
interpretation of him as an identity theorist. For identity seems strictly logically stronger
than supervenience, such that if X=Y then any supervenience relations of whatever modal
strength will hold between the properties of X and Y. Nonetheless, Spinoza’s claim—all
uses of mental and physical properties create intensional contexts—makes this inference
from identity to supervenience too quick, though I think it’s right, ultimately, that
supervenience poses no challenge to Spinozism.

It’s worth seeing how the story might go in greater detail, however. The
fundamental intuition behind supervenience claims is easy enough to describe.
McLaughlin and Bennett, for instance, characterize the basic supervenience relation as
follows: “A set of properties A supervenes upon another set B just in case no two things
can differ with respect to A properties without also differing with respect to their B
properties.” The A-properties are nothing over and above the B-properties, we might say. Phrased this way, the intuition is easy for Spinoza to accept, for 2p7 and 2p7cor state: “The order and connection of ideas is the same as the order and connection of things…That is, whatever follows formally from God’s infinite nature follows objectively in God from his idea in the same order and with the same connection.” Spinoza’s position in contemporary terminology is thus: any difference in A properties (i.e. any difference in physical properties) necessitates a difference in B properties (i.e. a mental difference).

Moreover, Spinoza’s idiosyncratic modal commitments allow him to sidestep many of the central debates surrounding various supervenience theses discussed in the literature. So Bennett and McLaughlin spend time discussing 1) the differences between logical, metaphysical, and nomological necessity, and 2) which type of necessity appropriately characterizes the relation between the physical and the mental. Moreover, they examine whether a weak, intermediate, or strong global supervenience thesis should be accepted, where the strong global supervenience is defined as follows:

A-properties strongly globally supervene on B-properties iff for any worlds w1 and w2, every B-preserving isomorphism between w1 and w2 is an A-preserving isomorphism between them.

Various worries about characterizing supervenience then arise. One might do so too weakly, so that dualism can gain a foothold, or too strongly, so that physicalism

---

370 Bennett and McLaughlin, “Supervenience” Intro

371 We are once again ignoring Spinoza’s nominalism for the moment.

372 Bennett and McLaughlin, “Supervenience,” 4.3.2
becomes an implausibly strong thesis about, say, the logical impossibility of Cartesian souls or other ghostly things.

But, at least on some readings, Spinoza can sidestep basically all of these issues, since Spinoza is arguably a necessitarian: every state of affairs is logically or metaphysically necessary; there is only one possible world. Don Garrett argues that three crucial propositions—1p16, 1p29, and 1p33—as well as much of Ethics Part II, commits Spinoza to necessitarianism.373 Perhaps the most striking of these is 1p33: “Things could have been produced by God in no other way, and in no other order than they have been produced.” Unshockingly, passages like these have nonetheless failed to produce consensus. Even before Garrett’s article was published in 1991, interpreters had spilled a great deal of ink on this question, and the tide has scarcely abated since then.374 Fortunately for my purposes here, few interpreters make Spinoza out to be much less than a strong determinist, in the sense both that the past plus the laws of nature determine a unique future, and in the sense that the actual laws of nature are necessary. This, combined with Spinoza’s other metaphysical positions, makes any putative supervenience violation difficult to interpret, for almost any reasonable formulation of it.

There is a general class of formulations of supervenience that Spinoza cannot accept, however. Any formulation of supervenience that extends beyond mental-physical

373 Garrett, “Spinoza’s Necessitarianism,” 205

374 By Garrett’s own catalogue at the time of publication (Garrett, “Spinoza’s Necessitarianism,” 192) Hampshire argues that Spinoza is a necessitarian, Curley that he is not, Matson, Jarrett, and Bennett maintain that he inconsistently affirms and denies the thesis, and Delahunty argues that ambiguity makes him incapable of committing himself either to the thesis or its denial. Since this paper, many more philosophers have weighed in. For another brief survey: Curley and Walski, “Spinoza’s Necessitarianism Reconsidered,” read Spinoza as a strong determinist (determinism is true and the laws of nature are necessary, in addition), Della Rocca, Spinoza seems to accept Garrett’s necessitarian reading, and Newlands, “Spinoza’s Modal Metaphysics,” has argued that Spinoza is committed both to necessitarianism and its denial, consistently, relative to different ways of conceiving.
correlations, with whatever modal strength, to require *explanations or dependence relations* of the supervening properties in terms of the “basic” properties will be unacceptable for Spinoza. So, for example, Spinoza cannot accept Terence Horgan’s suggestion that physicalism requires superdupervenience, and that superdupervenience places additional restrictions on the sorts of relations allowed between the basic and supervening properties:

Superdupervenience: Any materialistic metaphysics should countenance ontological inter-level supervenience relations only if they are robustly explainable in a materialistically acceptable way.\(^{375}\)

Ontological relations are objective relations between higher order and lower order facts or properties, as opposed to conceptual or semantic relations, for Horgan. Thus, the new requirement becomes not merely that there be such objective relations between mental and physical, but that the relations themselves be materialistically explainable. Spinoza, as I have characterized him, cannot accept any materialistically acceptable explanations, since 1p10 prevents any causal or conceptual connections between thinking and extended modes. Thankfully, however, Jessica Wilson has pointed out that, regardless of what one thinks of Horgan’s constraint as a move in the debate between physicalists and their opponents, it changes the debate about supervenience in particular fundamentally:

First, accepting such a strengthening of the supervenience connection means giving up the reasons for characterizing nothing over and aboveness in terms of supervenience in the first place…We have come full circle through the spectrum of supervenience-based accounts of nothing over and aboveness, to return to a reduction-based account…Hence even supposing that a superdupervenience-based formulation of physicalism were satisfactory, it would be disingenuous to

\(^{375}\) Horgan, “From Supervenience to Superdupervenience,” 563
claim that this vindicated supervenience-based accounts of nothing over and aboveness.376

Wilson’s point, which she applies not only to Horgan’s constraint, but also to Chalmers’ broadly logical supervenience requirement, is first and foremost that such restrictions change the subject. Rather than concerning themselves with supervenience in particular, they focus on the prospects for causal or functional reductions of the higher order properties to the lower. This is probably because such reductions are regarded as less problematic than the type-identity claims (ala Smart or Nagel) that were previously popular, but are now widely regarded as unworkable. At the very least, such restrictions change the subject from modal covariance to the question of some asymmetry between the lower level properties and the higher level ones.377

Jaegwon Kim also defines supervenience as a dependence thesis: “I take supervenience as an ontological thesis involving the idea of dependence...Supervenience, therefore, is not a mere claim of covariation between mental and physical properties; it includes a claim of existential dependence of the mental on the physical. I am assuming that a serious physicalist will accept this interpretation of supervenience.”378 Kim goes on to suggest that weakening supervenience makes it acceptable to too many other views—double-aspect theorists, neutral monists, emergentists, epiphenomenalists, and even substance dualists. Now, I think that some of these views have a much easier time incorporating supervenience than others, with the substance dualist certainly having the

376 Wilson, “Supervenience Based Formulations of Physicalism,” 451-453

377 Even speaking of ‘lower’ and ‘higher’ level properties, of course, makes no sense for Spinoza, for whom Thought and Extension are equally fundamental.

378 Kim, Physicalism, or Something Near Enough, 34
most difficult. Still, I think this is basically as it should be. If the physicalist’s only way of employing supervenience as an argument against competing positions is to define it so that only physicalism is consistent with it, and then to argue that consistency with the putative definition is required for any plausible philosophical position, she will not have made much headway. An argument of that sort boils down to stamping one’s feet about definitions, which is never a dominant strategy and is particularly implausible when there are so many other reasonable definitions of supervenience, some of which I considered above.

I conclude, therefore, that Spinoza faces no unique threat from supervenience. The real source of tension between Spinoza and the physicalist concerns the relative fundamentality of the physical and mental properties and entities: the physicalist sees an asymmetry where Spinoza sees equality. So the physicalist takes supervenience to give not merely correlations, but dependence relations, but this only because equality between the two is never on the table. It is not incidental, for the physicalist, which properties are framed as the supervening, as opposed to the basic, properties.

But in the argumentative context where Spinoza affirms the identity of mental and physical modes, it’s quite difficult to see what sort of general additional argument the physicalist might give to buttress her intuitions of asymmetry. Indeed, I think the primary source of resistance to any view like Spinoza’s has simply been that it has been unclear what the motivations for it might be. Another source of resistance might be the sense that the degree of causal and explanatory opacity required by Spinoza’s system is somehow
Opaque motivations plus the absence of any argument for 1p10 might be enough to explain why more care has not been taken by philosophers outside of the eliminativist camp to give compelling reasons for rejecting it. In any case, any such argument along those lines would change the subject in precisely the way that Wilson claims, so that the particular argument on those questions, rather than the supervenience claim, would be doing all the work against Spinoza’s view.

5.4 Spinoza’s Leg Up on the Physicalist

In this final section, I want to briefly consider a number of ways in which Spinozism has an advantage over most forms of traditional physicalism. In most cases, this advantage takes the following form: traditional physicalism faces a problem it solves badly or not at all, while Spinoza faces no similar problem or faces it and solves it quite easily. Because my initial discussion will be quite brief here, and because I devote chapter 6 to integrating Spinoza’s isolationism into the consciousness-heavy aspect of the general debate in the metaphysics of mind, my argumentative aims are quite limited. Still, I think the final Spinozistic advantage I discuss is quite powerful, and I’ll thus focus the majority of my attention here on it. Many, though not all, physicalists view their collective failure to ‘naturalize the mind’ —to provide a physicalistically-acceptable account of intentional content—as the greatest challenge to their worldview. Spinoza, by contrast, maintains that Thought and Extension are two of infinitely many equally fundamental ways of being. Thought’s primitive intentionality in this system appears no

379 As mentioned, this motivates Davidson’s rejection of causal opacity, though he accepts explanatory opacity.
more puzzling than what Spinoza calls the proportion of motion and rest that essentially characterizes Extended things. If the difficulties with this naturalizing project are as insuperable as they appear from recent discussions, therefore, traditional physicalists have plenty of reason to take Spinozistic isolationism seriously as an alternative to their view.

Before getting into the details of my case for Spinozism’s explanatory advantage over traditional physicalism, I want to finally discuss a potentially serious objection to Spinoza’s view. The objection from multiple realizability might have seemed pressing as soon as I interpreted Spinoza as an identity theorist, but it certainly seems so when actively comparing Spinoza’s view to other forms of physicalism. I will argue that, for several reasons, multiple realizability poses no threat at all to Spinoza’s view of the mind. For that reason, traditional type identity theorists, for whom multiple realizability is taken to be a serious objection, are worse off than the Spinozist.

We can pump the intuitions favoring the multiple realizability of mental states (MR) in various ways. Uncle Doug and his dog Dan can both be in pain, but there’s no state of Doug’s brain that’s the same in Dan’s. The human Doug has many neural structures that Dan simply lacks, even though they can both be in the same mental state, that of feeling pain. More wildly, Martians might feel pain, or believe that 2+2=4, without having brains at all. Perhaps Martian thoughts are realized in some entirely different substrate than the broadly biological one that Doug and Dan both share. Sameness of mental state is thus compatible, on this view, with difference in physical state, even when the differences are quite radical. Defenders of multiple realizability are often functionalists about the mind. So they want to defend both that mental properties
are realized, and perhaps must be realized, by some physical properties, but that they
need not be realized by any physical properties in particular. Multiple physical substrata
will do for making a mind.

Alarm bells should be ringing, for upon reflection none of the upshots of these
intuition pumps are particularly clear. As Louise Antony notes, even though MR
intuitions appear to be widely shared among both philosophers and ordinary people, there
are good reasons to question them. Antony first cites Ned Block’s thought experiment, in
which the citizens of China replicate the functional organization of the human mind. She
then mentions Lewis’s contentions that MR intuitions seem most stable in far-fetched
cases (when contemplating minds radically unlike our own), but that we tend to “treat
sameness of physical state as criterial” when making judgments about the mental states of
terrestrial creatures. Moreover, it’s not clear even that Lewis is right about this last
point. Doug and his dog Dan do not share any physical states, it seems, in precisely the
same sense in which Doug and his wife Dottie do not. Of course, the level of similarity is
much higher in members of the same species, and Lewis’s intuitions about similarity
relations may be guiding him here. But it’s pretty natural to think that no two organisms,
and even no single organism at different times, have precisely the same total physical
states or even the same total brain states. In that case, it’s not clear why the fact that
Doug’s dog Dan, or a Martian, might be in pain is any more puzzling for an account of
mentality than how Doug and Dottie both can be.

There are various responses one might make to this sort of puzzle about MR. For
my purposes here, the important point for interpreting Spinoza’s position has nothing, in

380 Antony, “Multiple Realizability: Keeping it Real,” 165
particular, to do with his identity theory of mind. Instead, Spinoza’s commitment to the PSR, plus his nominalism, make him unlikely to accept several of the basic premises required to get the MR intuitions going. Because establishing the scope and tenor of Spinoza’s nominalism would take me much too far afield, I’ll simply adopt Newlands’ view that Spinoza should be placed “in the family of conceptualists, for whom abstracta [universals, numbers, species, wills and intellects, good and evil] are modes of thought occasioned by but not grounded in mind-independent things…abstracta are confused representations of a finite mind”381 Newlands also provides a positive account on Spinoza’s behalf. For Spinoza, abstracta are mnemonic devices for minds incapable of seeing the necessary connections between each and every particular—“abstracta like biological kinds help us recall particulars through mental associations.”382 The crucial point here is that mental kind terms—pains, propositional attitudes more generally, etc.—are, like biological kind terms, abstracta as well, and perform similar functions.

Spinoza will thus not deny that there is some truth in the statement that Dottie and Doug, both being in pain, are in the same mental state. But this is just convenient shorthand for we finite minds, a concession to our limited inferential powers where particulars are concerned. Spinoza’s considered view must be that the PSR, as a general metaphysical principle, entails the PII, which precludes the existence of two qualitatively identical yet distinct things. A fortiori, no two organisms share the same total physical states, or the same overall mental states. As Spinoza says in Ethics 3p56: “There are as

381 Newlands, “Spinoza’s Early Anti-Abstractionism,” 259, while 252 gives examples of abstracta, for Spinoza.

382 Newlands, “Spinoza’s Early Anti-Abstractionism,” 260
many species of joy, sadness, and desire, and consequently of each affect composed of
these (like vacillation of mind) or derived from them (like love, hate, hope, fear, etc.), as
there are species of objects by which we are affected.” The problem of mental
realizability thus never gets going, since strictly speaking Dottie and Doug, much less
either of them and a Martian, do not share any mental states. The intuition that the same
mental state is multiply realizable thus never gets off the ground.

In a more traditional spirit, we might explain Spinoza’s evasion of the problem in
a different way. Spinoza is a token, rather than type, identity theorist. He believes that
particular physical modes are identical to particular mental modes, but makes no claims
about, and would not accept, arguments that assume mental state type identity and then
argue that such states are realizable in a variety of physical substrata. Like Jaegwon Kim,
Spinoza favors local rather than global reductions (in the sense of identity, though not in
the sense of explanatory reductions, given the EB). Given his other views, then, I think
that even if respecting MR intuitions were a desideratum of any viable theory in the
metaphysics of mind, Spinoza could evade objections along these lines. Moreover, to the
extent that one views that challenge for MR as a serious one for the traditional type
identity theorist, Spinoza’s view fares substantially better.

5.4.1 Preliminary Challenges for Physicalism, but not Spinozism

Before coming to the central challenge for the physicalist posed by the project of
naturalizing the mind, I want to mention several other areas where Spinoza seems to have
a prima facie advantage over traditional physicalism. Since I will not be delving into the

383 Kim, Physicalism or Something Near Enough
details here, my argument’s scope is very limited; I am happy to call to mind a number of
prima facie deficiencies in physicalism, even if appearances on one or another front are
deceiving.

We might start with the definition of physicalism. This problem takes several
forms. We have met one iteration of this problem, albeit obliquely, already. Nearly every
supervenience-based definition of physicalism seems insufficient for distinguishing
physicalism from its rivals. Wilson argues that only questionable modal assumptions
about the laws of nature (as contingent), as well as a tacit denial of holism (about the
same laws of nature) allows the physicalist to distinguish her position, via some
supervenience formulation, from dualist emergentism. Other problems with
supervenience discussed in the literature include the lone ammonium molecule problem,
the modal status problem, the epiphenomenal problem, and the blockers objection.

Crane and Mellor, following Hempel, forcefully push a different problem for
physicalism’s definition: it seems to face a dilemma of being either vacuous or false. The
thought is that the physicalist’s claim is that “the empirical world contains just what a
true complete physical science would say it contains.” This way of stating things
avoids falsity—for if the claim is supposed to be that the world’s entities are exhausted
by those postulated by current physical science, is it almost certainly false. Almost

384 Wilson, “Supervenience Based Formulations of Physicalism,”
385 Stoljar, Daniel, "Physicalism", especially 4.1-4.4; the blockers objection is given in Hawthorne,
“Blocking Definitions of Materialism.”
386 Crane and Mellor, “There is No Question of Physicalism,” 186; of course, I ignore here the vast
literature surrounding other ways of avoiding this sort of objection, including definitions of the physical as
not paradigmatically mental, etc. I am skeptical that these other ways of avoiding the dilemma succeed, but
I am content to merely present the problem without evaluating possible solutions for my purposes here.
certainly, current physical science is, at best, radically incomplete, and in some specific cases it seems inconsistent. But it embraces vacuity—who knows what sorts of entities will show up in our final physics, or whether it is even coherent to speak of a ‘final physics’ in the first place? And, more relevantly, how can the physicalist preclude paradigmatically mental entities from showing up in the final version of physics, should such a thing exist?

