From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Philosophy

2018

Reduction and Images of Reality

Tabris, Nathaniel Ariel

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp014t64gq879

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

REDUCTION AND IMAGES OF REALITY

Nathaniel Ariel Tabris

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT OF

PHILOSOPHY

Advisors: Shamik Dasgupta and Mark Johnston

April 2018

© Copyright by Nathaniel Ariel Tabris, 2018.

All rights reserved.

Abstract

There's an important but nebulous debate about the nature of metaphysics between the metaphysical realist and the antirealist. This dissertation reframes the debate by arguing that both the realist and the antirealist have an important insight missed by the other. Chapter 1 argues that there are things we can talk about only by immodestly positing them as metaphysical explainers. This thesis expresses the core insight of realists such as David Lewis, Kit Fine, and Ted Sider. Chapter 2 defends a constraint on what the realist can say: you shouldn't posit a metaphysical explainer if the resulting explanation would rely on an epistemic brute difference—a difference which you're completely unable to articulate. This constraint explains why many of the explanations offered by realists are in fact spurious, and it expresses the core insight of antirealists such as Hilary Putnam, Nelson Goodman, and Richard Rorty. Chapter 3 explains what we get when we put these two insights together. The two insights can be harmonized only by positing a holistic form of non-reductive metaphysical explanation.

Acknowledgements

To my advisors, Shamik Dasgupta and Mark Johnston: thank you! Shamik and Mark have allowed me to pursue a project when they (and I) didn't know where it was headed, continuously pushed me to do better, and just generally put up with me (which isn't easy!).

Daniel Garber, Hans Halvorson, Benjamin Morison, and Gideon Rosen each exemplify distinct philosophical virtues, and I am very grateful for the attention and help that I've received from them. In particular, I want to thank Dan and Hans for their generosity to me at a couple of points when I was especially discouraged about philosophy and my own work.

When I came to Princeton, I didn't know that Adam Lerner and Eric Hubble would also be here. After so many conversations and shared drafts over the years, I have no idea what this dissertation would look like were it not for them.

And to everyone who helped me think through these ideas, who gave me comments on earlier versions of the material, and who constituted an excellent community for doing philosophy: thank you! I want to mention especially Emad Atiq, Daniel Berntson, Ellie Cohen, Robbie Hirsch, and Raffi Krut-Landau.

Finally, I am deeply grateful to so many others—more than I will attempt to name—who have had a less direct effect on the peculiar content of this work: thank you!

S.G.D.

To Rachel, Roger and Beatrice

Contents

Abstract iii

Acknowledgements iv

1 Reduction and Metaphysical Immodesty **1**

1.1 The background 2

1.2 What is reduction? 8

1.3 Why not reduction? 24

1.4 Immodesty 30

2 Immodest, yet Articulate, Metaphysics **31**

2.1 Why posit a new distinction? 33

2.2 Bald-faced primitivism 45

2.3 Against positing brute differences 46

2.4 The persistence of brute differences 53

2.5 Immodesty without brute differences 68

3 Holism and Metaphysical Explanation **75**

3.1 Preliminary clarifications 76

3.2 Dichotomies and brute differences 81

3.3 Two metaphysical pictures 91

Bibliography **99**

Chapter 1

Reduction and Metaphysical Immodesty

Immodest metaphysics, as I call it, presumes that there are facts we can discover only by doing metaphysics. J. J. Thomson discovered the electron by positing it to explain the behavior of cathode rays. Arthur Evans discovered the Minoan civilization by positing it to explain artifacts uncovered in Crete. And the aim of this paper is to argue that metaphysics is in a similar business of discovery:

> **Immodesty** There are things we can talk about only by positing them as metaphysical explainers.

When Thomson posited the electron, he was positing a new *entity*—thus expanding our ontology—in order to answer a *why?* question. We'll focus on a different type of posit: positing a new *notion*—thus expanding our ideology—in order to answer a *what is it?* question.[1] This is what David Lewis did when he posited naturalness to explain the difference between green and grue.[2] And it's what Kit Fine did when he posited essence to explain part of the difference between the fact that necessarily Socrates exists if his singleton exists and the fact that necessarily the singleton of Socrates exists if he does.[3]

[1] By "new", I mean a notion which isn't *definable* in terms of old notions. Adding a definable notion doesn't let you say or explain anything you couldn't already say or explain.

[2] See "New Work For a Theory of Universals".

[3] See "Essence and Modality", p. 5.

Another way to posit a new notion—thus expanding our ideology—is by positing a new *category* of entity in order to answer a *what is it?* question. This is what happens if you posit universals to give an account of predication or posit haecceities to give an account of individuation.

The goal of this paper is to argue for **Immodesty**. The basic structure of the argument is simple:

1. We have good reason to say that some reductive views are false.
2. We need to endorse **Immodesty** to explain what it is for a reductive view to be false.
3. So, we have good reason to endorse **Immodesty**.

Let's briefly consider an example of how an immodest posit could figure in the explanation of what it is for a reductive view to be false. Take a dispute about the nature of mental properties. Some people think that mental properties are reducible to physical properties; other people deny this. One way to explain this dispute is to posit a notion of *fundamentality* and interpret the dispute as a dispute about which properties are fundamental. The reductionist—thus interpreted—claims that only physical properties are fundamental, while the non-reductionist—thus interpreted—claims that mental properties are (also) fundamental.

Let me be clear: I am not going to argue that we should posit fundamentality in particular (and I don't think we should!). In subsequent chapters, I address the question of which notions we should posit. Here, my thesis is more general. I shall argue that there is explanatory work to be done by positing new notions.

But before I get into the argument, I want to start by explaining who and what I take myself to be arguing against.

1.1 The background

Let me say something about the relation between **Immodesty** and metaphysical realism, and then something about how reduction relates to the dispute about metaphysical realism.

1.1.1 Realism and Immodesty

"Metaphysical realism" can mean many different things. We shall focus on the view that there's a metaphysical distinction amongst true sentences like "there are electrons" and "there are tables and chairs" or like "all observed emeralds are green" and "all observed emeralds are grue".

The realist thinks there's a further question we can sensibly ask about these true sentences: which of these sentences are merely true, and which are also *really* true or *fundamentally* true or obtain *in reality* or *carve at the joints*? For instance, we might say that the fact about electrons obtains in reality, but the fact about tables and chairs does not. Or, we might say that the fact about green is more joint-carving than the fact about grue. In this paper, I shall use "metaphysical realism" to pick out the view there's such a distinction to be drawn.[4]

One of the core commitments of metaphysical realism is **Immodesty**. The realist posits a notion—fundamentality, reality, or some such—and claims that this notion can be used to state new facts that we wouldn't otherwise have been able to state.