By contrast, Spinoza defends his monistic system on metaphysical grounds, and develops a highly intensional account of most of the major notions within it: mental and physical properties, different ethical conceptions, modality more generally, etc. So though recent developments in the natural sciences provide important psychological motivations, perhaps, Spinoza is not beholden to them in order to define his system’s central notions. Additionally, his view is an idiosyncratic form of the identity theory, so he need not worry about how the details of supervenience-based formulations turn out. However they do, his view respects the underlying intuition (even if he cannot accept every such supervenience-based formulation), while simply rejecting any further asymmetry claim between the properties of Thought and Extension.

Moving on from the definition of physicalism, several philosophers have argued that physicalism or closely related views lead to skepticism. For instance, Michael Rea has argued that naturalism might be inconsistent with discovering individuals’ modal

properties, realism about material objects and perhaps realism about other minds. In a similar vein, Alvin Plantinga has argued, roughly, that the probability of our cognitive faculties—perceptual, memory, inferential—being reliable, and thus of at least some of our beliefs being true, given a commitment to materialist naturalism and the theory of evolution, is quite low. This means that we should be skeptical about the deliverances of our faculties, and thus about our confidence in the theory of evolution discovered by them. Thus, materialist naturalism both appeals to evolutionary theory for justification, and undermines our confidence in it; the conjunction of the two is thus self-defeating.

The basic problem here, I think, concerns both the randomness of evolutionary development, and the extent to which evolution prescribes adaptive beliefs and behaviors, rather than beliefs that are true and behaviors that are conducive to knowing the truth. Of course, true beliefs might be adaptive, but there is no special reason to believe that they must be. Plantinga supports his case with a long list of worries along these lines from non-theist naturalists. Thus Darwin: “With me the horrid doubt always arises whether the convictions of man’s mind, which has been developed from the mind of the lower animals, are of any value or at all trustworthy.” And Churchland:

“Improvements in sensorimotor control confer an evolutionary advantage: a fancier style

---

388 Rea, *World Without Design*, 50 and elsewhere. To be clear, Rea distinguishes between naturalism and physicalism, characterizing naturalism as a ‘research program wherein one treats the methods of science and those methods alone as basic sources of evidence’ (50). He does this only to save naturalists from Crane/Hempel’s dilemma, which I discussed above, however, and obviously notes the tight connection between naturalism and physicalism, as evidenced in Tye, for instance, who seems fine with the equivalence, perhaps because he denies the seriousness of the above dilemma (“Naturalism and the Problem of Intentionality,” 129): “On the naturalist view, the world contains nothing supernatural...at the bottom level there are microphysical phenomena governed by the laws of microphysics, and at higher levels, phenomena that not only participate in causal interactions describable in scientific laws but also bear the same ontic relationship to microphysical entities.”

of representing is advantageous so long as it is geared to the organism’s way of life and enhances the organism’s chances of survival. Truth, whatever that is, definitely takes the hindmost.\footnote{Churchland, “Epistemology in the Age of Neuroscience,” 548}

At least one of Spinoza’s commentators, Stuart Hampshire, has suggested that Spinoza anticipates in many respects, and would thus welcome, modern evolutionary theory.\footnote{Hampshire, Spinoza and Spinozism, esp. 62-70 but also elsewhere} But whereas modern evolutionary theory posits adaptations occurring through random mutations leading to increases in fitness, Spinoza’s metaphysics requires that such changes occur of necessity. This does not mean, and Spinoza would certainly not believe, that evolutionary processes are guided or directed in any way. As he says in the Appendix to Ethics Part I: “All the prejudices I here undertake to expose depend on this one: that men commonly suppose that all natural things act, as men do, on account of an end” (G II/78).

Still, Spinoza’s necessitarianism, his theory of truth, and his overall epistemology make him far less vulnerable to criticism along the lines that Rea and Plantinga push. We can know with certainty, for instance, that “each idea of each body, or of each singular thing which actually exists, necessarily involves an eternal and infinite essence of God” (2p45); indeed, “the human mind has an adequate knowledge of God’s eternal and infinite essence” (2p47).\footnote{Marshall, The Spiritual Automaton, offers perhaps the best defense I’ve seen of the consistency of these claims with the rest of Spinoza’s metaphysical system.} Several classes of ideas—ideas of God’s essence, of certain common truths of mathematics and logic, and even the essences of finite modes,
considered in a certain way—are guaranteed to be adequate in the human mind. Realism about material objects, realism about other minds, modal knowledge, and general trust in at least some of the deliverances of our cognitive faculties thus cohere far better, it seems, with Spinozism, taken as a total metaphysical system, than with traditional versions of physicalist naturalism.

5.4.2 Physicalism Seems to Have a Problem with Intentional Content

The physicalist’s ontological commitments seem to require that she provide an account of how intentional states—states that are of or about things, directed at their objects, which may or may not exist, and perhaps display referential opacity—can be explained in terms of physical ones. Fodor states the challenge for the physicalist in his characteristically pithy way:

I suppose that sooner or later the physicists will complete the catalogue they’ve been compiling of the ultimate and irreducible properties of things. When they do, the likes of spin, charm, and charge will perhaps appear upon their list. But aboutness surely won’t; intentionality doesn’t go that deep. It’s hard to see, in face of this consideration, how one can be a Realist about intentionality without also being, to some extent or other, a Reductionist. If the semantic and the intentional are real properties of things, it must be in virtue of their identity with (or maybe of their supervenience on?) properties that are themselves neither intentional nor semantic. If aboutness is real, it must be really something else.

In what follows, I do not have the space to discuss one response to this challenge, which is uncommon but perhaps growing in popularity, and is exemplified in the range of works on powers in metaphysics. See, for example, Martin, The Mind in Nature; Molnar, Powers: A Study in Metaphysics, and several works by Heil. This is regrettable, but commensurate with my overall argumentative aims, since I nowhere purport to have provided a conclusive Spinozistic case against the physicalist. Rather, the goals of this chapter and the next are simply to show that the physicalist can offer the Spinozist no compelling arguments against his position, and that Spinozistic Isolationism seems to retain significant advantages over most ordinary forms of physicalism. Whether the views sketched by Martin, Molnar, and Heil count as ‘ordinary physicalism’ is an open question, and one worth further reflection, but not one I can discuss here. Morch, Panpsychism and Causation, does an elegant job of doing so, however.

Fodor, Psychosemantics, 97
One option we have seen already is to deny realism about the intentional; thus Churchland and Rosenberg embrace forms of eliminative materialism, which deny that beliefs and desires have representational content, and thus deny that they are about other things. This is not a popular response, but it’s at least arguably the most principled one that the physicalist can take. Rosenberg, who states the point forcefully while considering the natural physicalist response, is worth quoting at length:

If I believe that Paris is the capital of France, there has to be a clump of matter—some wet stuff in my brain—that is about Paris...But physics fixes all the facts, and it assures us that there cannot be clumps of matter—fermions and bosons—that just are, in virtue of their constitution, about other clumps of matter. So no original intentionality. You will doubtless be tempted to reply that it’s not just the clump of matter—the bit of porridge in the brain—that is, by itself, about Paris; it’s the neural circuits plus other clumps of matter causally connected to it in the right way. Original intentionality is a causal role with respect to other clumps of matter (remember that science only recognizes clumps of matter and fields of force). But piling up clumps of matter without original intentionality and having them participate in complex causal processes with one another won’t produce original intentionality—it’s still just fermions and bosons. There is no better proof of this than the limitations of teleosemantics.

Teleosemantics isn’t just the best naturalism can do to provide an account of original intentionality. It is the only possible account of it if the physical facts fix all the facts. Brain states and the behavior they bring about are among the most purposefully appearing things and events in the universe. The only way they can discharge their appearance of purpose in a world where physics has banned real purpose is via a Darwinian process of blind variation and natural selection. The essence of intentionality is purpose, as Dennett (1969), Bennett (1976), Dretske (1988), Millikan (1984), Papineau (1993), Neander (2006), Matthen (1988), and Lloyd (1989) have shown. But teleosemantics can’t individuate intentional content. No amount of environmental appropriateness of a neural state or its effects is fine-grained enough to give unique propositional content to the neural state, to confer on it the sort of specific aboutness that original intentionality requires. Teleosemantics can’t solve what Fodor calls the disjunction problem. So much the worse for original intentionality! If Darwinism about the brain can’t give us unique propositional content, then there is none. Because if Darwinism can’t give us content, nothing can. The conclusion to draw is that the brain does not acquire, store, or deploy its information propositionally, in ways that require original intentionality.395

395 Rosenberg, “Disenchanted Naturalism,” 25-26
I think we can glean at least three interesting arguments from this passage, all of which threaten realism about the intentionality of mental states, or prompt us to seek another route of securing it. They are, in rough outline:

**Impossibility**

1. Physics fixes all the facts.
2. If physics fixes all the facts, then no fundamental physical entities (i.e. fermions and bosons) are, by their nature, about any other entities.
3. If no fundamental physical entities are about any other entities, there is no original intentionality [definition of original intentionality].
4. There is no original intentionality [1,2,3].

**Deficiencies**

1. Physics fixes all the facts.
2. If physics fixes all the facts, then the only possible account of intentionality is some form of causal theory, or some form of teleosemantic proper function theory.
3. Neither extant causal theories nor extant teleosemantic theories can appropriately individuate intentional content—Fodor’s disjunction problem is insoluble given current resources.
4. There is (as yet) no non-deficient account of intentionality,

The final argument, which I find more plausible than **Deficiencies**, is not a deductive refutation of naturalistic theories of intentionality, but a more plausible pessimistic induction, which somewhat parallels arguments in the philosophy of science.
Pessimistic Meta-Induction

1. Past naturalistic theories of intentional content (i.e. Dennett (1969), Bennett (1976), Dretske (1988), Millikan (1984), Papineau (1993), Neander (2006), Matthen (1988), and Lloyd (1989) seemed promising and possibly true, but were ultimately unworkable or obviously false.

2. So, probably, future philosophers will work as unworkable or obviously false later naturalistic theories that supersede them.

3. So, probably, we have reason to believe that any current naturalistic theories of intentional content are unworkable or obviously false.

Since my point is not to defend Rosenberg’s arguments in detail, as I’ve excavated them from this passage, I will briefly just make several points. Of course, one obvious objection to Rosenberg is simply that physics doesn’t fix all the facts, which thereby undermines the first two arguments. The Pessimistic Meta-Induction for intentional content then just expresses a natural caution and need for further investigation into a physicalistically-acceptable account of intentionality, but not a serious challenge to its possibility.

Moreover, one might be generally skeptical of pessimistic induction arguments of this sort. Perhaps, for example, you think that substituting ‘past theories of physical phenomena’ for ‘past naturalistic theories of intentional content’ throughout the induction generates an obviously bad argument form, since some current explanations of physical phenomena will not be regarded in the future as unworkable and false, but rather will continue to be seen as true. This gets us into deep waters, since in order to evaluate this
response we’d need to know what view of current physical theory was being accepted. Is the explanation of some phenomenon in chemistry true, for instance, if it appeals to entities (i.e. protons, neutrons, electrons) that may very well have no place in our final theory? I cannot answer these questions here. I can note, however, that the *general skeptic* of pessimistic inductions in both physics and philosophy seems at least to assume to following premise: there is no reason to believe that physical and philosophical explanations differ in any fundamental way. This premise is required to make the parity between the two arguments even somewhat persuasive, and it seems to be to be false, though I won’t argue the point more here—in any case, noting its importance is enough for my purposes here.396

Returning to the first point about physics fixing all the facts, non-reductive physicalists and Spinozists agree that physics is not holistically fact-fixing, but for very different reasons. Non-reductive physicalists require irreducibility between ontological levels—between psychological and biochemical truths, and between truths of biochemistry and truths of physics, perhaps. But in arguing for this irreducibility a crucial problem arises for the view. Given the rather startling success of physics in unifying disparate areas of study—observations about movements and quantities of heat

396 See chapter 2 of the present work for extensive discussion of some reasons others have given—van Inwagen, perhaps most forcefully—for thinking that we should be skeptical of philosophical arguments and claims, in general, to a far greater degree than we are of any corresponding claims in the natural sciences. Of course, even if we accepted this claim, we might wonder whether naturalistic theories of content are any worse off than philosophical claims more generally. Perhaps this is the real source of resistance to arguments of this sort. To this, my short reply is: yes, naturalistic theories are worse off, in part because they are sufficiently specific to reveal substantial, and I think intractable, problems with individuating intentional content. Other theories (i.e. primitive intentionality for the dualist and perhaps for Spinoza) may be dissatisfying precisely because they are vague, but that is another question entirely. I do not, however, think that they are specific enough to face the pessimistic meta-induction worry, at least in a general way (i.e. while there is something sufficiently specific to be a ‘naturalistic theory of intentional content’, there is nothing specific enough to be a ‘dualistic theory of intentional content’ that can fall prey to the meta-induction).
transformed into statistical thermodynamics, to give a popular example—the argument for irreducibility faces serious empirical challenges. Moreover and more seriously, in arguing for irreducibility between ontological levels, the non-reductive physicalist must be careful not to employ premises that can be used in a structurally similar argument for the irreducibility of the mental to the physical. She must preserve inter-level irreducibility without justifying mental-physical irreducibility with objectionable ontological import.

By contrast, the reductive physicalist and Spinoza can agree that the truths of physics fix all the facts within the realm of the physical. (Of course, physics fixes absolutely no mental facts, for Spinoza.) Spinoza simply denies, as Rosenberg states above, “[that] original intentionality is a causal role with respect to other clumps of matter (remember that science only recognizes clumps of matter and fields of force),” though Spinoza accepts that “piling up clumps of matter without original intentionality and having them participate in complex causal processes with one another won’t produce original intentionality.” Because Spinoza agrees with Rosenberg that piling up clumps of matter without original intentionality never generates genuine intentionality, no matter what sorts of causal interactions the clumps of matter engage in, he supposes that clumps of matter, considered as such, must lack original intentionality. In this he agrees with Rosenberg.

He simply disagrees that the same stuff can be talked about in causal and explanatory terms as both physical and mental at the same time, given his commitment to the EB. But since Rosenberg simply assumes that intentionality results from the physical properties of the underlying material, he does not even address Spinoza’s position. So Spinoza should be unswayed by Rosenberg’s first two arguments, and he can happily
accept Rosenberg’s Pessimistic Meta-Induction for intentional content. In fact, Spinoza can do more than this, since the inductive failure of all naturalizing projects in semantics is simply a deductive result of Spinoza’s views in fundamental metaphysics. The explanatory barrier between the attributes positively requires that all such teleosemantic or otherwise naturalizing projects fail, given the basic metaphysical structure of God as the being of infinite attributes, each of which must be causally and conceptually isolated.

There are numerous additional arguments against physicalism that are interesting to consider alongside Spinoza’s own position, but I will instead conclude by summarizing the main results of this chapter. I argued that Spinozism can accommodate itself to physicalism, specifically by accepting the most plausible forms of many of the most promising arguments for physicalism. In the second part, I then claimed that Spinozism has substantial advantages relative to physicalism. To the extent that one finds a variety of objections to physicalism compelling while simultaneously appreciating their inefficacy against Spinoza, Spinozistic isolationism begins to look superior both to the traditional forms of dualism against which it arose, and to the varieties of materialism that supplanted it. The final chapter combines the results of the preceding ones to argue that Spinozistic isolationism can profitably be interpreted as a version of or very close rival to ‘Russellian monism’ (RM), with all the attendant advantages its various defenders take it to have. As we will see, however, Spinozistic Isolationism can solve

---

397 To mention one last example: James Ross, in “Immaterial Aspects of Thought” and again in Thought and World, poses a complicated argument, grounded in the determinacy of thought and the indeterminacy of every physical state, for the conclusion that any attempt to naturalize the mind, or generally to wholly explain thinking in physical terms, will fail. What’s fascinating to note is that Spinoza, I think, can accept the argument wholesale, since determinacy of a certain sort can be reserved as a characteristic unique to intentional or representational or, in short, Thinking states. But I do not have space to discuss the details here.
several problems that other Russellian monists cannot. Thus the version of Spinozism
offered here, buttressed by the argument for the explanatory barrier, which is central to
Spinoza’s metaphysical project, offers a perhaps unrivaled alternative in contemporary,
and not just historical, metaphysics of mind.
CHAPTER 6:

SPINOZISTIC ISOLATIONISM IN THE CONTEMPORARY DEBATE ABOUT
CONSCIOUSNESS

6.1 Spinoza’s Place in the Contemporary Debate

Spinoza’s argument for the explanatory barrier (Ch. 3) between the attributes supplies a key, missing ingredient in arguments for his major metaphysical theses, including substance monism, mind-body identity, and human ignorance of more than two attributes (Ch. 1). Spinozism plus an argument for the EB, or Spinozistic Isolationism, can offer powerful objections to dualism (Ch. 4) while retaining several advantages over, and accommodating the best arguments for, physicalism (Ch. 5). In this final chapter, I situate Spinozistic Isolationism within the constellation of the most promising contemporary views in philosophy of mind, once one’s dissatisfaction with dualism and physicalism has crystalized. Since dualism and physicalism are still the most widely held views, the territory beyond them is a morass of contested terminology. So I begin by trying to clearly state Spinozistic Isolationism (henceforth: Isolationism) in contemporary terms, without doing violence to the position’s innovative aspects. I then situate it with respect to a family of various forms of panpsychism or panpsychist Russellian monism, a cadre of currently popular contenders for bridging the physicalism-dualism gap.

One difficulty with mapping Isolationism onto contemporary terrain concerns a
fairly widespread practice of inconsistently distinguishing definitions of a position from central motivations for that position. Central motivations include not only the explicit arguments given for a given view, but also the desiderata that must be respected in drawing distinctions central to fairly characterizing the position, an issue that will be particularly important in understanding the Russellian monist views discussed shortly. I think philosophical positions should be evaluated holistically, in a manner that respects both their formal definitions and their central motivations and justifying arguments. So evaluated, Isolationism emerges as quite close to Russellian panpsychism; in fact, I suggest that the latter is just a parochial way of endorsing the former.

So it’s important to keep the asymmetry between what principles or argument premises Spinozistic Isolationists can use or fruitfully motivate more convincingly than some nearby contemporary Russellian view—uniformity intuitions, rejection of brute emergence, a top-down view of grounding within a given attribute—from what Isolationism requires for its defense. Spinozistic Isolationism is defended not with contemporary arguments but on Spinoza’s terms, and is supplemented by the previous chapters’ defense of the explanatory barrier, one of its central features (along with PSR). I argue later at several points that Spinoza can employ an argument or better defend a central motivation for some closely related view. When I do, note that it is no objection at all to Isolationism to point out that the argument being discussed fails in some other way or that this central motivation should be rejected. My way of drawing such connections is always conditional unless it is critical. If one finds an argument cogent, then Isolationists can accept it, better defend one of its key premises, or more convincingly respond to certain objections. That is the conditional aspect. The critical aspect is that, in the course
of exploring the Isolationist’s defense of some contemporary claims, the central motivations for the contemporary position are often revealed as in tension with each other in ways the Isolationist can avoid. This, of course, provides further grist, indirectly, for the Isolationist mill.

In contemporary terms, then, Spinozistic Isolationism is a token identity theory of mental states or events and physical states or events. From the Ethics’ 2p7: the extended mode and the idea of that mode are one and the same thing. Its key innovation is the explanatory barrier, which precludes any causal, conceptual, or explanatory relations between any thinking things, properties, states, or events, conceived as such, and any extended things, properties, states, or events, conceived as such. Consequently, Isolationists reject any cross-attribute (i.e. physical-mental or vice versa) grounding claims, since grounding claims express dependence relations that are either themselves explanatory or entail explanations across attributes. From another angle, we can start with what Isolationism is not. Isolationism is not Davidson’s anomalous monism, for reasons Davidson acknowledges. Most importantly, Isolationism requires that causal contexts be opaque, whereas Davidson explicitly affirms only explanatory, but not causal, opacity.

Isolationism is also not epiphenomenalist property dualism, since epiphenomenalists characteristically affirm that mental events are caused by physical

---

398 As mentioned in Chapter 5, Spinoza’s nominalism substantially complicates the statements of his position that immediately follow. I mostly ignore these complications, though occasionally I will point out some oddity of the Isolationist view, relative to prevailing expectations, which often results from precisely this aspect of Spinoza’s view.

399 Davidson, “Spinoza’s Causal Theory of the Affects,” fn. 13, pp. 110-111; these issues are discussed at greater (though perhaps not enough) length in the Introduction.
events, which themselves are causally inefficacious in the physical realm.\footnote{Robinson, “Epiphenomenalism,” Intro} By contrast, Isolationism broaches no asymmetry between the mental and physical. They are equally fundamental and equally causally efficacious, as the Isolationist’s commitment to token identity theory makes clear. Finally, Isolationism is also not, or not just, a garden-variety form of parallelist property dualism, for four reasons.\footnote{I also strongly endorse Strawson’s annoyance, voiced in \textit{Consciousness and its Place in Nature}, 242, fn. 120, with contemporary philosophers’ attempts to map views that either explicitly accept a non-traditional subject/property distinction, as does Strawson, or reject realist views of properties in favor of nominalism, both of which Spinoza arguably does. So if one can only understand Spinoza’s view within a property dualist framework, then I judge this preferable to not grappling with his views at all. It’s just important to recognize then that probably Spinoza’s view of properties and their relations to subjects is either strictly nominalist or much closer to Strawson’s view that there is no real distinction between subjects and their properties in many cases (certainly as regards finite modes of substance and their properties, and also potentially as regards the relationship between substance and its attributes).} Leibniz, the most famous defender of parallelism, seems to think that bodies result, or are grounded in, some features of monadic minds and their ideas or perceptions, even if this grounding or resulting relation remains somewhat unclear in Leibniz’s work.\footnote{Adams, \textit{Leibniz: Determinist, Theist, Idealist}, Ch. 9} This, of course, violates the explanatory barrier as understood it here. Second, Spinoza is a property pluralist—there are infinitely many attributes, rather than the typical parallelist two. Third, both Spinoza’s reasons for affirming infinitely many attributes, and his reasons for denying causal and explanatory relations among any two sets of attributes, are wholly general, independent of particular features of Thought or Extension. They are thus not held hostage to particular Leibnizian arguments about the impossibility of either inter-monadic causation or of explaining perception in mechanical terms (as in Leibniz’s mill).\footnote{See again Adams, \textit{Leibniz: Determinist, Theist, Idealist}, as well as \textit{Monadology} 17.} Fourth and finally, since identity theory is often taken to be a competitor to parallelism,
Spinoza’s embrace of the former seems to entail that he is not just a garden-variety proponent of the latter.

Isolationism fails to map smoothly onto any single contemporary position but fruitfully illuminates several of them while sharing some of their key motivations. So, in the interest of making contact with several lively current debates, I proceed in this chapter as follows. I first argue that there is substantial strategic overlap between Spinoza’s arguments for Isolationism and Chalmers’ Hegelian argument for constitutive Russellian monism. Nonetheless, the Isolationist cannot accept Chalmers’ formulation of the Russellian monist position, as Chalmers states it in terms of a grounding relationship between phenomenal and physical properties. But there are both inter-attribute grounding claims (i.e. between phenomenal and physical properties) and intra-attribute grounding claims (i.e. whether microphenomenal properties ground macrophenomenal or vice versa) to be adjudicated here. So I then argue that Isolationist considerations can be used to better motivate Priority Cosmopsychism at the expense of micropsychist Russellian monist views. This advantage results from the Cosmopsychist’s ability, supplemented by Isolationism, to provide a superior account of intra-attribute grounding and the most compelling hope for resolution of the combination problem for panpsychism.

After discussing Isolationism’s relationship to Chalmers’ style formulations of the Russellian monist position and weighing in on whether Micropsychism or Cosmopsychism offers the best account of the panpsychist position, I then discuss Isolationism with respect to Russell’s own views. I call attention to Stubenberg’s underappreciated contention that grounding plays no role at all in Russell’s formulation of his own position. Note that if this is right, the Spinozistic Isolationist may be able to
formulate his position in the Russellian vein without grounding, thus capitalizing on some of the argumentative considerations in Russellian monism’s favor without incurring the costs of non-Isolationism. I then end by discussing how Spinozistic Isolationism can evade two important objections to Russellian monist views as typically formulated.

In the end, I think Spinoza’s advantages, relative to nearly all the most widely discussed contenders in the metaphysics of mind, are substantial. Isolationism arguably best satisfies the Russellian monist’s desire to cut a middle path between dualism and physicalism. It best respects highly plausible constraints on the intelligible emergence of ‘higher’ phenomena from low-level ones, and solves the vaunted combination problem for panpsychism. Most importantly, Isolationism contributes asymmetrically to contemporary debates. It’s a comprehensive metaphysical system that provides more principled justification for certain key premises in these debates while not depending on their resolution for its defense. This asymmetry helps to further motivate the Isolationist’s rejection of any causal, explanatory, or grounding relations across attributes. But it also serves as indirect evidence for the benefits of defending positions firmly grounded in the comprehensive metaphysical systems of philosophy’s history.

6.2 Russellian Monism and Chalmers’ Hegelian Argument

Observant readers may have sensed the affinity between the arguments of the last two chapters, and what David Chalmers calls his Hegelian argument for panpsychism. This affinity is real, with historical roots deeper than Chalmers acknowledges. For the argumentative structure Chalmers employs—materialism as thesis, dualism as antithesis, Russellian monism as synthesis—corresponds to one clear line of interpretation of early
modern philosophy. On that reading, rationalist panpsychists like Leibniz and Spinoza remedy deficiencies in the dualist Descartes and the materialist Hobbes in various ways and with differing degrees of success.

One major caveat with this reading, it should be noted, is that what are currently taken to be the best arguments for the respective positions were not always deemed so. Papineau and others have noted, for example, that the causal closure argument for physicalism gained prominence only slightly later, at least partially in response to developments in the natural sciences. Despite this qualification, the contemporary debate plays out in much this way, with Derk Pereboom taking a more Leibnizian interpretation of Russellian monism and Nagasawa and Wager, as well as Goff, going in for more Spinozistic interpretations. But as new arguments become popular they engender new formulations of classic positions, which often substantially change the terms of the debate. Hence I want to outline several important affinities between the Isolationist strategy pursued in the present work, especially in chapters 3-6, and Chalmers’ Hegelian argument for Russellian monism. But I then give three reasons that Isolationists cannot ultimately accept Chalmers’ formulation of constitutive Russellian panpsychism.

In “Panpsychism and Panprotopsychism,” David Chalmers provides a now classic defense of what he calls the Hegelian argument for constitutive Russellian monism. The basic strategy is comprised of two parts, a set of arguments and a story about why the resultant view deserves the ‘Russelian’ moniker. Spinozistic Isolationism captures both, earning the Russellian designation in particularly beautiful fashion and arguably

---

404 Papineau, “The Rise of Physicalism,”
remaining much closer to Russell’s original driving intuitions, as we shall see in a later section. But first, I’ll say a bit about Chalmers’ definitions of the position, and about his arguments in support of it.

Chalmers defines *constitutive Russellian monism* as simply the conjunction of broad physicalism with the denial of narrow physicalism. Chalmers’ definition of the relevant terminology is worth quoting at length:

> Narrowly physical properties are “microphysical role properties, such as the dispositional property associated with having a certain mass, or the second-order property of having a property that plays the mass role. We can say that broadly physical properties are physical role properties along with any properties that realize the relevant roles: categorical bases for the mass dispositions, first-order properties that play the mass role...In effect, narrowly physical properties include structural properties of microphysical entities but exclude quiddities [the fundamental categorical bases of the microphysical dispositions characterized in physics], while broadly physical properties include both structural properties and quiddities. Here a structural property is one that can be fully characterized using structural concepts alone, which I take to include logical, mathematical, and nomic concepts, perhaps along with spatiotemporal concepts...[Then] we can distinguish narrow physicalism, which holds that phenomenal truths are grounded in narrowly physical truths, from broad physicalism, which holds that phenomenal truths are grounded in broadly physical truths.”

*Constitutive Russellian monism*, which then takes various forms depending on the character of the relevant quiddities, includes *constitutive Russellian panpsychism*, according to which two grounding relationships obtain. First, macroexperience is grounded in microexperience—this is the *constitutive panpsychist* aspect—and second, at

---

405 Chalmers, “Panpsychism and Panprotopsychism,” 261-262

406 Chalmers, “Panpsychism and Panprotopsychism,” 255-256
least some of the quiddities, or categorical bases for the dispositions given by physics, are microphenomenal properties—this is, for Chalmers, the *Russellian* aspect.407

Not all parties in the contemporary debate think that the best way of cashing out the difference between what Chalmers calls narrow and broad physical properties is in terms of structural/dynamical/dispositional properties, on the one hand, and non-structural-and-dynamical/categorical properties, on the other. As Alter and Nagasawa note, what we are really after here is some way of laying out a central distinction between two sorts of properties, those found in physics and those that do not show up there. As they put things, this central distinction might track: “Extrinsic vs. intrinsic properties; dispositional vs. categorical properties; relational vs. non-relational properties; or structural-and-dynamic vs. non-structural-and-dynamic properties; or even epistemic contrasts between physical dispositional properties and categorical properties about which we are ignorant.”408

But Alter and Nagasawa still state the Russellian monist’s view in grounding terms. For they define inscrutables as follows:

We will refer to properties (if such there be) that ground the physical/structural relations physics describes as *inscrutables*...[and then propose] that Russellian monism be understood as the conjunction of three claims: *Structuralism about physics*: the basic properties physics describes are structural/relational properties; *Realism about inscrutables*: there are inscrutables, the natures of which are not wholly structural/relational; *(Proto)phenomenal foundationalism*: at least some inscrutables are either phenomenal or protophenomenal properties.409

---

407 Chalmers, “Panpsychism and Panprotopsychism,” 254

408 Alter and Nagasawa, “What is Russellian Monism,” 427; but see also Pereboom, *Consciousness and the Prospects of Physicalism*, 89. As Alter and Nagasawa note, the last epistemic contrast does not seem quite up to the task of defining the metaphysical position they are interested in. I agree, but would simply add that the epistemic orientation of Russell’s work (and Spinoza, insofar as he is interested in the problem I’ve described in chapter three), while it shouldn’t show up in the position’s definition, might help illuminate unnecessary presuppositions in contemporary definitions (such as the focus on grounding, to be discussed later).

409 Alter and Nagasawa, “What is Russellian Monism,” 425
At this point, it’s worth noting only that not all ways Alter and Nagasawa discuss of drawing the central distinction necessarily involve the grounding of one category of property by another. For instance, neither the intrinsic/extrinsic nor the non-relational/relational distinctions appear to demand that the former category ground the latter. I return to this point below, especially in discussing Russell’s own views.

With the position defined, we can begin examining the considerations in its favor. Chalmers motivates this constitutive Russellian monist position with his so-called ‘Hegelian’ argument. In this argument, the original thesis is materialism, and its support is a version of what I last chapter called the Causal Closure argument for physicalism. Here’s Chalmers’ presentation of the argument.

1. Phenomenal properties are causally relevant to physical events.
2. Every caused physical event has a full causal explanation in physical terms.
3. If every caused physical event has a full causal explanation in physical terms, every property causally relevant to the physical is itself grounded in physical properties.
4. If phenomenal properties are grounded in physical properties, materialism is true.
5. Materialism is true.⁴¹⁰

Chalmers’ argument, unlike my earlier presentation of the Closure argument last chapter, does not foreground its commitment to the distinctness of mental and physical

---

⁴¹⁰ Chalmers, “Panpsychism and Panprotopsychism,” 251
properties. It also employs the language of grounding, in the third and fourth premises.

By grounding, Chalmers has in mind a constitution relation:

> Truths about consciousness are grounded in physical truths if all truths in the first set obtain wholly in virtue of truths in the second set obtaining. This requires at least that there is a metaphysically necessary connection between these truths, in that it is impossible for a world to be physically like ours without that world being phenomenally like ours.\(^{411}\)

After describing the remainder of Chalmers’ argument, I return to this point about the grounding relation’s role in the argument.

In opposition to the materialist thesis is the antithesis, dualism, which is supported by Chalmers’ preferred form of the conceivability argument, albeit one presented without the use of his controversial two-dimensional semantics. The argument’s guiding intuition is that in certain circumstances one can use epistemic possibility claims about the conceivability of physically identical but phenomenally different creatures—zombies or phenomenal inverts—to establish metaphysical conclusions. A simple statement is as follows, where P is the conjunction of the universe’s microphysical truths and Q is some phenomenal truth, like ‘Someone is conscious,” or “Tim is conscious.”

1. P&~Q is conceivable.
2. If P&~Q is conceivable, P&~Q is metaphysically possible.
3. If P&~Q is metaphysically possible, materialism is false.
4. Materialism is false.\(^{412}\)

\(^{411}\) Chalmers, “Panpsychism and Panprotopsychism,” 248

\(^{412}\) Chalmers, “Panpsychism and Panprotopsychism,” 249
As before, there are numerous remarks one might make on this version of the argument, and Chalmers has offered nearly endless precisifications and defenses of more complicated versions of the argument elsewhere. Typical physicalist strategies, naturally enough, include challenging the coherence of the initial conceivability claim in the first premise, as well as denying the link between conceivability and possibility, in whatever form, in the second. As the development of these debates will be familiar to most readers, and as it is not central to my argument here, I’ll not spend much time on them now.

Respecting the considerations that give both arguments their force against opponents drives the search for some middle way between physicalism and dualism. This is the central affinity between Chalmers’ strategy and the Isolationist one pursued here, where serious difficulties with dualist explanations of our knowledge in many quite ordinary cases (Ch. 3-4) are juxtaposed to similarly significant problems with physicalism (Ch. 5), by Isolationist lights. This point is worth stressing first because it highlights underappreciated common ground between Spinoza’s Isolationism and an extremely popular family of contemporary positions, the subject of extensive discussion in recent years.

It’s also crucial because occasionally Chalmers’ formulation of Russelian monism can seem nearly to exhaust the middle ground between physicalism and

\[\text{\ldots}\]

\textit{413} See, for example, “The Two Dimensional Argument Against Materialism.”

dualism. This impression is misleading, depending as it does at least partially on Chalmers’ decision to formulate the relationship between the microphenomenal and microphysical in grounding terms. For Chalmers, the microphenomenal provides the *categorical basis* or *realizes* the relevant microphysical dispositions, but this contention may be rejected. Still, the Isolationist does need to show clearly how he can capitalize on the substantial benefits of some such middle position while rejecting Chalmers’ characterization of it. In the process of doing so, several reasons in favor of the Isolationist’s choice thereby serve to put pressure on Chalmers’ desired formulation.

Chalmers desired Hegelian synthesis is constitutive Russellian monism, the conjunction of broad physicalism with the denial of narrow physicalism, and the constitutive Russellian panpsychist form is particularly important for our purposes. Russellian monism’s central virtue, on the standard story, is its ability to evade both the causal and conceivability arguments for materialism and dualism, respectively. It evades the conceivability argument because of the special character of the quiddities or inscrutables, that is, the properties that ground or realize the relevant physical roles. These quiddities, whether conceived as non-relational, intrinsic, categorical, or non-structural-and-dynamic, are either such that 1) we are ignorant of them, or 2) such that we cannot justify the claim that we can conceive of zombie versions of them.