And, many of the arguments against metaphysical realism are in fact arguments which target **Immodesty**. What I have in mind are arguments that the new notions posited by the metaphysical realist, because they are notions which cannot be explained in more familiar terms, are *unintelligible*. For instance, this is how Hilary Putnam explained his opposition to metaphysical realism:

> *Metaphysical* realism… is, or purports to be, be, a model of the relation of any correct theory to all or part of THE WORLD. I have come to the conclusion that this model is incoherent.[5]

What does it mean to say that metaphysical realism is *incoherent*? In a later paper, Putnam clarifies the objection:

> The main point was that metaphysical realism cannot even be intelligibly stated. I expressed this by saying that metaphysical realism is 'incoherent'. I did not mean

[4] For paradigm instances of realism in this sense, see David Lewis in "New Work for a Theory of Universals" and "Putnam's Paradox", Kit Fine in "The Question of Realism" and "The Question of Ontology", and Ted Sider in *Writing the Book of the World*.

[5] "Realism and Reason", p. 483.

> by that that it is inconsistent in a deductive logical sense, but rather that when we try to make the very vague claims of the metaphysical realist precise, we find that they become compatible with strong forms of 'antirealism'. Thus the attempts at clear formulation never succeed in capturing the content of 'metaphysical realism' because there is no real content there to be captured.[6]

The objection to metaphysical realism, then, is that talk of "reality" doesn't allow us to say anything new—it's "compatible with strong forms of 'antirealism' ", as Putnam puts it—because there is no further fact to be stated.

Along much the same lines, consider what David Chalmers says about the "heavyweight realism" who grants that mereological sums exist—in the ordinary sense of "exists"—but still thinks there's a further question whether mereological sums *really* exist. Against this view, Chalmers objects that "Intuitively, there is no further nontrivial truth to resolve concerning whether the mereological sum of two objects really exists."[7]

It's important to see that Putnam and Chalmers—amongst other opponents of metaphysical realism—aren't objecting to talk of "reality" or "the fundamental" as such.[8] They are objecting to a certain use or interpretation of this talk. In particular, they are objecting to *immodest* interpretation of these expressions—use which is intended to state *further* facts which we couldn't have stated without such expressions.

What's the sense in which these are *further* facts? Putnam and Chalmers don't make this explicit, but the most straightforward way to draw the line is in terms of explanation. For instance, Amie Thomasson explicitly contrasting her "simple realist view" with the type of realism (e.g., about properties) which she rejects:

> on the simple realist's view (as contrasted with the explanatory realist view), we can move from 'The house is red' to 'There is a property, redness, that the house has'. But we cannot use the latter to explain why the house is red—it is just a redundant way of

[6] "Replies", p. 353.
[7] "Ontological Antirealism", p. 103.
[8] Others include Shamik Dasgupta in "Constitutive Explanation".

> restating the former (introducing a new noun term for a property). So on the simple realist view, there are the disputed entities all right, but these are not 'posits' that are parts of 'theories', the inclusion of which is justified by their explanatory power.[9]

Along the same lines, Robert Kraut emphasizes that the key dispute about (e.g.) properties concerns their explanatory role—are they extra things which explain the familiar facts, or are they just a different way of describing those familiar facts? As he puts it, those who reject immodest realism claim that

> The "metaphysical facts," rather than providing justificatory and/or explanatory grounds for the norms that constrain linguistic behavior, serve to codify those very norms.[10]

Thus, the opponents of immodest realism which we're considering are happy to allow such talk of properties (etc.) when it is intended *modestly*. This is what Putnam calls "internal realism", what Chalmers calls "lightweight realism", and what Thomasson calls "simple realism".[11] More generally, there's a kind of deflationary metaphysician who is willing to admit new expressions so long these expressions are taken to be merely new ways of stating the same old facts.

Although it's a contentious label, I shall call these opponents of immodest metaphysical realism *antirealists*. Our antirealists—Putnam, Chalmers and Thomasson, to name a few—have no beef with modest, deflationary metaphysical realism; the key part of metaphysical realism to which they object is **Immodesty**.

And, **Immodesty** is precisely the part of metaphysical realism which I think is correct. My main focus is not to defend talk of "reality" or "the fundamental" as such.[12] Rather, my focus is to

---

[9] *Ontology Made Easy*, p. 157.

[10] "Universals, Metaphysical Explanations, and Pragmatism", p. 608. Similarly, although he doesn't put it quite as explicitly, it's also an important part of Thomas Hofweber's deflationary view of metaphysics that talk of properties (etc.) doesn't stand in need of the type of explanation offered by Armstrong, Lewis, or Sider (see *Ontology and the Ambition of Metaphysics*, chapter 11, esp. p. 282).

[11] "Realism and Reasons", p. 494; "Ontological Antirealism", p. 94–9; *Ontology Made Easy*, loc. cit.

[12] In fact, I argue in chapter 2 that positing these particular notions is a mistake, just as antirealists like Putnam contend!

argue against a reason for interpreting such talk in the deflationary manner proposed by Putnam, Chalmers, Thomasson, and other metaphysical deflationists.

Of course, you're free to use "reality" and "fundamental" in a deflationary manner to describe facts that you could have described otherwise. My argument is against the *antirealist* who claims there's nothing more than what can be stated by the deflationist. I argue that there are indeed further facts which we can't state without positing new notions.

1.1.2 Reduction and Antirealism

Many opponents of metaphysical realism have claimed that there's a connection between their rejection of *metaphysical realism* and their rejection of *reductionism*.

Consider the physicalist who claims that talk about (e.g.) the mental or the moral is defective unless it can be cashed out in terms of the physical facts. The reductionist—the physicalist in this case—claims that the mental and moral facts, if they are to be salvaged, must be reduced to physical facts. In a slogan: everything *real* is physical!

Many antirealists think that physicalism gives us an impoverished picture of the world, and they want to make sense of an alternative picture. They think that this can be done by rejecting realism. How? If there's no privileged structure—no "real" world in the realist's sense—then it's not the case that physics rather than psychology, ethics, politics or literary criticism is objectively privileged.

To take one example, Richard Rorty expresses this motivation for antirealism when he claims that by rejecting the realist's attempt "to isolate *privileged* representations" the antirealist is thereby "clearing the ground for morality and high culture rather than depriving them of 'objective truth'."[13]

Similarly, Nelson Goodman says that the "typical adversary" of the antirealist "is the monopolistic materialist or physicalist who maintains that one system, physics, is preeminent and all-inclusive, such that ever other version must eventually be reduced to it or rejected as false or

[13] *Philosophy and the Mirror of Nature*, pp. 174, 193 (emphasis in original).

meaningless."[14] And Mark Johnston notes that "Metaphysical Realism is the common element in scientism and mediaeval theology".[15]

On this way of understanding reductionism, the reductionist is someone who claims that familiar domains—psychology or ethics, for instance—cannot simply be taken at face value. These other domains depend for their factuality on the objectively privileged domain, the domain of facts which are *fundamental*, or obtain *in reality*, or which *carve at the joints*.

Let me note that it's not only the enemies of realism and reduction who understand things in this way. Ted Sider—to take one prominent realist—endorses a view on which all of the true but non-fundamental sentences are just a different way of describing the fundamental truths. As he puts it, "Every sentence that contains expressions that do not carve at the joints has a metaphysical semantics"; this is "a semantic theory" on which "meanings are to be given in purely joint-carving terms."[16] Thus, on Sider's view, if a sentence talks about non-fundamental things using terms which are not perfectly joint-carving, then the truth of this sentence depends on whether what it says can be cashed out in perfectly fundamental, joint-carving terms.

Many antirealists and realists alike see realism as an essential component of reductionism. They understand reductionism as the view which says there's a privileged domain of facts, and that the true sentences of every other domain can and must be cashed out as just a different way of describing the facts of the privileged domain.

If this is how we understand reductionism, then it's natural to think that we can reject reductionism by denying that there is any such privileged domain. One way to deny that there is any such privileged domain is to reject the intelligibility of the relevant notion of privileged—i.e., to reject that there are further facts we can state only by positing a notion of fundamentality (etc). Thus, on this way of understanding reductionism, one way to reject reductionism is to deny **Immodesty** and be an antirealist. And, in fact, this is how Rorty and Goodman describe their antirealist anti-reductionism.

[14] *Ways of Worldmaking*, p. 4.
[15] "Objectivity Refigured", p. 86.
[16] Ibid., p. 112.

I share the feeling that physicalism, scientism, and—more generally—reductionism provide impoverished pictures of the world. And, for this reason, I share the antirealist's desire to reject them. But I think that antirealists are mistaken to think that the best way to do this is by denying realism.

What we think it takes to reject reductionism turns on what we take reductionism to be. A central part of my defense of **Immodesty** is to argue that reduction has been misunderstood. In fact—I shall argue—it is precisely in order to reject reductionism that we should endorse **Immodesty**.

1.2 What is reduction?

In this section, I want to introduce three different ways of thinking about what it is to be a reductionist. Each of these views corresponds to a distinct account of what it would take to reject reductionism.

Let's assume that we're interested in reduction of the mental facts to the physical facts. We'll consider three types of proposals for how to say that the mental reduces to the physical: reduction as a claim about *equivalence*, reduction as a *realist* claim, and reduction as a *modest* claim.

1.2.1 Reduction as equivalence

One way to understand reductionism about (e.g.) the mental is as a claim about the *equivalence* of mental and physical facts. There are different types of equivalence. The most straightforward type—and the one what was traditionally tied to reductionism—is definitional equivalence.

> **Definitional Reduction** Mental expressions can be fully defined in terms of physical expressions. The meaning of every sentence about the mental is given by some physical sentence.

On this way of understanding reduction, the reductionist is committed to a pretty strong claim about how our language works. Clearly, some expressions have definitions in terms of other expressions. I can say: let "x is foo" mean "x is blue or green". In this case, "foo" is defined in terms

of "blue" and "green" because I stipulated the meaning of "foo". But in most cases, this isn't how expressions of our language get their meaning. Nobody stipulated the meanings of our mental expressions in this way. Thus, the definitional reductionist is committed to the idea that expressions of natural language—expressions which don't have their meanings fixed by stipulation—nonetheless have definitions in terms of other expressions of natural language.

Many people think that this is asking too much of the reductionist, and that it makes it too easy to reject reductionism. If reduction requires definability, then one way to reject reductionism is simply to deny that words have non-stipulative definitions. But the mere fact that expressions of natural language like "water" and "salt" don't have definitions shouldn't thereby rule out the idea that chemistry is—in some sense—reducible to physics.

1.2.1.1 Supervenience

An alternative to definitional equivalence is *supervenience*, or equivalence with respect to metaphysical possibilities.

> **Supervenience Reduction** Mental properties supervene on physical properties. There are no metaphysical possibilities which differ in the instantiation of mental properties without a difference in the instantiation of physical properties.

As Lewis put it,

> A supervenience thesis is, in a broad sense, reductionist. But it is a stripped-down form of reductionism, unencumbered by dubious denials of existence, claims of ontological priority, or claims of translatability.[17]

My worry for this way of understanding reduction isn't that it's false but that it's implicitly circular. Supervience is only a form of reduction if we are looking at the right sort of possibilities. But unless we have some independent grasp of the relevant notion of possibility, I worry that

[17] "New Work", p. 29.

we're saying reduction holds whenever you have supervenience in the possibilities which are the kind of possibilities relevant to reduction. And that's not a good account of what it is for one thing to be reducible to another!

There's a yet more serious worry: even if supervenience is sufficient for reduction, it's not necessary. And, Lewis' own corpus provide an example of reduction without supervenience. What I have in mind is his view about haecceitism and the identities of particular individuals.[18] On the one hand, Lewis didn't want to take identity-facts to be part of his basic ontology. He claimed that there were never pairs of possible worlds which were exactly alike with respect to the purely qualitative facts and different only with respect to which individual had which qualities. For instance, there's no world which is a qualitative duplicate of the actual world but in which I am in your shoes and you are in mine.

But on the other hand, Lewis also wanted to say that I could have been in your shoes. In order to say this, Lewis introduced a distinction between *possible worlds* and *possibilities*. Two distinct possibilities can be based on the same possible world, with a difference only in the counterpart relation under which we are considering the world. A counterpart relation maps an individual at one world to its counterpart at each other world, and usually that means that it maps you at each world to yourself at that world. But in order to get distinct possibilities, Lewis allows for counterpart relations which map (e.g.) me at a world to you at that world. Since the possible world together with this counterpart relation is what Lewis calls a "possibility", I get to say that I could have been in your shoes (holding fixed all the purely qualitative facts).

Once we have a distinction between possibilites and possible worlds it becomes important to ask: which of these is appropriate for a reductive supervenience thesis? Here, we have a worry about circularity.

But I now want to focus on a further worry for understanding reduction as a supervenience thesis, and it is this: Lewis himself wants to say that possibilities and identity-facts are "cheap", even thought they don't supervene on possible worlds! This shows that even if supervenience

[18] *On the Plurality of Worlds*, pp. 220–48, and especially pp. 230–5.

is one type of reduction, and even if we can independently characterize the notion of possibility which is relevant to reduction, supervenience isn't the only type of reduction. Thus, if you want to reject reductionism, it's not sufficient to deny supervenience.

1.2.1.2 Scrutability

So much for definitional equivalence and supervenience. Another way to characterize reduction would be in terms of *epistemic* equivalence. In particular, Chalmers has suggested that we should focus on what he calls *scrutability* theses—i.e., theses about the knowability of one fact (or class of facts) on the basis of some others.[19] Chalmers distinguishes many ways of formulating such theses, but most of the details don't matter for our purposes.

Let's state the basic idea of this way to understand reduction as

> **Scrutable Reduction** The truth of mental claims can be fully known merely on the basis of knowing the physical truths. There's no epistemic further question about the mental which isn't settled by the physical.

One kind of worry for this understanding of reduction targets the sufficiency of scrutability for reduction. This worry presses the possibility for us to have *a priori* knowledge of a necessary connection between two domains which is nonetheless a connection between irreducibly distinct domains. For instance, we might suppose that a Cartesian dualist who says that the mental and the physical are really and irreducibly distinct nonetheless claims to have *a priori* knowledge of the mental-physical bridge laws. In this case, the scrutability of mental facts from physical facts would not entail the reducibility of the one to the other.

But even if scrutability is insufficient for reduction, as long as it's a necessary condition it's possible for the anti-reductionist about (e.g.) the mental to be an anti-reductionist simply by denying the scrutability of the mental from the physical. What should we say about the idea that scrutability is necessary for reducibility?