---

415 This is particularly clear near the end of Chalmers, “Panpsychism and Panprotopsychism,” where he discusses all manner of different views for what the underlying quiddities might be like—non-structural and dynamical physical properties, ordinary phenomenal properties as in idealist versions, microphenomenal, and panprotophenomenal properties, or even non-structural physical properties (see also Pereboom, *Consciousness and the Prospects of Physicalism*, for more suggestions in this vein). All these options, combined with the question of whether the relevant grounding relations are a priori or a posteriori, and the question of whether the view is constitutive or some form of emergentist, seems to exhaust a great deal of the available space between physicalism and dualism.
Take (2) first. The Russellian monist’s claim, so interpreted, is that we cannot conceive of creatures that share all our broadly physical properties (i.e. the same narrowly physical properties, plus the categorical/intrinsic/non-relational, etc. base properties) but lack our phenomenal properties. In part, this results from the fact that phenomenal properties just are the categorical/intrinsic/non-relational bases; so duplicating them duplicates just what we must conceive as somehow different. So premise 1 of the conceivability argument is false. If instead we opt for interpretation (1), then the thought is that, given our radical ignorance of these categorical/intrinsic/non-relational properties, we are no longer justified in the inference from the conceivability of broadly physical zombies to their metaphysical possibility. So premise 2 of the conceivability argument is not true.

Meanwhile, Russellian monism evades the causal argument more straightforwardly. Each occurrence of ‘physical properties’ in the causal argument given above requires an interpretation, as either broadly or narrowly physical. If broadly physical properties are meant, then the Russellian monist simply accepts the argument’s conclusion. If narrowly physical properties are meant, then the closure premise can happily be rejected, since, as Chalmers puts it, “a full causal explanation of narrowly

---

416 This is the standard telling of the story. See, for example, Chalmers, “Panpsychism and Panprotopsychism,” 257—but note that this standard telling has been challenged by those who argue that we can conceive of beings that duplicate our broadly microphysical properties (i.e. structural properties plus the relevant micro-quiddities) but lack the human subject-level qualitative experiences we ourselves have. This objection thus contends that the Russellian monist as Chalmers defines her does not evade the conceivability argument as smoothly as my exposition here suggests. This objection seems strange to me, since it appears to presuppose that human subject-level qualitative experiences are not merely composed of their constituent parts (thereby embracing a form of holism) while rejecting holism in other ways (by embracing micropsychism at the expense of Cosmopsychism). It should be clear, however, that for the Isolationist view I am defending, which, as explained immediately below, rejects the conceivability argument due to its general modal commitments, after providing an argument for those commitments (the argument for substance monism and necessitarianism in Ethics Part I), these details about the conceivability argument pose less of a problem.
physical events will involve broadly physical properties; a causal explanation wholly in
terms of narrowly physical properties is incomplete."^{417}

Of course, a hard-nosed physicalist might quarrel with these claims to incompleteness, since the narrowly physical properties alone show up in our most developed natural sciences. But in objecting this way, of course, this hard-nosed physicalist makes clear her commitment to the rather implausible claim that the properties of physics suffice for exhaustively describing the world and all its truths. This strategy inherits the litany of defects discussed at length in the previous chapter, including problems with intentional content, defining ‘the physical’ for the sake of describing the view, and others, as well as the conceivability worries discussed here. Moreover, this hard-line view appears incompatible with most forms of non-reductive physicalism, the dominant way of endorsing the position.

Yet, despite the Hegelian affinities between the Isolationist’s and Chalmers’ strategy—materialism and dualism as unsatisfying thesis and antithesis awaiting synthesis—there are at least three key differences. First, the Isolationist is not motivated by the problem of integrating qualia, in particular, into the modern scientific worldview. Second, the Isolationist, as a Spinozist, has necessitarian modal commitments. Both these points affect the Conceivability argument, which the Isolationist cannot accept. Premise 2 of that argument, assuming that P&¬Q describes a non-actual situation, is not metaphysically possible on the standard reading of Spinoza’s modal views.\(^{418}\) But this has

\(^{417}\) Chalmers, “Panpsychism and Panprotopsychism,” 258

\(^{418}\) There is, of course, extensive dissent on this question, references to which were given in previous chapters. Garrett, “Spinoza’s Necessitarianism,” gives a classic statement of the necessitarian reading, while Curley and Watski, “Spinoza’s Necessitarianism Reconsidered,” offer dissent, and
nothing to do with qualia in particular—it is simply a general consequence of Spinoza’s modal commitments.

More centrally, as I discussed in detail at the end of the last chapter, Spinoza’s Isolationist sees central and insuperable difficulties with integrating both the intentional/representational and the qualitative aspects of Thought into the realm of Extension. Unlike Chalmers and many other contributors to the debate over qualia, who often seem to tacitly assume that the problem of representation is soluble in physicalist terms while that of qualia is not, Spinoza thinks both issues stand or fall together.419

Again, though, his reasons are first and foremost general, based on the explanatory barrier between the attributes. Of course, he can also accept many of the arguments, some of which were discussed last chapter, in favor of the impossibility of reducing intentional properties and relations to physical ones or explaining such properties in physicalist-friendly terms. But the more general nature of the Isolationist’s strategy, it should be noted, provides the first example in this chapter of the broader asymmetry emphasized above. For the physicalist gives no objection at all to Spinoza’s Isolationism by claiming that one or another particular argument discussed at the end of the last chapter fails, since Spinoza’s position rests on independent grounds.

Third and most importantly, Chalmers description of the Russelian monist position uses the notion of grounding to explicate its key claims. The Russelian monist,

Newlands, “Spinoza’s Modal Metaphysics,” seeks to reconcile the two camps. Even if Spinoza is not best interpreted as a necessitarian, however, it remains wholly uncontroversial that the problem of qualia is not central to his thought.

419 This claim is compatible with Della Rocca’s contention in Spinoza, for example, that Spinoza is a representationalist about qualia. In that case, the representational is the mark of Thought and qualitative properties are just representational properties, both equally resistant to being caused or explained by anything Extended, on pain of violating the EB.
according to Chalmers, accepts broad physicalism while denying narrow physicalism, thereby affirming that phenomenal truths are grounded in broadly physical truths. These broadly physical truths include the microstructural properties along with the corresponding quiddities. The Isolationist, as already noted, must reject this cross-attribute grounding relation. The extensive network of considerations in favor of this move cannot be rehearsed in detail here; they constitute the majority of the present work, especially chapters 3-4. But I do want to call attention to two key points, which I will tackle in reverse order in what follows. First, as Stubenberg has convincingly argued, I think that Russell’s central insights can be clearly stated and defended without employing the notion of cross-attribute grounding central to the contemporary debate. This gives the Isolationist space to define her view with Russell’s help, even if Russell would reject both her reasons for rejecting cross-attribute grounding (i.e. the explanatory barrier) and her final metaphysical picture (i.e. Spinozistic Isolationism).

Second, however, it’s crucial to see that there are really two grounding questions in play here, between the (micro)phenomenal and the (micro)physical, on one hand, and between microphenomenal and macrophenomenal, on the other. Chalmers displays a clear preference for a “bottom-up” conception of grounding—note that he just defines Russellian monism as the view that macrophenomenal truths are grounded in broadly microphysical truths. But this definition, as I think Chalmers would agree, is more heuristic than normative. Indeed, debating this latter, intra-attribute grounding question contributes prominently to discussions of the relative merits of so-called micropsychist and cosmopsychist monistic views, especially as regards the most prominent objection to panpsychism, the combination problem. So before turning to Isolationism’s relation to
Russell’s formulation of his views, I discuss Isolationism’s contribution to this intra-
attribute grounding debate and its resolution of the combination problem.

6.3.1 Varieties of Panpsychism

We can start by describing two types of panpsychism, which diverge precisely on
the question of whether grounding within the realm of the phenomenal (within Thought,
for the Isolationist) is “top-down” or “bottom-up.” The micropsychist proceeds bottom-
up and affirms that some (usually all) microphysical entities, like quarks and electrons,
are conscious. The cosmopsychist moves from the top down, arguing that the cosmos is
conscious and that ordinary (i.e. human and animal), as well as any micro-consciousness
is ontologically posterior to and derivative from it.

Following Chalmers’ discussion, I will be concerned with constitutive forms of
the panpsychist thesis. Constitutive Micropsychism holds that macroexperiences, like my
conscious experience of seeing red or experiences of other humans and perhaps higher
animals, are wholly or partially grounded in microexperiences. Constitutive
Cosmopsychism, relatedly, holds that ordinary conscious experiences, as well as any
microexperiences, are wholly or partially grounded in the ontologically prior conscious
experience of the cosmos. Of course, in speaking of Spinoza’s Isolationist view in
particular, God or Nature takes the place of the cosmos.

Chalmers contrasts constitutive panpsychist views with emergent panpsychist
views in the standard way, whereby the emergentist will either rely on contingent laws of
nature or some law-like but non-constitutive relations among micro-, macro-, or cosmo-
phenomenal experiences. These issues arose last chapter in discussing Jessica Wilson’s
opposition to Karen Bennett’s views, and occur ubiquitously in the literature. Chalmers notes, and I agree, that such views typically inherit problems structurally similar to the dualist’s, and thus do not offer the desired ‘Hegelian synthesis.’ For similar reasons, I’ll also only briefly discuss Chalmers’ panprotopsychism.

On panprotopsychism, the fundamental micro-properties are not phenomenal, but they constitute phenomenal properties when arranged correctly, are not merely structural properties, and a priori entail phenomenal properties. Chalmers is careful to define protophenomenal properties so that the emergence of phenomenal properties from them is not, or does not seem, brute in an objectionable way. But I see no way of 1) preserving non-brute emergence while 2) providing any content to the definition of protophenomenal properties, and am thus content to ignore them.

Naturally, this move is too quick to be satisfactory to those who favor panprotopsychism, but there are extended defenses available of my claim here. Amy Kind, for instance, argues at length that any panprotopsychist view collapses either into a version of neutral or physical monism. The thought here is not, notably, that panprotopsychism does not constitute an alternative to physicalism or dualism; it incontestably does. But this is not all that Chalmers wants the view to do, as becomes clear by his presentation of it as a “new antithesis” to panpsychism. Here, if we are to

420 Chalmers, “Panpsychism and Panprotopsychism,” 259-60

421 Kind, “Pessimism about Russellian Monism,” 419-420 (though see also 408, 411-414) puts the point this way: “Development of this [Russellian monist] account will involve a choice between two options: phenomenal monism and physical monism. This choice hinges on the issue of whether the inscrutables are phenomenal. If this choice is not flatly the choice between dualism and physicalism, it is at the very least not too far removed from that choice. And so we are thus essentially back where we started.” See also Strawson, Consciousness and Its Place in Nature, for similar worries.

422 Chalmers, “Panpsychism and Panprotopsychism,” 259
take the thesis-antithesis-synthesis structure of his paper at all seriously, panprotopsychism must be regarded as an interesting alternative to the preceding panpsychist synthesis, including the physicalism and dualism it subsumes. I just join Amy Kind in rejecting this further claim. Panprotopsychism is either not distinctive enough to provide a true alternative to physicalist or phenomenal monism, or it buys sufficient distinctiveness at the cost of inheriting problems structurally similar to those afflicting brute emergence views.

With panprotopsychism placed to the side, the debate between Constitutive Micropsychists and Constitutive Cosmopsychists can be seen as integrating two independent but mutually-reinforcing questions: what is the proper understanding of the grounding relation, and what is the best solution to the combination problem for panpsychism. I discuss the interaction of these two questions, and the Isolationist contribution, in the next section.

6.3.2 Priority Cosmopsychism, Micropsychism, and Spinozistic Isolationism

Even though Isolationists cannot be Russellian monists in Chalmers’ sense, because of their rejection of cross-attribute grounding, I have suggested that Isolationism can fruitfully contribute both to the question of which version of the grounding relation within Thought is best, and to solving the combination problem for panpsychism. Here I flesh out this suggestion by starting with an even more general question: once one has been convinced, either by Chalmers’ Hegelian argument or on other grounds, to embrace fundamental mentality, what sort of view should one defend?
This is not necessarily a question with an obvious answer. Strawson, whose view I’ll discuss in more detail in a later section, notes that one might be a micropsychist but not yet a panpsychist, holding that some but not all of fundamental reality is experiential. However, he notes that such a view is wildly theoretically unattractive, and might even count as a form of dualism.\(^{423}\) Now, while I don’t think that merely labeling a position as dualist counts as any argument at all, I do think that there are strong considerations in favor of regarding fundamental mentality as ubiquitous.

However, the need for additional argument here is not always explicitly recognized. Sometimes, philosophers working in this area pay lip service to the prospect of non-uniformity but then make arguments that seem to commit them to its impossibility. For instance, Chalmers strictly defines distinctively Russellian panpsychism as “the view that some quiddities are microphenomenal properties.”\(^{424}\) But he then characterizes the advantages of Russellian panpsychism in a way that tacitly presupposes that all quiddities will be microphenomenal in the same way. He says:

The Russellian panpsychist addresses two metaphysical problems—what is the place of phenomenal properties in nature, and what are the intrinsic properties underlying physical structure?—and in effect answers both of them at once.\(^{425}\)

The presupposition announces itself in the general framing of the two problems. For clearly, if only exactly five quiddities were microphenomenal properties, many microphenomenal properties were not quiddities, and some physical dispositional properties still had their categorical bases left unaccounted for, Russellian monism

\(^{423}\) Strawson, *Consciousness and its Place in Nature*, 25

\(^{424}\) Chalmers, “Panpsychism and Panprotopsychism,” 254

\(^{425}\) Chalmers, “Panpsychism and Panprotopsychism,” 254
wouldn’t be much of a general solution to these two metaphysical problems. Only once uniformity reenters by the back door does this way of framing things make any sense.\footnote{Another way to see the implausibility of micropsychism but not panpsychism relies on a point made by Thomas Nagel, “Panpsychism,” whose basic thought is that we cannot restrict which atoms end up in brains (thereby becoming conscious) and which do not, so that a view on which there is a privileged class of conscious fundamental particles, requiring such a restriction, will not work.}

Alter and Nagasawa echo positively this ‘solve two problems at once’ strategy, and on the face of it, the suggestion has its advantages.\footnote{Alter and Nagasawa, “What is Russellian Monism,” 444} As they put it, consciousness wants a job to do in the current picture of the world, and physics has an open position for grounding some spatiotemporal structure. But only if phenomenal properties are uniformly distributed, grounding all of the structure that needs grounding, will both problems actually be resolved. But skeptics of the Russellian monist project are unlikely to be swayed by this offer to solve two problems at once as an argument, at least as it stands. For even if one ignores those structuralists who deny that physics really does need help grounding its fundamental structures, the Russellian monist’s suggestion is just a little too convenient.\footnote{On this ontic structural realist picture, see Ladyman and Ross, \textit{Every Thing Must Go}.}

For one can appreciate Russell’s insight—about the highly abstract, structural information provided by modern physics—wholly apart from one’s commitments on the mind-body question. Many physicalists agree with Russell about the content of physics while ignoring or dismissing the panpsychist positions discussed here.\footnote{Moreover, as we will see at the end of this chapter, some physicalists vociferously deny Russell’s claim on at least one interpretation.} Only if one starts out with both problems in view does the solution of solving them at the same time seem elegant, even necessary. Otherwise it just looks \textit{ad hoc}. Indeed, it might look \textit{ad hoc}.
hoc even if its elegance is actually a sign of its truth, and thus serves much better as contemplative material for the converted, rather than as ammunition for the evangelizing.

Moreover, as Stubenberg has pointed out, and as we will see more in a later section, Russell’s own, more radical view is agnostic about the ubiquity and uniformity of the underlying phenomenal properties. Russell thinks that we only know the intrinsic nature of events here, in our own case. Strictly speaking, provided that other events, including events in the conscious lives of others, have the structural role properties physics ascribes to them, they may be fundamentally different, or much the same, intrinsically. The state of things with others might well be unimaginable, and in any case it’s impossible to know what it is. This is just one indication of a point Stubenberg has made repeatedly, that Russell’s orientation in developing his view was in some ways fundamentally different, and decidedly more epistemological, than the constellation of positions that presently bears his name.430

Spinozistic Isolationism, by contrast, fares much better in motivating uniformity of the microphysical as well as the microphenomenal, though it does not, of course, motivating the latter’s grounding the former. With its explicit commitments to naturalism and the PSR, it thus justifies Chalmers’ desired solution—ubiquitous mentality—better, in at least two ways. First, one can wield the PSR against the specific possibility, and many similar ones, outlined above—that exactly 5 quiddities might be microphenomenal properties, but that there remains many microphysical role properties without intrinsic bases and many microphenomenal properties doing no work at all.

430 See Stubenberg, “Russell, Russelian Monism, and Panpsychism,” for more on the relationship between Russell’s actual views and the contemporary views that claim common lineage with them.
For absent some other cause for why things occur in this specific way, Spinoza concludes, they could not do so. In 1p8s2 of the *Ethics*, Spinoza considers a superficially similar example: “If twenty men exist in Nature…it will not be enough (i.e. to give a reason why twenty men exist) to show the cause of human nature in general; but it will be necessary in addition to show the cause why not more and not fewer than twenty exist.” Here, though, Spinoza’s thought is not that a seemingly arbitrary state of affairs could not come about—they can and do ubiquitously—but just that closer examination will reveal causes for the exact, from our vantage arbitrary, situation that does result. By contrast, the possibility above, and the one that Russell seems to explicitly consider, whereby the physical role properties might be uniform while the intrinsic properties were radically different in each case, is supposed to answer a *fundamental* metaphysical question. So it’s hard to see, unlike with contingent matters of fact, how an arbitrary answer might be justified with reference to a larger structure, since the structure described is already metaphysically basic. Arbitrariness regarding fundamental metaphysics seems undesirable regardless of one’s view of the ultimate relationship between the microphenomenal and the microphysical.

More generally, the explanatory barrier serves as a crucial plank in Spinoza’s argument for substance monism, a key component of the Isolationist position outlined here. But if that argument succeeds, then explaining either phenomenal or physical uniformity is no problem at all. Each derives, and in the same way, from the results of *Ethics* Part I-II, which establish how God’s infinite modes manifest uniformly, and according to the same order, across God’s infinitely many attributes. The Isolationist’s God/Nature/Cosmos-first strategy, if it can be made to work, thus provides a more
satisfying justification for Chalmers’ position than his own ‘solve two problems at once’ argument can do.