[19] Chalmers motivates the shift from a definitional to an epistemic connection in *Constructing the World*, pp. 7–16.

A key move that Chalmers makes in responding to worries is to argue that the meaning of our expressions must themselves be scrutable. Why does he need to say this? The most obvious way for there to be reduction without scrutability would be for two descriptions to pick out the very same things in the world without this being something that's scrutable. For instance, it might be that our claims about the mental are just a different way of talking about the physical facts—thus, the mental is reduced to the physical—but there's no way for us to know this even if we know all the physical facts. So, to rule out this kind of case, Chalmers argues that "a view on which the theory of reference is a source of inscrutability should be rejected."[20]

But I think there's a problem with this move. In effect, Chalmers is assuming a reductive view about linguistic meaning. He's assuming that facts about what our expression mean are scrutable—because reducible to—facts about use and facts about the putatively reductive base. But if we reject such a reductive picture, then I think there's room for us to deny that you can reject reductionism by rejecting scrutability.[21]

Setting aside this possible source of inscrutability, I think there's a further worry for this way of thinking about reduction, and it's the same further worry that I pressed against understanding reduction as supervenience: there are facts which need to be included in the complete scrutability base which—plausibly—we still want to be reductionists about. In particular, Chalmers notes that "some indexical truths ('I am David Chalmers', 'It is now 3 a.m.') are not scrutable from objective physical truths alone" and "some negative truths ('There is no nonphysical ectoplasm') are not scrutable from positive physical truths alone", and for this reason Chalmers says we should include both indexical truths and a "totality sentence" in the scrutability base.[22]

But, the fact that you can't figure out which person you are simply on the basis of the objective physical facts doesn't mean that there's some irreducible further fact about your identity. And likewise, the fact that some set of facts gives a complete description of the world is plausibly not

[20]Ibid., p. 176.

[21]In fact, I think the non-reductionist ought to deny the scrutability of meaning. Unfortunately, the argument for this claim would take us too far afield; I plan to address this in other work.

[22]Ibid., pp. 110–1.

some further fact (God didn't have to make all the creatures and then make a totality fact in order for these to be all the things he made; he simply had to stop creating!).

Of course, one might hold a non-reductive view on which indexical facts or totality facts are genuine further facts over and above the physical facts. But one might also hold a reductive view of such facts while also acknowledging that they must be included in a minimal scrutability base. This shows that even if scrutability is one kind of reduction, scrutability isn't the only type of reduction. Thus, if you want to reject reductionism, it's not sufficient to deny scrutability.

1.2.2 Realist Reduction

If you can deny that indexical facts—to take one example—are definable or supervene on or are scrutable from the non-indexical physical facts, and yet still not be a reductionist about indexical facts, what more do we need to say to distinguish a reductive and a non-reductive view?

This is where the notion of fundamentality or some other immodest posit comes in. If we understand reductionism as a view about what's fundamental (or what has some other metaphysically privileged status) then we can say that we're reductionists about (e.g.) indexical facts by saying that only non-indexical facts are fundamental.

Thus, we can understand one kind of reductive view about the mental as the thesis

Realist Reduction Physical facts are fundamental; mental facts are not fundamental.

More generally, reductionism understood in this way is the claim that some restricted domain of facts are the privileged ones—the ones which describe how things "really" are—and everything else we say is either false or just another way of describing the world already described in a better way by (e.g.) the privileged facts of physics. Even if we can't define the non-fundamental expressions or know which ones are true purely on the basis of the fundamental truths, the idea of **Realist Reduction** is that we can still be reductionists about something by saying it's non-fundamental rather than fundamental.

Alternately, instead of positing a notion like fundamentality, we might posit a relation of metaphysical dependence or explanation—i.e., something like ground. We could then understand reduction as something like

Ground Reduction Physical facts ground mental facts.

Whether we posit a notion like fundamentality or a notion like ground, the basic idea here is to understand reductionism as a view which privileges one way of describing the world—the physical, for instance—over all others. For simplicity, I'll address my comments to the approach which posits fundamentality; what I say generalizes to approaches which posits notions like reality, joint-carvingness, or ground.

If we understand reduction along these lines, then we could think of the reductionist and non-reductionist as making different claims about what's fundamental. The reductionist about the mental says that only physical facts are fundamental, and the non-reductionist says: no, the mental facts are also fundamental.

But there's another way to reject reductionism, so understood. We can reject the relevant notion of privilege—e.g., fundamentality. We can say that it's not the case that physical facts are privileged over mental facts because there's no such thing as metaphysical privilege; there are simply lots of different way of talking about the world. And, as I discussed above, this is how antirealists take themselves to have a way to reject reductionism.

1.2.3 Two types of negations

It's important to understand what exactly the antirealist is doing here. It would be a mistake to think, for instance, that the antirealist is simply refusing to answer questions about reality, or that the antirealist is simply offering a different view about reality. Neither of these interpretations take seriously the idea that the antirealist can reject the very notions of "reality" or "fundamentality".

To see the antirealist's strategy, we need to consider the form of the argument given against metaphysical realism. In particular, many antirealists—Putnam, Goodman, and Rorty being no-

table examples—claim that metaphysical realism is to be rejected because it is *unintelligible*. More precisely, as we saw Putnam explain, the objection to the putative view called "metaphysical realism" is that there only appears to be a coherent view to be rejected, but upon closer inspection there turns out to be no such view.

1.2.3.1 Theism and atheism

Let me briefly illustrate the structure of this type of argument.[23] By stipulation, I shall use 'God' to mean a being who completely transcends experience. Consider now the question: *Does God exist?*

The most straightforward response is a simple *yes* or *no*. If you say *yes* then you are a theist; if you say *no* then you are an atheist.

The Positivist, however, says neither *yes* nor *no*. Instead, the Positivist responds by rejecting the question. Talk of a being who completely transcends experience is *meaningless*—says the Positivist—because a claim is meaningful only if it has empirical significance.

The Positivist rejects both simple theism and simple atheism. Still—and this is worth emphasizing—it's correct to say that the Positivist's criterion of meaning would give us an argument against theism and *for* atheism. The Positivist and the simple atheist agree about what there is; their disagreement concerns the coherence of the possibility that there is something more than what they agree there is. Thus, the Positivist counts as an atheist because the Positivist denies that there is anything which God could be.

More generally, one way to argue against theism is to claim that

(i) the theist intends to be committed to "something more" than what both the atheist and the theist countenance, and

(ii) the theist's description of the "something more" in question is unintelligible (incoherent, meaningless).

[23] This illustration was central to Denys Turner's undergraduate lectures on atheism at Princeton University in Spring 2017. I am grateful for his insight on the topic. See also his related discussion of negative theology in *The Darkness of God*, especially ch. 2.

1.2.3.2 Reduction and Immodesty

One of the central questions in this paper is: what happens when we reject the ideology posited by the immodest metaphysician—e.g., by the metaphysical realist? As I've noted, some prominent antirealists think that we end up with a non-reductive view.