Henceforth, then, experientiality will be taken as ubiquitous. The further question of this section is what form one’s panpsychism should take. Micropsychists include Chalmers (insofar as he finds constitutive panpsychism promising) and also Galen Strawson.\footnote{Similarly, from here on I distinguish Micropsychists from Cosmopsychists, with reference to the question of whether grounding is bottom-up or top-down, rather than distinguishing Micropsychists from Panpsychists, with reference to whether mentality is ubiquitous. I do so because I take the latter question as settled by the immediately preceding material.} The latter is admirably clear about his commitment to the further principle—“experience is impossible without an experiencer,” a subject of experience—examination of which brings us to the most serious problem for the micropsychist.\footnote{Strawson, \textit{Consciousness and its Place in Nature}, 26, as well as fn. 46} If there are ubiquitous experiences at the fundamental level and if the principle above is correct, then \textit{subjects are ubiquitous} at the fundamental level. There is a point of view for every electron on the micropsychist’s telling. Of course, this seems laughable to many people, but sheer counterintuitiveness is not the central issue at this stage in the dialectic.

Rather, the real question for those who have traveled this far down the panpsychist path is the so-called combination problem.\footnote{Strawson, \textit{Consciousness and its Place in Nature}, 26, acknowledges this as one of the central challenges for his view. All of Part III of Bruntrup and Jaskolla (eds), \textit{Panpsychism: Contemporary Perspectives} is dedicated to it.} William James provides the classical statement of the issue, which I quote at length to give a sense of its scope:

Where the elemental units are supposed to be feelings, the case is in no wise altered. Take a hundred of them, shuffle them and pack them as close together as you can (whatever that may mean); still each remains the same feeling it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean. There would be a hundred-and-first feeling there, if, when a group or series

---

\footnote{355}
of such feelings were set up, a consciousness belonging to the group as such should emerge. And this 101st feeling would be a totally new fact; the 100 original feelings might, by a curious physical law, be a signal for its creation, when they came together; but they would have no substantial identity with it, nor it with them, and one could never deduce the one from the others, or (in any intelligible sense) say that they evolved it.

Take a sentence of a dozen words, and take twelve men and tell to each one word. Then stand the men in a row or jam them in a bunch, and let each think of his word as intently as he will; nowhere will there be a consciousness of the whole sentence. We talk of the ‘spirit of the age,’ and the ‘sentiment of the people,’ and in various ways we hypostasize ‘public opinion.’ But we know this to be symbolic speech, and never dream that the spirit, opinion, sentiment, etc. constitute a consciousness other than, and additional to, that of the several individuals whom the words ‘age’, ‘people’, or ‘public’ denote. The private minds do not aggregate into a higher compound mind.434

As others have observed and Chalmers has argued for at length, calling it the combination problem is a bit of a misnomer. Chalmers’ reflection on this passage and other features of our conscious experience suggests, at least, 7 problems, organized according to the subjective, qualitative, and structural character of this experience.435 I think Spinozistic Isolationism turns many of these problems on their heads, however, and my primary aim is to sketch the outlines of that argument here. I will not, however, take pains to show precisely how or to what extent all of Chalmers’ problems are resolved, as

434 James, Principles of Psychology, as quoted in Chalmers, “The Combination Problem for Panpsychism.”

435 Chalmers, “The Combination Problem for Panpsychism,” 182-3, gives a complete list, in typically thorough fashion, as follows: the subject-summing problem of how micro-subjects combine to yield macro-subjects; the quality combination problem of how micro-qualities yield macro-qualities; the structural combination problem of how micro-structure combines to yield the macro-structure of our experience; the unity of consciousness problem, the boundary problem of how microexperiences yield a bounded consciousness, the awareness problem of how microexperiences yield awareness of qualities, and the grain problem of how underlying heterogeneous experiences combine to yield experiences with the smooth, homogeneous character of some of our macro-experiences.
it’s not 100% clear that it’s always fruitful to take others’ mapping of the territory so seriously.\textsuperscript{436}

But before tackling the combination problem directly, it’s worth pausing to note the special significance of the underlined sections from James, given the discussion of Strawson to follow. For James, the real significance of the combination problem lies in the fact that it requires \textit{brute emergence}. As he says, “This 101\textsuperscript{st} feeling would be a totally new fact.” Panpsychism on this view fails because it cannot avoid the sort of brute emergence it was explicitly introduced, on some ways of telling the story, to combat.

Galen Strawson’s recent defenses of panpsychism make particularly clear the connection between his view and the denial of brute emergence. Here I discuss it briefly to show how the Isolationist can further buttress no brute emergence arguments for panpsychism, as prelude to arguing that the Isolationist offers the best defense of the Cosmopsychist’s top-down notion of grounding and corresponding solution to the combination problem. Together, these two arguments show from yet another angle how panpsychism arises naturally out of Isolationism, which can best account for the motivations that made the panpsychist view attractive in the first place.

\textbf{6.3.4 Strawson’s No-brute Emergence Argument for Panpsychism}

In a series of recent papers, Galen Strawson has defended an argument from the impossibility of certain forms of emergence to what he calls ‘real materialism,’ which

\textsuperscript{436} See Montero, “What Combination Problem?” for several good reasons for being skeptical about Chalmers’ characterization of the combination problem in particular.
turns out, surprisingly enough, to be a form of panpsychism.\textsuperscript{437} Strawson aims for a view he characterizes as properly ‘physicalist’ and properly ‘realistic’—realistic about experience that is. His central claim is that any view with both these features entails panpsychism, given the necessity of avoiding certain objectionable forms of emergence. He starts by simply assuming \textit{Physicalism}: every real, concrete phenomenon in the universe is physical.\textsuperscript{438} By ‘concrete,’ Strawson has in mind anything that is temporally or spatiotemporally located, and ‘phenomenon’ is his most general term for anything that exists. While many philosophers now lay claim to this physicalist mantle, Strawson thinks, they do so only by courting one of two other forms of incoherence, giving up on ‘being realistic about experience’ or accepting objectionable forms of emergence.

These costs are worth paying, physicalists think, because of a tacit, anti-Russellian presupposition that \textit{they know a great deal about the nature of the physical}. In particular, Strawson attributes to them \textit{NE}: Physical stuff is, in itself, in its fundamental nature, something wholly and utterly non-experiential.\textsuperscript{439} In not ‘being realistic,’ physicalists, tacitly presupposing that “the nature or essence of all concrete reality [including experience] can in principle be fully captured in the terms of physics,” must deny or explain away the existence of conscious experience.\textsuperscript{440} For experience inevitably resists, Strawson thinks, this radical reduction. Even the most resolute physicalists think

\textsuperscript{437} He defends a version of the argument in Strawson, \textit{Mental Reality}, expands upon it in “Real Materialist Monism,” and does extended battle with a variety of commentators in Strawson, \textit{Consciousness and Its Place in Nature}, the book on which much of my discussion draws.

\textsuperscript{438} Strawson, \textit{Consciousness and its Place in Nature}, 3

\textsuperscript{439} Strawson, \textit{Consciousness and its Place in Nature}, 11

\textsuperscript{440} Strawson, \textit{Consciousness and its Place in Nature}, 4
that there is a *conceptual* gap between the realm described by physics and that of experience! In persistently talking as if there were a radical gulf between two sorts of things, the mental and the physical, which their explicit position commits them to unifying, physicalists reproduce the dualism they reject. They end up, tacitly though only rarely explicitly, committed to denying the reality of experience, whose existence—and here Strawson follows Russell exactly—is the thing each and every human being knows clearest of anything.

This inference—that any non-panpsychist brand of physicalism is tacitly if not explicitly committed to eliminativism about conscious experience—is obviously controversial. It plays no role in my argument here. Rather, I want to focus on Strawson’s discussion of the most common proposal for evading the incoherent conclusion of the argument thus far.

1. *Real Physicalism (RP)*: Every real, concrete phenomenon [including experience] in the universe is physical.

2. *NE*: Physical stuff is, in itself, in its fundamental nature, something wholly and utterly non-experiential.

3. So, every real, concrete phenomenon [including experience] is, in itself and fundamentally, wholly and utterly non-experiential.

Strawson states the argument as a *reductio*. To avoid the absurd conclusion, that experiential phenomena might be essentially and fundamentally non-experiential, Strawson suggests that there are two options. First, one can embrace panpsychism, the denial of NE: physical stuff is not, in its fundamental nature, wholly and utterly non-
experiential. Second, physicalists can defend some notion of emergence, according to which the experiential emerges from something fundamentally non-experiential.

Emergence is an admittedly fuzzy concept, but one that, Strawson thinks, his opponents require to state their views. The basic thought is the conscious experiences of beings like us emerge from fundamentally non-experiential phenomena combined in certain ways and joined by specific causal relations. Strawson’s favorite example of an intelligible—though he explicitly claims that this notion is metaphysical, rather than epistemological—emergent phenomenon is liquidity. No single H2O molecule is liquid, but many of them, suitably arranged at the appropriate temperatures, are. Sufficient knowledge of the relevant scientific facts gives us insight into this phenomenon (and perhaps God-like knowledge of these facts makes the emergence metaphysically necessary). Liquidity emerges at the top but remains absent from the most fundamental level. Strawson’s physicalist opponents, he thinks, must assume that consciousness is, or could be, like this.

Strawson thinks that the physical property of liquidity—note, as opposed to watery experiential properties, for example—*wholly depends* on what he calls other P-properties: size, shape, mass, charge, etc.441 Any metaphysically possible emergent state of affairs will wholly depend, Strawson thinks, on its underlying basis.442 The physicalist who rejects panpsychism must defend the total metaphysical dependence of experiential

441 Strawson, *Consciousness and its Place in Nature*, 13

442 Again, it’s worth noting the compositional feature of Strawson’s view, since Isolationism rejects it, but once again the difference need not be important. The requirement of total metaphysical dependence is more important, at this point, than the choice of whether this dependence is top-down or bottom-up.
phenomena on the underlying wholly and fundamentally non-experiential, which can be
described in the language of physics. This, Strawson thinks, they cannot do.

But, Strawson’s opponents can reply, actuality entails possibility! Fundamental
phenomena are wholly non-experiential, but experience is real, so experience must
emerge, in the total metaphysical dependence sense he intends, from the non-experiential.
The debate descends into quibbling about examples. Strawson gives two compelling
analogies for cases of brute, or impossible, emergence: the extended from the wholly
non-extended, and the spatial from the wholly non-spatial. As he rightly notes,
presentations of classical physics that employ point particles are not to be trusted—
infinities spring up all over and must be dealt with in successor versions of physical
theory.443 On this basis, he argues for the incoherence of the emergence of the extended
from the non-extended. Such cases, according to Strawson, are directly analogous to the
emergence of the experiential from the non-experiential, and are, by definition, a miracle
every time they occur. For “it is true by hypothesis that in brute emergence there is
absolutely nothing about X, the emerged-from, in virtue of which Y, the emerger,
emerges from it.”444 Only the widespread prevailing commitment to both RP and NE, real
physicalism about experience joined to the non-experientiality of the fundamental, makes
this notion seem possible, rather than strictly miraculous.

443 See Strawson, *Consciousness and its Place in Nature*, 15-18, as well as fn. 26-27 especially for
the physics examples. Though these examples are perhaps somewhat controversial, they should not be
unduly so. One explicit motivation for string theory, for example, by way of reconciling general relativity
and quantum mechanics, is to eliminate some of the infinities or absurdities that arise from the conflicts
between, for example, point-sized and wave-like phenomena.

444 Strawson, *Consciousness and its Place in Nature*, 18
But Strawson’s opponent just fails to see the analogy. Liquidity is the relevant comparison for the emergence of experience from the non-experiential, rather than extension from the non-extended. Strawson, the objection goes, illicitly relies on epistemic premises to sort cases of emergence that are possible, or intelligible, from those that are not. But, it continues, such epistemic judgments are not reliable. Conceivability and possibility are not as closely linked as Strawson’s argument seems to require.

Spinozistic Isolationism can fruitfully contribute here, but we must be careful to describe precisely how. One natural thought, common to both Strawson and to many of Spinoza’s interpreters, is to explicitly commit to, and employ, something like the Principle of Sufficient Reason. If there is no intelligible cause or reason for the generation of B—the extended or the experiential—from A—the non-extended or the non-experiential—then B cannot emerge from A. All emergence is intelligible, and this intelligibility is a guide, albeit a perhaps fallible one, to metaphysical possibility.

As I argued in chapter two, however, this move won’t work. There, I discussed Della Rocca’s attempt to use the PSR to refute skepticism, and I argued that he relies on an interpretation of the PSR according to which ‘intelligible’ entails ‘intelligible to us.’ The same equivocation haunts Strawson’s argument here. If the intelligibility required is only in-principle intelligibility, then arguments by analogy and disanalogy are ineffective. The fact that liquidity seems like a poorer model for the emergence of experience than the (impossible) model of extension from non-extension need not have any metaphysical, as opposed to epistemic, import. The difference here is between failing to see how the experiential intelligibly emerges from the non-experiential and seeing that the experiential fails to intelligibly emerge.
On the other hand, if Strawson’s argument requires ‘intelligibility to us,’ then he faces several problems. First, he must deal with all the counterexamples to every direction of the conceivable $\leftrightarrow$ possible inference. Relatedly, he must defend the principle’s plausibility in light of those counterexamples. This might include providing a non-hand-wavy explanation of the emergence of liquidity, his stock example of *intelligible* emergence, from only *fundamental* physical properties, something that, depending on physics’ progression, might actually be much harder than he believes. Only having two fleshed out stories of emergence side by side will allow us to evaluate which is intelligible and which is not, as well as which most closely approximates the story we tell about experience and the fundamentally non-experiential.\footnote{See Priest, “Sylvan’s Box,” for an extended argument that we wouldn’t be able to tell even then, though with a different upshot than what I want to press here.} And that already ignores the inevitable disagreements of interpretation.

But the *distinctively isolationist* aspect of Spinozistic Isolationism can provide what Strawson needs. For then all that Strawson actually requires for his argument to succeed is the falsity of physics-alism. This is the implausible view discussed earlier, and accepted by almost no one, that the nature of experience can be exhaustively characterized in the language of fundamental physics.\footnote{See Strawson, *Mental Reality*, 62-65 for an argument against physics-alism, though its denial is something many philosophers already accept without much argument.} After that point is granted, there are two fundamental conceptual vocabularies, one for the mental and one for the physical. By the chain of reasoning outlined in chapter one, the explanatory barrier between the attributes guarantees that these two realms are causally and conceptually isolated from each other. The argument of chapter three provides the EB with its
justification. To reject the EB is to cede the possibility that we can provide no compelling explanation for a whole host of ordinary events, of the sort discussed there and in chapter four. But, with Strawson as with Spinoza (1p10s): “From these propositions it is evident that although two attributes may be conceived to be really distinct (i.e. one may be conceived without the aid of the other), we still cannot infer from that that they constitute two beings, or two different substances.” The Spinozistic Isolationist’s further advantage, of course, is that she has the philosophical machinery to run her further argument for monism. This guarantees substantial unity once the radical causal and conceptual isolation between the mental and physical has been established.

It’s important to emphasize the difference between the PSR-centric and EB-centric approaches just contrasted. Employing a PSR-centric approach, which arguably undergirds or at least renders more reasonable Strawson’s own forceful insistence on the disanalogy between the cases of liquid and consciousness emergence, makes a very specific demand, on the epistemic interpretation. In considering the particular cases at hand, one must be able to intelligibly explain, or justify the metaphysical possibility of, liquidity’s emergence from the base set of P-properties (mass, charge, shape, etc.). Any such explanation must fail for the disanalogous case, whether that of the extended from the non-extended, or the experiential from the non-experiential. Specific explanations for putatively emergent phenomena, one metaphysically possible and the other not, must exist to be contrasted, and only after this occurs is one of them found wanting.

The Spinozistic Isolationist, however, focuses on the explanatory barrier between the attributes as a general thesis. The isolationist does not purport to have a ready explanation for the emergence of liquidity, extension, or experience. Nor does she think
such explanations can be ruled out merely because we cannot, either currently or in principle, fathom them—she endorses the metaphysical, rather than epistemic, interpretation of the PSR. Using the EB in the current discussion is all the more effective because the isolationist’s commitment to it, unlike Strawson’s commitment to the particular analogies and disanalogies he chooses, is motivated elsewhere and on more general grounds.

The isolationist argues that an almost transcendental presupposition for preserving the possibility of explanation in ordinary cases, *unrelated to the current debate about the origins of mentality*, requires the EB. But then, EB in hand, it turns out that additional arguments can supply the conclusion Strawson desires: realism about experience and a plausible definition of physicalism together entail panpsychism. The explanatory barrier’s generality and independence from any specific origin in the mind-body debate makes its use there all the more striking. Moreover, the combination of this Isolationist argument against emergence, along with the Isolationist contribution to solving the Cosmopsychist combination problem that follows below, offers a powerful package for the Isolationist to bring to the contemporary debate.

### 6.3.5 Grounding Top-Down to Solve the Combination Problem

Having shown how Spinoza’s Isolationist can offer a compelling defense of “no brute emergence” arguments of the sort Strawson favors, I now complete the Isolationist package by arguing that two exciting recent defenses of Priority Cosmopsychism, by Phillip Goff and Nagasawa and Wager, could benefit from the tools of Spinozistic Isolationism. So supplemented, they provide both a superior account of intra-phenomenal
grounding and hope for a compelling resolution to the combination problem. Jonathan Schaffer’s work on priority monism figures large in both defenses, and I think that all these authors are heavily indebted to Spinoza in various ways. Still, some of Spinoza’s more idiosyncratic views, including perhaps the explanatory barrier between the attributes, might make one reluctant to rely on Spinozistic theses explicitly. To the extent that the Isolationist argument I’ve given here is successful, then, Spinoza’s influence on these recent defenses of Priority Cosmopsychism seems worth airing openly.