Whether they're right about this depends on how we should understand reductionism. For instance, denying **Immodesty** is certainly not sufficient for rejecting **Definitional Reduction** or **Scrutable Reduction**. But, rejecting **Immodesty** could still play a part in your rejection of one of these forms of reduction: first, you'd deny the equivalence of (e.g.) the mental and the physical facts, and second, you'd deny that there's any further sense in which the mental facts are defective given this lack of equivalence to the physical facts.

As I'll argue shortly, I think that it's the non-reductive position rather than the reductive position which depends on **Immodesty**. And in fact, I think that we should understand the reductionist as making the same sort of move—the rejection of further ideology—which the antirealist tries (unsuccessfully) to leverage against the reductive position.

Before I explain how I see the reductionist employing this move, I want to discuss a case in which I think the antirealist is able to successfully employ this move.

1.2.3.3 Quantifier variance

Suppose we have two mereological simples: objects that don't have any parts. Some people claim that whenever we have two simples, there's also a third thing: the mereological whole which has each of these simples as a proper part. And some people deny there there's any such object, because they deny that anything has proper parts.

The first sort of person is a *universalist* about mereological composition, and the second sort is a *nihilist* about mereological composition. The universalist thinks that there are lots of mereological wholes—things with proper parts—while the nihilist thinks there are none—just simples.

Some philosophers think there's a substantive debate between universalism and nihilism (as well as the various intermediate positions which say simples *sometimes* compose a whole). But

other philosophers worry that the universalist and the nihilist are talking past each other, because each party can interpret the other as speaking the truth in some slightly different language. For instance, the universalist can interpret the nihilist as making claims in a language which implicitly restricts the quantifiers to simples. When the nihilist says "nothing has proper parts", the universalist can interpret this to mean that nothing (which is a simple) has proper parts.

Similarly, the nihilist may interpret the quantifier used by the universalist as a plural quantifier. When the universalist says "there's something which has Alice and Bob as proper parts", the nihilist can interpret this to mean that there are some things—e.g., Alice and Bob—such that Alice and Bob are amongst these things.

The quantifier variantist claims we can show that (i) there's a language which makes everything the universalist says come out true, (ii) there's a language which makes everything the nihilist says come out true, and (iii) there's a non-crazy translation from each language into the other. Let's suppose that this can be shown.

Does it follow that the dispute between the universalist and the nihilist isn't substantive? By itself, it does not.[24] For all we've said, it's possible that one of these languages is metaphysically privileged. For instance, if only simples exist in reality, then while we can still talk like universalists, we wouldn't be correctly describing fundamental reality when we do so. Or, if reality contains things with proper parts, then of course we can use a restricted quantifier which doesn't include these things in its domain, but in doing so we would be using a less fundamental quantifier.

At least—for all we've said—the dispute could turn out to be substantive in this way. If it's a dispute about which quantifier is more fundamental, then the fact that we can translate between the quantifiers does nothing to undermine the substance of the dispute.

Thus, we would need to add some further claim in order to say that the dispute is not substantive. What is this further claim?

Sider think that we need to add the claim that the two quantifiers—the one which makes the universalist's talk come out true, and the one which makes the nihilist's talk come out true—are

[24] Cf. Sider, *Writing*, p. 69.

equally fundamental. Because they are equally fundamental, neither is metaphysically privileged over the other, and so it's not the case that either universalism or nihilism is true *rather than* the other.[25]

But I think this isn't what the quantifier variantist should say. Instead of accepting the notion of fundamentality and saying that both quantifiers are equal with respect to this notion, we can deny that there's any such notion.

If we take this approach, the quantifier variantist can still say—loosely speaking—that the quantifiers are equally fundamental. But strictly speaking, what they should say is that neither quantifier is more fundamental than the other because there's no such thing as fundamentality.

Thus, what the quantifier variantist should add in order to say that there's no substantive dispute just is the claim that there's no further question. There are two different ways of talking—and that's all!

1.2.4 Modesty and debates about reduction

In this section, I'll explain how I think we should understand reduction.

To cut to the chase, I think we should understand reduction as a modest view. And part of why it's a modest view is that—like the quantifier variantist I just described—the reductionist is making a "that's all!" claim. But to see this, we need to think more carefully about the structure of a debate about reduction.

1.2.4.1 Reduction vs. non-reduction about what?

The reductionist and the non-reductionist disagree—for example—on whether the mental is reducible to the physics, or whether the moral is reducible to the natural.

One important question about any such debate is about how we demarcate the domains. How do we draw the line between the physical and the non-physical, or the natural and the non-natural, and so on?

[25] Ibid., pp. 46–7, 69–70.

Each way of drawing a line corresponds to a distinct version of physicalism, or naturalism, or whatever label is best apt for the demarcated domain. This then gives rise to questions about which version of physicalism (etc) is best, or which best earns the name "physicalism" (etc).

However we answer those questions, there remains a dispute about the relation between whatever we're calling (e.g.) the "physical" domain and some other relevant domain—say, the mental. My concern is with characterizing this kind of dispute. How should we describe various alternate positions regarding the relation between the two domains?

What I propose now is a particular structure for thinking about these sorts of debates. This structure will provide a simple and minimal way of stating two opposing positions.

Here is an outline of the structure. The two parties in the dispute have full agreement about one domain—say, claims made in purely physical or natural terms. The two parties have partial agreement about the second domain—say, claims made in mental or normative terms. The two parties disagree about the relation between the two domains, and this disagreement corresponds to a partial disagreement about the second domain.

1.2.4.2 Full agreement about one domain

Let me continue to take, as an illustration, a debate about the physical and the mental.

First, let me stipulate that the two parties in the sort of dispute I have in mind *fully* agree about one domain—in this case, claims stated in physical terms. It's not important that the two parties agree to call these claims "physical", but it is important that there's some set of sentences such that (i) both parties in the dispute agree about which of these sentences are true and (ii) they both agree about what's required of the world in order for these sentences to be true. In other words, they agree about the truth and the metaphysical nature of claims within the first domain.

Such full agreement is an idealization. In actual disputes, there will rarely if ever be complete agreement about the physical domain—i.e., which physical claims are true, and what these claims mean. Often these disagreements about the physical facts are incidental to the dispute about the

mental facts and can be ignored in practice. In such cases it's simpler for us to suppose that these disagreements have already been settled.

Let me emphasize that both parties agree about the nature of the physical facts. Without this, we won't be able to distinguish physicalism and non-physicalism in terms of the relation between the physical and the non-physical domains. For instance, you don't count as a physicalist if you say the mental is nothing over and above the physical because you're a panpsychist or an idealist. The debate between the panpsychist and the dualist is tricky, and I want to focus on the simpler debate in which both parties agree about the nature of one domain.

It's not important for our purposes that the parties agree about some way of demarcating the physical domain from the non-physical, or even that they think there's anything that makes the claims in this domain count as "physical". Calling the claims "physical" could be nothing more than a matter of convenience, not a description of some essential feature which all the claims have in common. Various views about the domain will give us different versions of "physicalism", but as long as the two parties in our dispute agree about the domain, the debate will have the structure that we're considering.

Finally, let me note that it's not important for our purposes whether the physical claims are in fact true. The present goal is to characterize disputes without regard to the correctness of either position in the dispute. We can bracket the question of the actual truth and actual nature of the physical so long as the two parties agree about the truth and nature of the physical claims.