Goff, Nagasawa, and Wager define Priority Cosmopsychism (henceforth: Cosmopsychism) as the view that all physical ultimates instantiate phenomenal properties, and that there is exactly one basic physical ultimate, the cosmos, with its one basic cosmic consciousness. Ordinary consciousness, of the sort we are familiar with in our own cases, is ontologically posterior to and derivative from the basic cosmic consciousness.447 They then identify four central benefits of Cosmopsychism relative to the micropsychism that we have seen defended by Chalmers and Strawson, and that constitutes the dominant form of the view. The first is largely implicit, but comes through in their reliance on Schaffer’s work: considerations from physics suggest that the universe is an entangled system, in some sense prior to, more fundamental than, and not strictly composed of, its subsystems.448 As this is a complicated, controversial argument in the philosophy of physics that I cannot engage without significant digression, however, I set it aside here. Second, Cosmopsychism is consistent with infinite decomposition or


448 Schaffer, “Monism: The Priority of the Whole,” 32
atom-less gunk; to the extent that gunk is possible, it threatens micropsychism just as much as microphysicalism. To the extent that gunk is possible, it threatens micropsychism just as much as microphysicalism. Third, the cosmopsychist offers a better account of the grounding relation, or, slightly differently, the micropsychist requires that an implausible version of the grounding relation obtain. Finally, Cosmopsychism supposedly solves the most serious form(s) of the combination problem. I focus on the final two problems here.

Goff gives the most detailed defense of the claim that Cosmopsychists offer a better account of the grounding relation between microexperience and macroexperience. His argument follows the broadly Hegelian line sketched in this chapter, whereby panpsychism offers an exciting middle road between two views with irremediable problems. He then notes that both Micropsychism and Cosmopsychism are live options for the panpsychist, each requiring a different conception of grounding, which Goff defines as “a non-causal explanatory relation that holds between facts,” where the derivative is nothing over and above the grounding entity or entities.

Each brand of panpsychist, whether micro- or cosmo-, cashes out this ‘nothing over and above’ talk differently. For instance, Goff thinks the two most plausible ways of doing so are grounding in terms of truthmaking (‘bottom-up’), or grounding in terms of subsumption (‘top-down’). A quick example for each should serve to draw the contrast. With grounding as truthmaking, we can say that truths about tables are grounded in truths about particles arranged in certain ways, table-wise, and the particles ground, and are the

449 See Schaffer, “Monism: The Priority of the Whole,” 32, as well as Nagasawa and Wager, “Panpsychism and Priority Cosmopsychism,” 113. I do not love this argument, since it relies on a claim—that metaphysical theses are necessary if true—that seems difficult to defend (i.e. it might be true, but justifying it seems precarious). For only with that claim in place can the mere metaphysical possibility of atomless gunk falsify micropsychism or microphysicalism.

450 Goff, “Cosmopsychism and Micropsychism,” 3, online
truthmakers for, all facts about tables. With grounding as subsumption, X grounds by subsumption Y iff Y is a partial aspect of X, as, for instance, a substance grounds its properties or the whole of space grounds particular regions of it.\textsuperscript{451}

The final move in Goff’s argument is to argue that grounding by truthmaking leads to problems for the micropsychist. The truthmaker grounding theorist’s basic dilemma is the following. Consider the relationship between the grounding, or fundamental entities and the grounded, or derivative, ones. On the truthmaking view, truths about derivative entities depend wholly on truths about fundamental entities. Do the derivative entities exist in their own right, or not? If they do not, the micropsychist is in trouble, since the experience we have direct knowledge of is of the macro-variety, which is, according to the micropsychist, grounded in the micro-experiential and thus derivative and hence nonexistent. So suppose instead that the derivative entities exist. Still, the proponent of truthmaker grounding will want to suggest that the grounded entities are metaphysically more fundamental, existentially privileged.

But this sits poorly when we again turn our attention to micropsychism. As Goff puts it: “When I ask what it’s like to be Bill I’m not interested in anything more fundamental than the o-subject [ordinary subject] that is Bill’s conscious mind and its states.”\textsuperscript{452} The panpsychist gets going in the first place because she cannot accept dualism but also cannot see how microphysical entities and properties are supposed to compose or generate her conscious macroexperiences. Being told that microexperiences, rather than

\textsuperscript{451} Goff, “Cosmopsychism and Micropsychism,” 7-8; the subsuming view of grounding for space is recognizably Kantian, and the substance-property grounding relation seems to accord well with the dominant assumptions of early modern philosophy, as in Leibniz and Spinoza, for instance.

\textsuperscript{452} Goff, “Cosmopsychism and Micropsychism,” 10

368
the ordinary experiences of subjects like you and me, are metaphysically fundamental seems to undermine the original motivation for panpsychism. For we have lost the intelligible connection between the truths we were interested in explaining in the first place—truths concerning ordinary experiences—and their more fundamental grounds.453

By comparison, the cosmopsychist need not relegate our ordinary conscious experiences to a second-class existence outside of fundamental reality. For the cosmopsychist can employ, rather than grounding by truthmaking, grounding by subsumption, according to which the basic cosmic consciousness has individual ordinary subjects, like you and me, as partial aspects. As Goff puts it: “Similarly the cosmopsychist claim that my conscious mind (the substance rather than the property) is an aspect of the cosmic mind is consistent with the claim that my conscious mind is part of Reality as it is in and of itself. There is no inconsistency in holding that both a whole and its aspects are privileged.”454

On my view, there are two huge problems with Goff’s reasoning as a standalone argument for Cosmopsychism as opposed to micropsychism. First, the literature on grounding is vast and widely contested, but Goff’s argument appears to rely on the truthmaking and subsumption views as constituting an exclusive and exhaustive disjunction.455 Second and more fundamentally, even supposing that the disjunction is exclusive and exhaustive, it’s not clear what progress cosmopsychism makes if its only

453 Goff, “Cosmopsychism and Micropsychism,” 11
454 Goff, “Cosmopsychism and Micropsychism,” 13
455 See Bliss and Trogdon “Metaphysical Grounding,” for a good survey of some of the recent literature on the topic.
motivations concern getting the better of the grounding debate. For micropsychism was rejected because it made the truth conditions for ordinary conscious experiences inaccessible, dependent on micro-experiences that were metaphysically privileged. But cosmopsychism, while denying that the individual’s conscious mind is somehow absent from privileged metaphysical reality—since it is a partial aspect of it—still makes individual conscious experiences derivative. It still leaves the puzzle of explaining how truths about my consciousness might be subsumed into truths about the larger consciousness of which my mind is a part. One might still think that we have lost the intelligible connection between the fundamental cosmic consciousness and the ordinary conscious experiences we were interested in explaining in the first place. These questions might well be answerable, but the symmetry in explanatory deficiencies here makes these grounding considerations at best only a partial support for the cosmopsychist over the micropsychist.

So, of the four advertised benefits of Cosmopsychism, its potential for solving some or all aspects of the combination problem seems the most important, since the combination problem in some form is often taken as fatal to panpsychism. Why is the combination problem regarded so seriously? I highlighted the key point in James’ initial formulation, given above. The dialectic goes like this. Panpsychism evades the causal and conceivability arguments that threaten dualism and physicalism, respectively. It responds to the problem of the brute emergence of experience or mentality by positing experience or mentality at the fundamental level. However, if it turns out that we can make no more sense of, for example, a macro-subject’s being composed of micro-subjects, than we can of the microphysical giving rise to the experiential, then the advantages seem minimal.
For the physicalist can now retort that, given the unintelligibility in both cases, physicalism accords much better with our other knowledge and/or ordinary intuitions. Specifically, it seems to better respect our sense that there’s nothing that it’s like to be a quark or an electron, and that it is crazy to assume otherwise. The micropsychist tolerates this violation of our intuitions because she thinks it offers enormous benefits—the avoidance of brute emergence—that her position now appears to avoid in name only. The combination problem, according to those who wield it as part of an argument against panpsychists, reproduces the difficulties panpsychism was introduced to solve. And, of course, if the Cosmopsychist allow requires some objectionable emergence of the ordinary consciousness from the cosmic one, she will be vulnerable to the same charge. In which case we are perhaps best left embracing one of the a posteriori forms of physicalism we initially hoped to leave behind.

Therefore, if the Cosmopsychist can blunt the force of this objection, she will have made a serious advance on panpsychism’s behalf. Nagasawa and Wager suggest that Cosmopsychism can do just this. However, as becomes clear in their defense of Cosmopsychism’s solution to the combination problem, they are actually worried about two fairly different aspects of the problem, the grain problem and the boundary or derivation problem. Let’s consider each in turn.

According to Chalmers, who relies on earlier work by Lockwood, Maxwell, and Sellars, the grain problem is to explain “how microexperiences come together to yield homogeneous macroexperiences, such as a homogeneous experience of red, instead of an
enormous jagged array of distinct qualities." More carefully, Chalmers characterizes the grain problem for panpsychism, though it is clear that a similar problem appears to afflict the microphysicalist. It is this: how do microphysical particles come together to yield homogeneous macroexperiences, on the microphysicalist’s view? This parallel illustrates the general worry that micropsychism is only a very tenuous advance on the dominant physicalist theory.

Nagasawa and Wager argue by analogy that while this problem seems serious for micro-versions of both physicalism and panpsychism, it poses no challenge for the Cosmopsychist. They say:

Suppose, *per impossibile*, there is an absolutely perfectly smooth painting, which is analogous to a smooth, homogeneous phenomenal experience. Such a painting cannot be an aggregate of small dots, which are analogous to phenomenal or protophenomenal properties of physical ultimates, but it can be a segment of a larger painting that is equally smooth and homogeneous, which is analogous to the cosmic consciousness.457

As analogies go, this seems pretty conclusive, at least if Nagasawa and Wager are careful with their terminology. If the perfectly smooth painting under discussion manages to be a *segment of* larger painting without being a *detachable part of* it, then their disanalogy seems to hold. The Cosmopsychist has resources that the micropsychist and microphysicalist do not. For suppose one has a perfectly homogeneous phenomenal experience of redness. One can abstract away a segment of this experience, perhaps by referring to the “bottom half” of one’s visual field, but one does not thereby commit

---


457 Nagasawa and Wager, “Panpsychism and Priority Cosmopsychism,” 121

372
oneself to arguing that the unified phenomenal experience is composed of top and bottom halves of the visual field. If one were committed in this way, then jaggedness would plausibly be introduced at the edges, and the same problem would recur for Nagasawa and Wager’s view as for the micro-versions they criticize. But it seems like the process of abstraction need not be equivalent to a process of decomposition into proper parts.

Of course, the right thing to say about the grain problem might simply be that phenomenal experiences actually aren’t as smooth or homogeneous as they are sometimes presented as being. But in that case the problem dissipates for everyone, microphysicalists and micropsychists and cosmopsychists alike, since in all cases either composition or decomposition can introduce slight heterogeneity without any issues. So the upshot seems to be either that the grain problem poses trouble for none of the relevant competitors, or that the Cosmopsychist has an advantage in addressing it.458

The boundary or derivation problem is both more interesting and potentially more serious. It arises against the backdrop of the subject-summing aspect of the combination problem, which we saw stated forcefully by James in the large quote with which we began. The existence of 100 micro-subjects does not necessitate the existence of a 101st macrosubject that is identical to one of the ordinary, organic human consciousnesses whose experiences we are interested in grounding. This point can be pressed in numerous ways. Here are just two, both due to Phillip Goff.459 First, he argues, the panpsychist’s basic argument, whether given by Strawson or Chalmers, relies to a certain extent on a

458 Stubenberg, “Russell, Russellian Monism, and Panpsychism,” 84-87 notes that a view that returns to some of Russell’s traditional motivations also seems like a promising candidate to evade this problem, and I think he makes a persuasive case for it.

459 What follows summarizes two arguments given in Goff, “Experiences Don’t Sum,” 54-59.
denial of the transparency of the physical. That is, the Russellian claim that we know nothing about the intrinsic nature of the physical is what allows the Russellian monist to make the identity or quasi-identity claims that she does. It allows her to say that the intrinsic nature of the brain is phenomenal. The mind is the brain, with the first term referring to its intrinsic aspects and the latter to its extrinsic relations with other things.

However, it also relies to some extent on the transparency of the mental, since if conscious experience were not transparent, then the panpsychist would have given no strong reason for rejecting a posteriori physicalism. This latter view just states that the conceptual gap between mental and physical does not threaten their empirical identity, and to reject it requires giving this conceptual gap ontological significance. But then the micropsychist has a problem, since it turns out that her view requires her to say that ordinary macroexperiences, of the sort we are all familiar with, are in actuality composed of billions of microexperiences, each with its own microsubject. But our ordinary experiences do not present themselves in this way, not even a little, and to that extent micropsychism seems to both rely on, and tacitly reject, the transparency of the mental. This doublespeak undermines the original motivations for the view.

Goff’s second argument focuses on a specific macroexperience, like feeling a severe pain. On the micropsychist’s view, this SEVERE PAIN is composed of billions of LITTLE PAINs arranged in the right ways. The problem, on Goff’s view, is that experiences are how they are perceived as being, and the micropsychist cannot make sense of this fact, since LITTLE PAINs feel like almost nothing at all, while SEVERE PAIN has a fundamentally different quality. But this novel quality is objectionably brute in precisely the way that the panpsychist criticizes physicalist explanation as being.
Goff takes both these arguments as indicators of a more fundamental conceptual truth, which is that ‘experiences, and subjects, don’t sum.’ This subject-and-experience-summing objection is often regarded as the most serious aspect of the combination problem. And initially at least, it may seem like Nagasawa and Wager’s, or Goff’s own, versions of Cosmopsychism have a precisely analogous problem. For a natural thought is that the Cosmopsychist’s view consists in arguing that the ordinary conscious experiences we are familiar with, the ordinary conscious subjects we are, compose a cosmic conscious experience of a cosmic subject. And explaining this seems not a bit easier than explaining how an ordinary conscious experience could be composed of the experiences of microscopic subjects and their experiences.

This, however, as Nagasawa and Wager note, is not the Cosmopsychist’s view. Rather, the Cosmopsychist argues that the cosmic conscious experience and the cosmic subject are ontologically prior to, and also not strictly composed of, the ordinary conscious experiences and subjects with which we are most familiar. In Schaffer’s terms, while aggregates depend on their parts, integrated wholes do not, and indeed mereological terms probably do not characterize the relationship between integrated wholes and their aspects correctly in many cases.\(^{460}\) So the Cosmopsychist straightforwardly averts the challenge from supporters of the claim that ‘experiences and subjects don’t sum’ is a conceptual truth.

Still, the Cosmopsychist does need to explain in what sense ordinary consciousness experience, and ordinary conscious subjects, might be derived from

\(^{460}\) Schaffer, “Monism: The Priority of the Whole,” especially 47; his discussion of entangled systems makes this particularly clear.
Nagasawa and Wager respond in two ways. Their first argument is again by analogy, and is similar to the preceding example. They say:

Consider, for example, a visual experience. A visual experience can be considered to be a unity which may be segmented into distinguishable color experiences (e.g., experiences corresponding to red and green hues) or experiences of separable regions in space (e.g., experiences corresponding to the right-hand side and the left-hand side of the visual field). Yet the whole visual field is considered to be a unity that is more fundamental than the segments.\textsuperscript{461}

This analogy works much better for experiences than for subjects. It makes good sense to suppose that aspects of a single phenomenal experience might be separable in analysis into various segments that nonetheless do not strictly compose the unified whole, but instead depend upon it as partial aspects, rather than proper parts. But the analogy starts to break down when we recognize that the natural development of it in this case is to stipulate further that each of the segments feels him or herself to be an isolated, wholly separable and independent subject. For even if the subjects are wrong and if they do not exist independently as they suspect, how this might be the case is left wholly unexplained by further consideration of Nagasawa and Wager’s analogy. And it is arguably this aspect of the problem that is most pressing for the Cosmopsychist to answer. Deriving ordinary subjects from cosmic—Spinoza would say, divine—subjects is the real trick.

Nagasawa and Wager’s second strategy is to appeal to a set of arguments Schaffer gives in support of priority monism, and to argue that they can be adapted to the Cosmopsychist’s goals. So, they think, the priority monist can explain how the cosmos as a unified whole manages to be locally heterogeneous in the way our world seems to be.

\textsuperscript{461} Nagasawa and Wager, “Panpsychism and Priority Cosmopsychism,” 121
She can argue that the cosmos is locally heterogeneous because it either 1) instantiates the determinable property of being heterogeneous, or 2) because it has a set of regionalized properties, or because 3) it instantiates various properties adverbially, being, for example, here-sunny and there-rainy.\(^{462}\) Now, even phrased in this highly general way, I don’t think that this second strategy is at all promising. In fact, it seems substantially less interesting than even the first argument by analogy, since each response merely restates the problem as a proposed solution. So, for example, if the objector wonders how the Cosmopsychist can say that medium-sized ordinary subjects are derived from a single cosmic-subject, then the Cosmopsychist responds either that the cosmic subject has the determinable property of being multi-subjective. Or, she can say that it has the regionalized properties of being Michael-like here, Kant-like there, and Schaffer-like in a third region, or, yet more implausibly, that the cosmos instantiates Michael-consciousness-here and Schaffer-consciousness-there. Even bracketing how plausible (or mostly implausible) the views of properties sketched here seem to be, we in any case seem no closer to answering the question we started with. Given the (apparently widespread) experience of perceiving the world from a point of view distinct from that of other subjects, the task of showing how this point of view derives from a wider cosmic one seems no closer to completion.