That's what we need to specify about the first domain, about which the two parties fully agree. Let's consider now the second domain—in this case, the mental.

1.2.4.3 Partial agreement

As I said, the two parties *partially* agree about this second domain. More precisely, while the two parties disagree about the nature of the mental claims—they disagree about what's required of the world in order for these claims to be true—they agree about the truth of various mental claims. For instance, let's suppose they agree that John believes that Mary sees red. Let's suppose more

generally that both parties are happy to use the terms from folk-psychology and that they agree about which mental claims stated in these terms are true.

Like the agreement about the physical domain, this degree of agreement an idealization which allows us to ignore incidental disagreements concerning particular mental claims.

By considering only debates with this structure, I should note that we're thereby omitting the position which simply denies that there are any truths in the second domain—for instance, the eliminativist about the mental, or the error-theorist about the normative. Clearly there is an important disagreement between this type of (e.g.) physicalist and the non-physicalist, as well as an important disagreement between this type of physicalist and the physicalist who is happy to grant the truth of mental claims.

For our purposes, I want to focus on type of physicalist who agrees with the non-physicalist as much as possible while remaining a physicalist. In part, this is because my aim is to clarify the role which metaphysical realism or **Immodesty** might play in making sense of certain disputes, and there's no need to bring in any new notions in order to make sense of the dispute with the eliminativist or the error-theorist. (And in part, I don't feel the need to focus on the dispute with the eliminativist or error-theorist because I don't find this type of physicalism or naturalism at all plausible.)

1.2.4.4 Partial disagreement

The two parties in the disputes we're considering fully agree about one domain and agree about the truth of claims in the second domain. Where do they disagree? They disagree about the nature of the claims in the second domain, and in particular, they disagree about how much of what it takes for these claims to be true is already covered by the claims in the first domain and how much isn't.

To return to the question of the physical and the mental, the physicalist and the non-physicalist in this dispute disagree about whether the nature of the mental is exhausted by what they agree about, or whether it involves "something more" than this.

Of course, in some sense they can both agree that a mental claim involves "something more" than the physical claims—it is a distinct *sentence*. And perhaps it corresponds to a distinct *proposition*, depending on what we mean by "proposition".

For now we can note that whatever the sense is in which both the physicalist and the non-physicalist agree that the mental claim involves "something more", the non-physicalist goes beyond this when she claims that the nature of the mental involves "something more" than what's described by the physical claims.

Obviously this characterization of the debate is schematic. Our present concern is not with the question of where to draw the line between physicalism and non-physicalism, or natural and non-naturalism, or any other pair of more and less reductive views. There are many places one could draw the line, and each of these corresponds to a distinct, more specific dispute; we use labels like "physicalism" and "non-physicalism" to imprecisely pick out clusters of these disputes.

The particular views in the dispute will be determined by what exactly the two parties agree about and what exactly they disagree about. They both grant that there's some sense in which the mental can be "something more" that the physical without giving up physicalism; slightly different versions of physicalism will correspond to more or less restrictive ways in which the would-be physicalist can draw the line around the kind of "something more" which is compatible with her view.

Our focus is not on the question of where to draw the line but on the question of the nature of the line which we're drawing. How might we characterize the difference between more and less reductive positions?

When the debate can put into the structure I've proposed—full agreement about one domain, partial agreement about the other domain, disagreement about whether there is "something more" than what both parties agree about—then the nature of the distinction between the particular reductive and non-reductive positions is determined by what exactly the two parties in the dispute agree about. The non-reductive position is thus characterized by contrast with the reductive po-

sition. The reductionist gives a description of the world and says "that's all!" while the non-reductionist says there's "something more".

1.2.4.5 Modest Reduction

On this way understanding reduction, you can count as a reductionist without committing yourself to any immodest notions. You don't need to say that the physical facts have some metaphysically privileged status—the kind of thing that you'd posit a notion like fundamentality to state. All you would need in order to give your account of the world are the physical notions and the linguistic (or conceptual) notions which both the physicalist and the non-physicalist accept. Being a physicalist doesn't depend on some further fact, beyond the facts that can be stated with these notions; it depends on the claim that there is no further fact beyond these.

> **Modest Reduction** Reductionism (about some domain) is the view that there are no *further* facts beyond the facts of this domain. You're a reductionist because you say "that's all!" after describing the reductive domain.

If **Modest Reduction** counts as a form of reductionism, then what would we need to say in order to reject reductionism? We'd need to say that there are further facts beyond the agreed upon domain. The non-reductive view about the mental, for instance, would say that the truths about the mental domain are *further* facts over and above the facts in the physical domain.

Let me emphasize, I'm using "that's all!" as a technical phrase: it's precisely the rejection of the kind of new ideology which goes with **Immodesty**. This is the move the quantifier variantist can make to say that there's no further question about which quantifier is metaphysically privileged, and it's the move that the reductionist can make to say that (e.g.) talk of the mental requires nothing over and above the physical facts that both parties accept.

1.2.5 Rejecting Modest Reduction

Recall, the antirealist thought that they could reject reductionism by rejecting immodest ideology—e.g., the notion of fundamentality. This is an effective way to reject **Realist Reduction**. But, as I just argued, we don't have to understand reductionism in this way. We can instead understand realism as a modest claim. The modest reductionist gives a description of the world in modest terms—i.e., ideology which can be accepted by both the reductionist and the non-reductionist. And, what makes this a reductive view is that the reductionist says "that's all!", whereas the non-reductionist says there's "something more" than what's been described.

Ideally, if you wanted to reject such a modest reductionist view, you would do it by saying what the "something more" is, beyond those things described by the reductionist. For instance, you might posit a notion of fundamentality, and say that you're a non-reductionist about (e.g.) the mental because you—unlike the reductionist—are committed to the fact that some mental properties are fundamental.

You don't have to posit a notion of fundamentality in order to explain the alternative to **Modest Reduction**. In fact, in chapter 2 I argue that this is the wrong way to explain your non-reductive view.

But even if you don't posit a notion of fundamentality, you can't reject **Modest Reduction** without endorsing **Immodesty**. The way to reject reductionism, so understood, is to say that there's something more than what can be described already in modest terms. You say that there are further facts which we would need new ideology to talk about, and we'd need this ideology to explain the difference between reductionism and non-reductionism. And to say this just is to endorse **Immodesty**.

1.3 Why not reduction?

Many antirealists—those who reject **Immodesty**—think that the best way to reject reductionism is to reject metaphysical realism and **Immodesty**. But I've argued that this is wrong. There's a

version of reductionism which doesn't depend on the kinds of posits which the antirealist rejects—e.g., posits like fundamentality. Thus, the antirealist hasn't ruled out this kind of reductionism.

Suppose that I've convinced you it's harder to reject reductionism than the antirealist thinks. Grant—as I have argued—that we should understand reductionism as **Modest Reduction**, and that in order to reject reductionism so understood, you would need to endorse **Immodesty**.

At this point, you're free to say that it's not so bad if we can't rule out **Modest Reduction**. Maybe supervenience and talk of fundamentality is as much of reductionism that we have good reason to reject.

My goal in this section is to explain the cost of this move. We do have good reason to rule out **Modest Reduction**, and thus—from what I've argued already—we do have good reason to endorse **Immodesty**.

To make the case, I'll focus on a particular reductive dispute: a dispute about the nature of predication. I'll focus on this case because I think it best shows just how much you have to give up if you're not going to reject reductionism about anything.

1.3.1 Minimal predication

What does the world have to be like in order for there to be truths of the form *a* is *F*?

On one view, there are no such truths; this would be an error-theory about predication. Perhaps Parmenides comes the closest to endorsing an error-theory about predication by claiming that there are only truths with the form *a* is *a*. But let's set this view aside. It's very implausible that there's nothing more than truths like these.

One way for there to be "something more" than what Parmenides grants is for there to be truths about distinctness as well as truths about identity. If we grant distinctness facts, then we can say that there are (e.g.) three distinct things. But again, it's very implausible that there's nothing more than such truths.

One way for there to be "something more" than just identity and distinctness facts is for there to be facts about the inclusion of one thing in a group of things—e.g., that *a* is one of *a* and *b*

and *c*. Having got this far, we might also admit names for collections of things—let's say that "F" corresponds to the collection of *a* and *b* and *c*.

We now have truths of the form *a* is *F*. If we stop here and deny that there is anything more than this, we have a simple version of class nominalism. If we instead go on to admit not just classes of things but also set-theoretic constructions of things—sets of sets of things, ordered pairs of things, sets of ordered pairs of things, and so on—then we would have the resources for a much richer version of nominalism.

1.3.2 Putative accounts of predication

Is there anything more to predication than what we've accepted thus far? There are various things we might add: universals, tropes, a "resemblance" relation. All of these differ from what the nominalist countenances in that the things which the nominalist takes as corresponding to predicates are abundant—for every group of individuals, there's a class or a set or a concept. Universals, tropes, and "resemblance" are sparse—not every collection of individuals share a universal or have perfectly resembling tropes or resemble each other in some respect.

But just saying that there are sparse properties is insufficient to ensure that we're committed to something more that what the set-nominalist already accepts. This is because the nominalist can say that there's a set of properties (i.e., a set of sets) such that not every collection of individuals has one of these properties in common. Of course, the set-nominalist is committed to not just one such restricted collection of properties but many—every collection of properties corresponds to some sparse notion of properties. Should we distinguish the views by asking whether there are one or many notions of sparse properties? This too seems problematic since the nominalist grants that there is a property which picks out a unique class of properties—call it "*being a privileged class of properties*". Mere formal differences are not a promising way to tell apart nominalism from its more inflationary alternatives.

1.3.3 Something more than minimal predication

What is the nominalist failing to accept?

One suggestion is that the nominalist cannot explain *objectively privileged* genuine respects of similarity. Of course, the nominalist can say that there's some class of sets which count as the objectively privileged genuine respects of similarity. But the anti-nominalist should deny that any of the nominalist's sets could satisfy this description as it's meant by the anti-nominalist. Why? Because it's essential to the anti-nominalist's use of the expression that it involves something more than what the nominalist accepts.

Similarly, another suggestion is that the nominalist cannot explain the *qualitative* natures of properties. Of course, the nominalist can say that qualities are just distinct sets of individuals. But the anti-nominalist should deny that any mere set could count as a quality—at least, as the anti-nominalist understands qualities. Clearly, says the anti-nominalist, there's more to qualities than what the nominalist gives us.

In each case, what's essential to the non-reductive alternatives to nominalism is that they say there's "something more" than what the nominalist accepts. This is true even though it seems there's no neutral way of describing for the nominalist this "something more".

For instance, the anti-nominalist can say that she is committed to genuine qualities which the nominalist rejects. But what's the difference between a genuine quality and the sets which the nominalist calls "qualities"? The anti-nominalist says there's a difference, but what is it?

Any predicates which we use to try to describe in positive terms what the nominalist is missing can be interpreted by the nominalist as corresponding to sets—which the nominalist already accepts. Of course, the nominalist may not be correctly interpreting the predicates—this is what anti-nominalists should say—but the possibility of the nominalist's interpretation shows that what the anti-nominalist says only rules out nominalism if we presuppose a way of interpreting predication which is precisely what we were trying to describe.

In this way, the dispute about predication provides the simplest illustration of a position which cannot be precisely characterized in terms that both parties accept and understand. Because the

nominalist has available an interpretation of what the anti-nominalist says, it's not sufficient to pin down the anti-nominalist's position simply to state it. The anti-nominalist would also have to rule out the mistaken interpretation of what she says. But this just is the task of ruling out nominalism.

Thus, the anti-nominalist cannot rely on a precise *positive* characterization of what it is which she accepts and the nominalist rejects. The anti-nominalist does not reject nominalism by describing in positive terms something which the nominalist cannot—as interpreted by the nominalist—accept. Rather, the anti-nominalist rejects nominalism by accepting everything that the nominalist says there is, and then saying there is "something more" than this.

1.3.4 The cost of nominalism

Given the obscurity of the dispute about predication and given the difficulty in fully articulating the alternative to nominalism, why would anyone care about this dispute? Why would anyone have a reason to endorse or reject nominalism?

Not—at least in my case—because I have some direct awareness that there are (say) universals or tropes or that there is a more inflated relation of *instantiation* than what the nominalist admits.

For me, one central reason to reject nominalism is that the nominalist is unable to make sense of more inflationary positions about other things, and especially about the mental and the moral. I don't just mean that nominalism entails reductive views about other things—although I think it does. What I mean is that the nominalist doesn't even have the resources to say what the disputes about the nature of (e.g.) the mental or the moral are actually about.

From the nominalist's perspective, disputes about the nature of the mental or the moral can only be disputes about which classes correspond to which words—or something like that—and not the kind of deep and substantive dispute which people in the dispute take themselves to be having. In a nutshell, I find nominalism problematic because it entails a deflationary interpretation of every other dispute in metaphysics.

Let me be clear: I'm not saying that reductionism about predication is false because it entails reductionism about other things. I'm saying that reductionism about predication is false because

it entails that we can't even make sense of non-reductionism about other things! And I think that's a pretty implausible claim.

1.3.5 Non-reductionism without posit

There's an important objection to my claim that in order to reject **Modest Reduction** we must endorse **Immodesty**. There's a possibility which, thus far, I have overlooked: it might be that there's something more than what the reductionist accepts, but this something more doesn't involve positing anything new because it's something with which we have direct acquaintance.

Take the non-reductive view about predication. Clearly—I think—we want to say there's something more than just individuals and arbitrary classes of individuals. I suggested than in order to say this, we should say that there's something which would be explained by positing a new notion. But, instead, we might think that we are directly acquainted with the thing that the reductive nominalist is failing to recognize. In particular, we're directly acquainted with the sensible qualities—the genuine respects of similarity between (e.g.) two patches of blue.

Similarly, one might think that we're directly acquainted with what the reductionist about the mental and the moral are omitting: *thinking*, in the case of the mental, and *experiencing pleasure*, in the case of the moral.