Moreover, it’s worth recalling the way that Goff characterizes the debate between the microphychist and the cosmopsychist on the question of grounding. There, in considering truthmaker grounding, Goff argued that the truthmaker theorist faces a dilemma concerning the status of the derivative entities. For it seems like she must say

\(^{462}\) Nagasawa and Wager, “Panpsychism and Priority Cosmopsychism,” 122-123
either that derivative entities do not exist, or that they exist but are less fundamental. Goff suggested that the grounding by subsumption theorist need not say that derivative entities do not exist, nor need he say that they are not part of fundamental reality. However, I argued that one might think there is still a perfectly good sense in which the derivative entities are less fundamental, and that this requires explanation. Or, slightly differently, one might argue that even though one is not forced to say that the derivative entities are less fundamental or somehow less real, once one has top-down grounding by subsumption rather than bottom-up truthmaker grounding, one has good reason to do so. In debates over Schaffer’s priority monism, for instance, critics have taken both tacks. Horgan and Potrc have argued that existence monism—the view that there is only one object, rather than merely one basic object—should be preferred to priority monism, and Della Rocca has argued that any defensible monistic view is forced to say that the derivative entities do not fully exist.463

My point is not to endorse these objections, or to reject them for that matter. I simply want to emphasize just how under-motivated the Cosmopsychist project, as defended by Goff, Nagasawa and Wager, is. The two major challenges for the view are first to defuse the summation problem by arguing that ordinary conscious experiences and subjects are not parts of, in a straightforward compositional sense, of the cosmic experience or subject. Second and equally importantly, they must provide some story to explain how ordinary subjects are derived from the cosmic one, and they must defend this story as an advance on the bottom-up view of explanation preferred by both the

463 See, for example, Horgan and Potrc, “Existence Monism Trumps Priority Monism,” Della Rocca, “Rationalism, Idealism, Monism, and Beyond,” and his unpublished Whitehead Lectures at Harvard University.
micropsychist and microphysicist. Otherwise, the combination problem in one its
guises—I think the subject summing or derivation versions are worst for Micropsychists
and Cosmopsychists, respectively—threatens to dissolve the motivations prompting the
move to panpsychism in the first place. In that case, either a posteriori physicalism or
old-fashioned interactionist dualism looks like a better bet.

Can Spinozistic Isolationism succeed where contemporary defenses of
Cosmopsychism come up short? Ultimately, I think that it probably can, but defending a
specific Spinozistic response to the subject derivation problem in detail is not my primary
aim here, for two reasons. First, I’ve discussed the issues in this chapter without the
precision required for really evaluating the proposed solution. Now, I don’t think this is a
defect unique to the discussion here—as we have seen, many of the debates in this area
proceed by appeal to simple analogies or bare intuitions of just what sorts of combination
or derivation of subjects and/or experiences are possible. But shared imprecision flatters
no one.

The more fundamental reason I won’t defend a specific Spinozistic response to
the subject derivation problem, however, is that discussing this particular challenge for
Cosmopsychism is, on my view, impossible to pursue fruitfully in isolation from other,
more foundational metaphysical questions. Spinoza’s manner of tackling these questions
is particularly illustrative in this regard. To give just a flavor for what I mean, consider
first that in the *Ethics* Spinoza argues carefully that God necessarily exists (1p11), that
absolutely infinite substances cannot be divided (1p12-13), and that therefore corporeal
substance cannot be composed of parts on pain of absurdity (1p15s). When Spinoza later
argues that Extension is an attribute of God (2p2), he completes an extended argument
for the coherence and actuality of the claim that substantial wholes ontologically proceed their parts, which do not compose them. This extended argument rewards further scrutiny in a way that the analogies used by Nagasawa and Wager do not, because it comprises one section of Spinoza’s more comprehensive overall argument in the Ethics. Similarly, Spinoza is sensitive to, and tries to explain, how individual ordinary subjects can unite to act as a larger subject. He explores this topic at greatest length in the Physical Digression of Ethics Part II, where he discusses the case of composite bodies and defines what it means to be an individual. There, he says (2p13lemma3d):

> When a number of bodies, whether of the same or of different size, are so constrained by other bodies that they lie upon one another, or if they so move, whether with the same degree or different degrees of speed, that they communicate their motions to each other in a certain fixed manner, we shall say that those bodies are united with one another and that they all together compose one body or individual, which is distinguished from the others by this union of bodies.

But, of course, Spinoza’s Parallelism, which states that the order and connection of ideas is the same as the order and connection of things, requires that an analogous explanation hold true for thinking subjects in their relation to a larger thinking subject. Whether the specifics of Spinoza’s physics, or the corresponding claims in mechanistic psychology, are plausible is somewhat beside the point here.

The basic idea is clear enough, however, and can be seen in its Spinozistic context by considering some remarks of Gregg Rosenberg, who forcefully pushes the boundary problem formulation as the most serious challenge posed by James’ famous passage. He says:

> There are good combination rules for many different kinds of things, in which these things can combine to become another unitary thing of either the same kind or a different kind. This is not a conceptual problem. It’s a problem about
understanding nature deeply enough to know when things interact and combine, and when they don’t… The difficulty is really the odd empirical fact that bounded phenomenal fields exist surprisingly at a mid-level of the physical world, at a scale corresponding to physical activity in animal brains.\(^{464}\)

The above quote from the *Ethics* shows that Spinoza is interested in precisely this same question: understanding nature well enough to know when and how things interact and combine, or, from the broadest perspective, derive from God or Nature so as to act together to produce a single, particular effect. This is not to say that Spinoza’s explanation, either in the bodily or mental case, is completely satisfying. After all, he lacked many of the relevant scientific facts, and was well aware of this. But what this does suggest is that naïve readings of Spinoza’s view, according to which, for example, he commits both to micro-subjects (like electrons and quarks) and implausible macro-subjects (like countries and sports teams, or like the fusion of any two human consciousnesses), are too hasty. In fact, Spinoza would probably reject both micro-subjectivity, in the robust sense of performing recognizably mental actions, and macro-subjectivity above the animal level and below the divine. He might do so precisely because the bodies and minds involved do not interact closely enough to bring about unified effects. The key insight, however, is that whether such subjects can be identified, for Spinoza no less than Rosenberg, might well be an empirical question, answered after investigation into the ways that bodies combine to bring about certain effects, rather than an armchair speculation. Additionally, the plausibility of Spinoza’s own combination rules would require evaluation based on the ways that his position has been interpreted in the secondary literature, which itself requires extensive digression into those discussions.

\(^{464}\) Rosenberg, “Causality and the Combination Problem,” 226-227
The basic thought, however, is that Spinozistic Isolationism is grounded within a comprehensive metaphysical system, anchored in empirical investigation of how and why things combine, in a way that Nagasawa and Wager’s Cosmopsychism is not.

What is crucial is the difference in methodology. This point runs through the entire argument of this work, and bears repeating here. When faced with the derivation or boundary problem as an objection to Isolationism as a version of Cosmopsychism, the Spinozistic Isolationist has far greater resources at her disposal. For she does not require ad hoc analogies or a specific theory of properties (ala Schaffer) developed solely in order to respond to the objection at hand. Rather, as we saw numerous times this chapter, the Isolationist can support existing responses—some aspects of a broadly Hegelian argument, the argument against brute emergence, the Cosmopsychist solution to the combination problem—with a comprehensive system whose consequences, derived from more general principles outside the specific ongoing debate, offer responses to the particular objection at hand. This methodological difference, as I’ve stressed in various ways in each chapter, is Spinozistic Isolationism’s greatest advantage. I think that the arguments of chapters 4-6 suggest that, to the extent that the argument for the explanatory barrier is successful, Spinozistic Isolationism is the most promising of the oft-discussed contenders in the metaphysics of mind. It offers compelling objections to dualism and physicalism, and provides much more comprehensive motivation for the crucial premises in arguments typically given by contemporary Russellian monists. In the ideal case, it thus hopes to reap the benefits of each position without paying its costs.
6.4 Isolationism and Russell’s Own Views, and Two Final Objections

Before concluding, I want to briefly discuss Isolationism’s relationship to Russell’s own views, specifically as regards the question of whether Russell’s view, and thus Russellian monism, requires the notion of grounding for its formulation. I finish by examining two final ways that Spinozistic Isolationism is superior to standard contemporary forms of Russellian monism.

As Alter and Nagasawa tell the story, contemporary defenses of Russellian monism are both inspired by, and remain close to the original intentions expressed in, Russell’s own formulations of his views. For instance, in their article “What is Russellian Monism,” Alter and Nagasawa describe the formulation I reproduced above, with structuralism about physics, realism about the inscrutables (defined as the grounds for the physical structure), and proto/phenomenal foundationalism, as “a general formulation of the [Russellian] view: one that expresses its main components, which are common to all versions.” If this were true, the Isolationist aspiration to capitalize on the argumentative advantages of Russellian monism in any form would be seriously hampered, since the Isolationist would have already rejected something essential to the family of philosophical views. But Alter and Nagasawa are wrong—grounding is not essential.

Russell himself states his motivations for developing his views in the following, characteristically insightful way, one that makes clear how disconnected his views are from those of contemporary “Russellian monists,” such as Chalmers or Alter and Nagasawa. He says:

465 Alter and Nagasawa, “What is Russellian Monism,” 425
Science is at no moment quite right, but it is seldom quite wrong, and has, as a rule, a better chance of being right than the theories of the unscientific. It is, therefore, rational to accept it hypothetically. It is not always realized how exceedingly abstract is the information that theoretical physics has to give. It lays down certain fundamental equations which enable it to deal with the logical structure of events, while leaving it completely unknown what is the intrinsic character of the events that have the structure. Nothing whatever in theoretical physics enables us to say anything about the intrinsic character of events elsewhere. They may be just like the events that happen to us, or they may be totally different in strictly unimaginable ways. All that physics gives us is certain equations giving abstract properties of their changes. But as to what it is that changes, and what it changes from and to—as to this, physics is silent.466

This is a fascinating passage for numerous reasons. As expected, we get the foundational distinction between the abstract, logical structure of physics, and the intrinsic, non-relational natures of the objects that are structured. One of Russell’s key insights is that if we pay close attention to the concrete results of actual physics, we recognize our freedom to interpret or give content to the intrinsic aspect of the events that are structured by physical laws. For there is absolutely no conflict between almost any interpretation of the intrinsic aspect of physical phenomenon A and A’s playing a given role within a causal nexus. The two sorts of properties perform fundamentally different jobs, only one of which physics even attempts to describe.

But, as Stubenberg has rightly emphasized, nothing in this passage at all suggests any notion of the “intrinsic character of the events that have the structure” as somehow grounding what Russell calls “the logical structure of events.” This case can be made on philosophical and textual grounds. Purely philosophically, it’s clear that the intrinsic properties of a thing do not always ground its extrinsic properties. For instance, we might accept Pereboom’s distinctions:

466 Russell, “My Philosophical Development,” 51
Intrinsic properties are non-relational properties, and extrinsic properties are relational properties…P is a purely extrinsic property of X just in case P is an extrinsic property of X and P has no intrinsic aspects… P is an absolutely intrinsic property of X just in case P is an intrinsic property of X, and X’s having P does not reduce to parts of X having purely extrinsic properties…[and] P is a relatively intrinsic property just in case P is an intrinsic property of X and X’s having P reduces to parts of X having purely extrinsic properties.467

With these distinctions in mind, consider a solid gray colored patch. Pereboom gives the example of **being one [object] among many** as an example of a purely extrinsic property the gray patch might have. We might also consider purely Cambridge properties of the patch, such as **being thought of by me**. Clearly, these extrinsic properties are not grounded in any of the solid gray patch’s intrinsic properties. Rather, its Cambridge properties depend on me and my activities, rather than on anything about it intrinsically.

Textually, Stubenberg notes: “The most detailed expression of the grounding idea has been given in terms of dispositional and categorical properties…But the careful reader discovers that this line of thought does not appear in Russell’s writings.”468 In speaking of the question of whether the intrinsic/extrinsic or nonrelational/relational distinctions might signify Russell’s commitment to the grounding idea in another form, Stubenberg is equally pessimistic. He notes:

So it is not the case that Russell sees physics as a peculiar body of doctrine that speaks only of relations, without bothering to provide the necessary relata. He is moved by other concerns. What Russell brings to the situation is a set of philosophical beliefs and goals. Ontological and epistemological beliefs push him to prefer logical constructions to entities that have ‘very convenient properties,’ like space-time points and electrons. His views about self-knowledge and non-deductive inference push him to believe that our percepts, plus the events we can infer from them, are real and better known than the entities that can only be

467 Pereboom, *Consciousness and the Prospects of Physicalism*, 93-94

468 Stubenberg, “Russell, Russellian Monism, and Panpsychism,” 79
known through inference, such as electrons...[But] there is no necessity about any of this...The project of logically reconstructing the particles of physics might be carried out from a completely different starting point, having nothing to do with percepts or their intrinsic features. All that is required is a set of real entities that is better known than the inferred theoretical entities of physics.\textsuperscript{469}

Russell, on Stubenberg’s telling, undertakes the project of logically constructing more distantly-known entities from better-known ones. In his case, the percepts whose intrinsic character we know serve as the logical construction base set, but there is no thought that the constructed entities are grounded in these intrinsic percepts.

Clearly, Spinoza’s Isolationist has still different concerns, explicitly rejecting the grounding formulation of contemporary defenders of Russellian monism while not demonstrating explicit interest in Russell’s project. Plausibly, this is because theoretical physics had not advanced, by Spinoza’s time, to the point where Russell’s project could have been salient. Still, what’s important to see here is that if Alter and Nagasawa’s central distinction is drawn in terms of intrinsic and extrinsic properties, then Spinoza can help himself to it without accepting cross-attribute grounding and thereby violating Isolationism. And if Isolationism has the virtues that I’ve argued that it has, then the present work gives additional reason to favor different no-grounding formulations of the view, either by returning to what Russell himself wrote, as Stubenberg urges, or by developing Isolationism as a promising instance of some key Russellian intuitions.

One might think, however, that Spinoza’s commitment to infinitely many attributes actually bars any smooth integration of the view into the Russellian monist camp. The worry, simply, is that the Isolationist has far more properties than she can use!

\textsuperscript{469} Stubenberg, “Russell, Russellian Monism, and Panpsychism,” 81-82
Some central distinction between two sorts of properties, one characterized in relational or extrinsic terms, and the other characterized as non-relational or intrinsic, motivates the distinctively Russellian monist. But the Spinozistic Isolationist, one might think, has countless such distinctions to make. What’s special about the distinction between, in Spinoza’s terms, extended and thinking properties, such that it maps smoothly onto the one that so impresses Russell and contemporary defenders of similar views?

Russell thinks that the intrinsic character of our own percepts might differ radically from the nature of events that happen elsewhere, as the quote above makes clear. But Spinoza’s Parallelism and Mind-Body identity theory do not allow him to argue that physical events might have radically different intrinsic characters (i.e. that some physical events might have, while others lack, a corresponding mode of Thought). Pace Russell, who seems to consider this very possibility in arguing that all events other than those going on here, in me, might have radically different natures, it does not even seem open to Spinoza to argue that all electrons, for instance, have thinking intrinsic natures, while all quarks have intrinsic natures of the third or fourth, etc. unknown attribute.

A possible, elegant, Spinozistic solution to this objection—that his substance of infinite attributes cannot be smoothly integrated into the Russellian framework in which a distinction between two sorts of properties is central—might respect Russell’s insight while at the same time subsuming it into a larger picture. In a sense, being a Russellian panpsychist might just be a parochial, human way of being an Isolationist. On the view sketched here, Russell’s central binary, between the relational properties of physics and the intrinsic ones of consciousness, is central because Thought and Extension are the
only two attributes cognitively accessible to human beings. Melamed’s solution to Bennett’s challenge to explain human ignorance of all but these two attributes, which I discussed in chapter one, consists in arguing that the explanatory barrier operates not only between attributes, but also within Thought itself. The same thinking mode has infinitely many aspects, each of which might serve as the intrinsic, non-relational base for some extended or third, fourth, fifth, etc.-attribute property.470

On this picture, Spinoza would require the additional claim that only Thinking properties are truly intrinsic. Perhaps this could be secured in two steps, first by arguing that all Thinking properties are of the same, perhaps representational, kind (pace those who sharply differentiate qualia from representational properties) and then by arguing that only such representational properties can be truly intrinsic.471 With these two steps secure, one can imagine a third-attribute counterpart to the actual Russell, master of the science of attribute three rather than of physics, drawing the same central distinction between Thought and Attribute Three, and being similarly frustrated when the Isolationist suggests her reply. The central distinction in that scenario is as it is because Thought and Attribute Three are all the attributes that are known there. Only from the Spinozistic Isolationist’s perspective are all the central distinctions drawn by all the Russells of

470 We must be careful here, of course, to avoid idealism by illicitly importing a bottom-up approach to grounding here. This is one reason I discuss an alternative to the position outlined here later in the chapter. For now, though, it’s enough to echo Melamed, “Spinoza’s Metaphysics of Thought,” 678-79, at the end of the paper where he puts forth his brilliant solution to the problem of human ignorance of other attributes: “Spinoza’s view of Thought as being more complex and powerful than any other attribute presents us with a surprising position on the mind-body issue. Spinoza is neither a materialist, nor a reductive idealist, and even the ‘double aspect theorist’ label would not adequately fit him. His view is one that grants clear preeminence to Thought without adopting reductive idealism. Here, perhaps, lies Spinoza’s real philosophical genius.”

471 Della Rocca, *Spinoza*, makes an argument supporting the first step here, and it might be thought that Leibniz’s thought provides resources for defending the second step. But saying exactly how this might go is not a project I’ll pursue more here.
various attributes other than Thought seen as what they are: mere consequences of the
claim, as *Ethics* 2p7 has it, that the order and connection of ideas is the same as the order
and connection of things.

Isolationists, I think, can thus happily wear the label of Russellian panpsychists,
in a sense that remains faithful to Russell’s original intentions, as Isolationists, like
Russell, do not formulate their view in grounding terms. Isolationists can propose an
argument structurally similar to Chalmers’ Hegelian argument while better motivating
key aspects of the final panpsychist view, such as commitment to uniform, ubiquitous
phenomenal properties, rejection of brute emergence and acceptance of top-down
grounding, and the Cosmopsychist solution to the combination problem. Finally, as I
show now, Isolationism can convincingly respond to two powerful objections to
Russellian monist views as traditionally formulated.

Daniel Stoljar and Alyssa Ney propose a set of important, complementary
objections to Chalmers-style Russellian monism that Isolationism can smoothly combat.
Stoljar, for instance, begins by observing that Russellian monists typically defend their
position by drawing some central distinction, as was discussed earlier. For example, they
might appeal to intrinsic vs. extrinsic, or structural-and-dynamical vs. non-structural-and-
dynamical properties. But, as Stoljar argues, “distinctions in this class…have proved
difficult to control.”472 This becomes problematic because every argument that seeks to
distinguish physical truths from truths about consciousness, in order to establish
Russellian monism, takes a similar form. Stoljar considers three such arguments, the
general form of which is:

472 Stoljar, “Russellian Monism or Nagelian Monism,” 325
389
1. Every physical truth is a truth of a certain kind, i.e. one that concerns relations/one that concerns extrinsic or only comparatively intrinsic properties/is expressible with sentences of a certain sort.

2. For every truth T of that kind, if T a priori entails a truth T*, then T* is of that kind too.

3. No truth/not all truths of consciousness are truths of that kind.

The problem with employing unstable distinctions in the first premise is that the Russellian requires precisifications of each distinction such that P1 and P3 can both be true together. This, Stoljar argues, is quite hard to do. For example, if one draws the central distinction in terms of relational and non-relational truths, Stoljar argues, then “there is no reason to suppose that physics cannot tell us about properties that are not relations…[and] it is false that from truths about relations only truths about relations follow…[and] furthermore some truths about consciousness are themselves relational.”

The proponent of the general argument given above might then follow Pereboom in claiming that every physical truth is either purely extrinsic or relatively intrinsic, while some truths about consciousness concern only absolutely intrinsic properties. In this case, Stoljar’s response is more complicated, but the gist of his strategy is to challenge either the claim that any truths of consciousness concern absolutely intrinsic properties, or the claim that no physical truths do. Regardless, Stoljar thinks, the Russellian monist will have a difficult time making her case.

From a complementary perspective, Alyssa Ney argues that we might accept the central distinction between intrinsic and extrinsic properties, suitably qualified perhaps,

---

473 Stoljar, “Russellian Monism or Nagelian Monism,” 331
while simply arguing that physics provides intrinsic truths. In Stoljar’s framework, this vitiates the truth of P1. Ney starts by accepting David Lewis’s distinction between intrinsic properties, which things have in virtue of the way they themselves are, from extrinsic properties, have in virtue of their relations or lack thereof with other things.\footnote{Lewis, as quoted in Ney, “Are There Fundamental Intrinsic Properties?” 221-222} 

She then makes a forceful preliminary point, before proceeding to her central argument. The preliminary point is simply that just because physicists describe properties with reference to what they do, or with reference to their causal or other effects on the properties of other things, this does not entail that they are not intrinsic. For example, suppose electrons are composed of two sub-particles, each of which, in itself, has a perfect rhombus shape, which is, unfortunately, not observable in principle with our best equipment or any possible extension of them. Still, having the property of \textit{being perfectly rhomboidal} might characterize these sub-electronic particles as they are in themselves, thus blocking the Russellian monist’s P1.

However, Ney admits that there appears to be a deeper intuition about the sorts of properties physics, in particular quantum mechanics, describes that buttresses the claim that no absolutely intrinsic properties can result from this characterization. I will not discuss all the details here, but the basic outline of the argument is as follows:

First, one establishes that there exist certain entangled states, and that such states characterize systems \textit{merely in terms of extrinsic properties}. Second, one argues that the initial quantum state of all the particles of the universe is in an entangled state. Finally, one argues that systems never evolve out of entangled states—once entangled, always entangled.\footnote{Ney, “Are There Fundamental Intrinsic Properties?” 227}
Ney carefully describes several ways one might motivate, and block, each step in the argument, depending on one’s interpretation of quantum mechanics. For our purposes, these details are not important. Her crucial claim is simply that the first premise is false, since while quantum mechanics never does attribute intrinsic properties to particles in entangled states, “on any mainstream, realist approach, quantum mechanics will attribute intrinsic properties to something else: the wavefunction.”\(^{476}\) If once learns to appreciate, as any realist about QM must, according to Ney, that the wavefunction [of the whole Universe] is a physical object in itself, and that the properties of particles and other macro-objects are constituted by or grounded in features of the wavefunction, then one can deny P1. Physics gives us insight into the intrinsic properties of things—the character of the Universe’s wavefunction—and thus does not offer categorically different information from what we get in conscious experience, on at least this way of drawing the central distinction.

Combined, Stoljar and Ney press the following dilemma for the Russellian monist as she is traditionally motivated. In order to argue for the view, one needs either an unstable central distinction that cannot be made more precise without undermining one’s justification for jointly accepting P1 and P3, or one can use the intrinsic/extrinsic distinction but P1 is therefore false. Either way, the position offers no help against the physicalist. This last claim, of course, seems a bit too strong. After all, in this chapter we entered Russellian monism both through Chalmers’ Hegelian argument, which relies on the conceivability argument and is at least somewhat independent of the issues just

\(^{476}\) Ney, “Are There Fundamental Intrinsic Properties?” 231
discussed, as well as through Russell’s own views, also justified on different, more epistemic grounds.

Even with this caveat, however, we can see that Ney’s counter-argument to the micropsychist Russellian monist was not to contest things at the microscopic level. Rather, she argued that though intrinsic properties are absent from particle physics, they reappear in cosmology, at the level of the wavefunction of the whole Universe. This cosmic level gives us insight into the intrinsic properties of physical things and blocks, absent some other way of drawing the central distinction that does not replicate these issues, one of the Russellian monist’s best arguments against the physicalist.

As should be clear, however, this move offers no threat to the Spinozistic Isolationist, for whom the move to the cosmic level poses no problems. As we have interpreted Spinozistic Isolationism, it’s compatible with the view to run the standard Russellian arguments for a micro-version of the intrinsic/extrinsic distinction, which Ney accepts, while denying the relevance of that distinction at the level of God or Nature. Since Spinozistic Isolationism is dependent on completely different sorts of argument for its ultimate defense, it does not require that the difference between the intrinsicality of conscious experiences and the extrinsicality of at least some classes of physical properties hold at every level. Indeed, Ney’s appeal to the wavefunction of the Universe evokes one of Spinoza’s most puzzling claims, in Ep64, that some examples of infinite modes include “motion and rest” and “the face of the whole universe.” If there is any more presciently metaphorical anticipation of the universal wavefunction in the history of science or philosophy than Spinoza gives here, I do not know of it.
Of course, if the Isolationist defends the picture sketched previously, on which only Thought contains intrinsic properties, she cannot then accept intrinsic physical properties at the cosmos level. So she can choose to defend the previous picture at the cost of finding some other way of responding to Ney and Stoljar here, probably by employing the very argument—that only Thinking properties could be intrinsic—that she used to motivate her position in the first place. Alternatively, she can draw the central distinction differently, so that one category of properties need not ground the other (preserving Isolationism) but so that she can accept Ney’s conclusion that physics gives information about intrinsic properties at least in the case of the Universe’s wavefunction (thereby avoiding the present objection). I see no in-principle obstacles to either tack, and I thus conclude that Spinozistic Isolationism escapes these two criticisms as well, further cementing its advantages relative to other contemporary views.

As I’ve already mentioned, however, fully fleshing out some Spinozistic solutions—to the derivation problem of finite subjects from God or Nature, for instance—requires more space than I can give here. Fortunately, however, much of the rest of the Spinoza secondary literature devotes itself to these and related tasks. What I hope to have done is to have made plausible a rejuvenated Spinozistic contributor to contemporary debates in metaphysics and philosophy of mind, grounded in Spinoza’s central, puzzling, and previously insufficiently defended claim that each attribute is causally and conceptually isolated from every other.
6.5 Conclusion

Scholars often characterize the history of early modern philosophy and the temporally corresponding scientific revolution as the overthrow of Aristotelianism in both philosophy and science. Descartes, writing to Mersenne, notes of his *Meditations*:

I may tell you, between ourselves, that these six meditations contain all the foundations of my physics. But please do not tell people, for that might make it harder for supporters of Aristotle to approve them. I hope that readers will gradually get used to my principles, and recognize their truth, before they notice that they destroy the principles of Aristotle.477

If the *Meditations* contain the principles of a mechanized Cartesian science of Extension wedded to an understanding of the mind as a separate, thinking substance that causally interacts with extended matter, Spinoza extends Cartesian mechanism to the mind itself.478 And though strict mechanism in science is no longer the dominant paradigm, and indeed has not been so for quite some time, mechanistic notions of causation and explanation are very much still with us. Mechanistic or efficient causal theories compete with Humean counterfactual or regularity theories that appear to deny the existence of genuine causation at all. They do so alongside a widely shared commitment to characterizing causation as occurring between events, rather than agents or substances. This has been implicit both in my presentation here, and in the arguments of most (though not all) philosophers I’ve thus far discussed.

I have made my arguments here within this broadly mechanistic causal and explanatory framework, shared in large part by contemporary dualists and physicalists

477 Descartes, Letter to Mersenne, MP 2

478 For a recent detailed treatment of this idea as a comprehensive reading of Spinoza’s philosophy, see Marshall, *The Spiritual Automaton*. 
alike. In the first chapter, I suggested that the explanatory barrier between the attributes provides central premises for many of Spinoza’s major and most distinctive philosophical theses. I then argued that no compelling Spinozistic response to skepticism was to be found elsewhere in his system, before arguing in chapter three that the introduction of the EB solves a skeptical or explanatory problem within all interactionist systems. Chapters four and five extended Spinozistic arguments, grounded now in the prior defense of the EB, against contemporary forms of dualism and physicalism, respectively. Finally, chapter six made the case for Spinozistic Isolationism as a form of or close rival to Russellian monism, specifically Cosmopsychism, and argued that Spinozistic Isolationism provides the most comprehensively motivated and well-defended form of the position.

But all this occurred while standing astride the vanquished, on the post-Cartesian telling, Aristotelian framework. And undoubtedly, the Aristotelian framework of that time got what it deserved, though it takes more care than it might seem to show just exactly how it was scientifically and philosophically deficient. For my purposes, however, it’s worth keeping in mind that my defense of Spinozistic Isolationism occurs in a very particular context, even if that context’s presuppositions are widely accepted in philosophy.

Contemporary Neo-Aristotelians, of course, reject both the mechanistic causal and explanatory framework presupposed by the early moderns and their descendants, and its concomitant replacement of causation by agents for causation between events. Moreover, they do not regard issues with qualia or conscious experience as the most difficult issues in the philosophy of mind. Rather, they think that properly understanding
intentional content and human rationality are the central questions.\textsuperscript{479} Invariably, they are quite sensitive to the stakes of these dual rejections of Aristotle. To give just one example, David Braine argues:

If it is once granted that ascriptions of ‘mental states’ must denote the state (or disposition of parts and relations of parts) of some physical or non-physical set up or apparatus, it will immediately be taken that, when they are historically explanatory of some physical movement, this is because the state which they denote (qua state of a set up or apparatus) is mechanically causative of what they explain—‘mechanically causative’ in the sense of being analogous to mechanical causes in their mode of action, external to and determinative of their effect...What we have been dealing with has been a theory of meaning that demands some degree of isomorphism with the physical structures which arise at the level of human acquaintance to be a condition of meaningfulness or possession of content in factual statements.\textsuperscript{480}

The final lines, concerning an isomorphism between mental states with meaningful content and the underlying physical structures, seems to allude to something like Spinoza’s Parallelism. The thought, sketched by Braine here and by others elsewhere, is something like the following: given this mechanistic account of causation and explanation, something like the mind-brain identity theory favored by many in the 50s and 60s is almost inevitable. What I have tried to suggest here is that Spinozistic Isolationism, rather than this modern identity theory, offers the best way of satisfying the criticisms of dualism, reductive and non-reductive physicalism, and panpsychism, respectively.

\textsuperscript{479} Ross, \textit{Thought and World}, gives just one of numerous examples I could cite here. Feser, \textit{Philosophy of Mind: A Beginner’s Guide}, is particularly clear about the stakes of this change in focus, on both the qualia/rationality and mechanistic/Aristotelian causation fronts. I tried to do some justice to this intuition by focusing both on the contemporary consciousness debates, in Ch. 6, and on issues with intentionality in a physicalist framework, in Ch. 5, but of course much more needs to be, and can be, said.

\textsuperscript{480} Braine, \textit{The Human Person}, 244, 247; similar arguments can be found too in Ross, \textit{Thought and World}. 397
This means, however, and given the context just described, at best only that if *Aristotelian forms of explanation were rightly overthrown on philosophical and scientific grounds*, then Spinozistic Isolationism is the most compelling contender in metaphysics, both of mind and more generally. This premise is widely shared by those philosophers discussed here, and thus does not threaten most of the arguments I have made in this work directly. Many days, this rejection seems broadly justified to me. Increasingly, however, I wonder whether challenges to the predominant mode of causation and explanation, alongside some of the ramifications of actually endorsing the ethics in the *Ethics* that results from Spinozistic Isolationism, might not justify a reconsideration of some, perhaps quite different form of the vanquished Aristotelianism within the mainstream, contemporary debate.\(^{481}\)

Even if that is the case, the current project interacts fruitfully with other exciting work in Spinoza studies, by Karolina Hubner and others, which seeks to explain causation in Spinoza in less purely mechanistic terms, along the lines of Aristotelian formal causation.\(^{482}\) An interesting extension of the current project would explore to what extent, if any, the present conclusions depend on or would be changed by radically reconceiving causation in Spinoza as formal, rather than efficient. I suspect that a comprehensive interpretation of all causation in Spinoza as formal causation would make

\(^{481}\) I have not discussed the ethical ramifications of Spinoza’s view at all here—that is a topic for another time. Several recent works, including Marshall, *The Spiritual Automaton*, Newlands, *Reconceiving Spinoza*, LeBuffe, *From Bondage to Freedom*, and others have, however. I do not agree with all of their conclusions, naturally. But I am closest to those interpreters who suggest that Spinoza does not have the resources to generate robust ethical normativity, which I think might be the gravest problem afflicting his philosophy.

\(^{482}\) For an excellent example of this sort of project, see Hubner, “On the Significance of Formal Causation for Spinoza’s Metaphysics.”
the extensive causal and explanatory opacity on which the present work depends easier, rather than more difficult, to justify. For it seems easier to argue that formal causes create referentially opaque contexts than to argue the same about efficient causes; the hegemony of the former would thus help motivate the ubiquity of opaque contexts in Spinoza’s system. If that’s the case, then the real showdown might be between views that capture key aspects of Spinoza’s own position—neo-Aristotelian formal causation, on one hand, or Leibniz’s Parallelism and commitment to the PSR, which seem to employ some formal causal Scholastic elements, on the other—rather than between endless epicycles of modern dualism and physicalism. Regardless, there is great benefit in studying Spinoza, the most consistent and most daring of the early moderns, who is guided by the Principle of Sufficient Reason in its most uncompromising form. For seeing how his arguments can be further strengthened by the justification for the EB given here can serve as a call to action to those who think that a remade Aristotelian synthesis or a Leibnizian idealism offers the best or only alternative to the various dualisms and physicalisms defended nowadays. All of which, I have argued, collapse into or are defeated by the Spinozistic Isolationism defended here.

483 The question of whether Spinoza’s PSR is the best, and not just the most uncompromising, form of the PSR is one that Pruss, The Principle of Sufficient Reason, takes up with great vigor, if not always with due fairness to Spinoza’s view. Still, Pruss employs a commitment to some fairly strong versions of the PSR in service of a classical theistic, rather than Spinozist, metaphysics. I do not love many of the arguments that Pruss makes. But for my money it is this debate—between classical theists committed to their PSR and Spinoza committed to his, with ethics and metaphysics as the primary battleground—that is the one worth watching.

484 Of course, I have by no means come close to discussing all the various physicalisms and dualisms defended nowadays. Among the most promising forms of dualism or arguments for it not discussed explicitly here, I will mention Sophie Gibb’s ‘double prevention’ model of emergentist dualism, which she argues, I think rightly, can respect the causal closure of the physical. On this, see Gibb, “Defending Dualism,” as well as “Mental Causation and Double Prevention.” For another compelling argument for dualism left undiscussed, see Unger, “The Mental Problems of the Many,” and Empty Ideas, especially Ch. 8. But don’t make the absence of explicit argument more than it is. Any alternative to the
Spinozistic Isolationism defended here should show why the explanatory or skeptical problem to which the EB responds is resolvable by other means, why the EB fails to solve it, or why it is not a problem after all, but in a way that amounts to more than just insisting that it isn’t. Without that background, these other promising forms of dualism will face 1) the stricture against mixing mental and physical forms of explanation, 2) the violations of the PSR inherent in their positions, or both. The same could be said for specific physicalist views not specifically discussed here too. In both cases, Spinoza can offer substantial in-principle reasons for doubting that these alternatives succeed, even without demonstrating this in detail. This situation can be contrasted with the dialectic vis a vis Aristotelianism or Leibnizian idealism, where most parties to the contemporary debate discussed here seem to share a similar confidence in its dismissal. I think this makes a Neo-Aristotelian or Leibnizian alternative, compatible with the givens of contemporary science, far more compelling than either physicalism or dualism as a rival to Spinozistic Isolationism.
WORKS CITED

Primary Texts

Spinoza


Descartes


Sextus Empiricus

Hume


Leibniz


Secondary Texts


https://home.cern/about/physics/dark-matter