If the only things that we want to be non-reductionists about (or, to make sense of non-reductive views about) are things like this—things that the non-reductionist claims to have direct acquaintance with—then presumably there's no need to posit something further, and there's thus no need to endorse **Immodesty**.

I don't have a decisive response to this objection. However, I'll note that this option is an alternative to **Immodesty** only if it handles all of the cases. In particular, if we want to be non-reductionists about predication, then we'd need to say that all genuine respects of similarity—the things that the reductionist doesn't accept—are directly accessible to us. In other words, we'd need to say that they are all phenomenal qualities.

So, if you want to be a non-reductionist without endorsing **Immodesty**, you could instead adopt a position which is, in effect, a form of idealism. You could say that to be is to be perceivable. But—I think—that's a pretty high price for avoiding **Immodesty**.

1.4 Immodesty

In this paper, I have argued that we have good reason to endorse

> **Immodesty** There are things we can talk about only by positing them as metaphysical explainers.

The argument has two main parts. First, I argued that it's the non-reductionist rather than the reductionist who is committed to "something more". And this means that you can't reject reductionism simply by rejecting **Immodesty**, as various antirealists would like to do; instead, you should reject reductionism by endorsing **Immodesty**.

Second, I argued that we do indeed have good reason to reject reductionism. In particular, there's a reductive position about predication on which predicates are all just arbitrary classes. Maybe you're happy to accept this position, but it strikes me and many others as a pretty implausible view. Thus, in order to reject this reductive position about predication, we should say there's something more than just arbitrary classes.

One possibility is that the "something more" is something we can talk about without posit, because it's something that's directly accessible to us. I haven't given an argument against this possibility, but I and many others find the resulting idealism to be implausible.

Thus, in order to reject reductionism about predication, we should endorse **Immodesty** by saying there are facts, including facts about the nature of predication, which we can only talk about by positing some new ideology.

Chapter 2

Immodest, yet Articulate, Metaphysics

In chapter 1, I defended a key thesis rejected by antirealists:

> **Immodesty** There are things we can talk about only by positing them as metaphysical explainers.

When David Lewis posited naturalness, it was supposed to be a *new* notion—not definable in terms of notions we already had—which could be used to explain the difference between green and grue.[1] When Kit Fine posited essence, it was supposed to be a *new* notion—not definable in other terms—to explain part of the difference between the fact that necessarily Socrates exists if his singleton exists and the fact that necessarily the singleton of Socrates exists if he does.[2]

In chapter 1, I addressed one worry raised by antirealists about these explanations. The antirealist argues that we can take the familiar facts at face value, without positing anything metaphysical to explain them. Thus, according to the antirealist, there's no explanatory work for these posits to do. I have argued that this thought is wrong; there is explanatory work to be done. In particular, we should endorse **Immodesty** in order to make sense of non-reductive views. But, this doesn't show that every immodest explanation by posit is in good standing—it doesn't show that these specific posits can do the explanatory work which needs to be done.

[1] See "New Work For a Theory of Universals".
[2] See "Essence and Modality", p. 5.

Lots of people—even some metaphysicians!—are skeptical about the *intelligibility* of the explanations given by such posits. Lewis told us that green is more natural than grue. But what does this mean? The skeptic—Hilary Putnam, to take one example—worries that this "explanation" is spurious.[3] At best, it is just a different way of talking about the things that were to be explained; at worst, it is unintelligible nonsense.

The task of this paper is two-fold. First, I'll argue that this worry should be taken seriously. I'll do this by defending a particular constraint on immodest positing:

> **Articulation** We shouldn't posit a metaphysical explainer if the resulting explanation would rely on an epistemic brute difference.[4]

Second, I'll explain the upshot of this constraint. One possibility is that it rules out very little of what the immodest metaphysician wants to say, so we should continue positing as we please. Another possibility is that it rules out everything the immodest metaphysician wants to say, so we should reject immodest metaphysics. I'll defend a third possibility. We can reconcile **Immodesty** with **Articulation** only by positing some fairly peculiar notions.

Our discussion will be framed around a particular explanatory project: explaining the difference between two positions about the nature of persistence. Broadly speaking, there are two camps about the nature of persistence. Four-dimensionalists think what it is for an object to persist across time is very similar to what it is for an object to be extended across space. Three-dimensionalists deny this. They say that persistence across time is importantly different from extension across space.[5] How exactly is persistence supposed to be similar or different from extension? This is the question we'll address.

---

[3] See "Realism and Reason" and the remarks reported by Lewis in "Putnam's Paradox".

[4] Roughly, something is an epistemic brute differences *for you* if you are thinking about it by means of "primitive" and "undefined" concepts. I'll say more about this later.

[5] As Kit Fine puts it, three-dimensionalists "have thought that there is a distinctive way in which material things are present in time as opposed to space…. The 3D-ers have then held that things exist in time but are extended in space while the 4D-ers hold that things are extended both in space and in time" ("In Defence of Three-Dimensionalism", p. 699).

The basic plan of the paper is to first motivate **Immodesty**, then motivate **Articulation**, and finally to see what we get when we put them together. We'll do this by attempting to explain the difference between three- and four-dimensionalism.

Section 1 addresses **Immodesty**. I argue that there's work to be done by a posit because we can't give a satisfactory account of the dispute about persistence without positing a new metaphysical distinction.

Sections 2 through 4 address **Articulation**. In section 2, I try to distinguish three- and four-dimensionalism by positing an unexplained, epistemically brute difference. In section 3, I argue that this is a bad idea; more generally, I defend **Articulation**. In section 4, I argue that many of the sophisticated attempts to distinguish three- and four-dimensionalism are no better.

And in section 5, I suggest that we can harmonize **Immodesty** with **Articulation**. It's possible to posit new notions without also positing brute differences. This yields a new account of the difference between three- and four-dimensionalism.

2.1 Why posit a new distinction?

The goal in this section is to argue that we can't give a *satisfactory* account of the difference between three- and four-dimensionalism unless we accept

> **Immodesty** There are things we can talk about only by positing them as metaphysical explainers.

There are lots of disputes which can be characterized in *modest* terms. These are disputes which you can explain while consistently denying **Immodesty**. You can fully explain these disputes using notions which are in good standing outside of metaphysics: empirical notions, logical notions, or otherwise familiar notions. When a dispute can be fully explained without positing any new metaphysical notions, let's say that it is a *modest dispute*.

Should we interpret the dispute between three- and four-dimensionalists as a modest dispute? Of course, we could use the names "three-dimensionalism" and "four-dimensionalism" for posi-

tions in some modest dispute. And in fact, many people have stipulated that they're using the names for a dispute about, e.g., whether a persisting object has instantaneous temporal parts at each of the times when it exists.

You're welcome to engage in one of these modest disputes if that's what you want to do. But, I shall argue, modest disputes aren't the disputes we should be having if we want to have a dispute about the nature of persistence. My general strategy here will be to show that many of the modestly stated questions used to distinguish three- and four-dimensionalism really aren't questions about the nature of persistence.

Let me start with an illustration of the strategy. Some three- and four-dimensionalists disagree about the ontology of time. Many four-dimensionalists are eternalists: they think that the past and the future exist. Many three-dimensionalists are presentists: they think that only the present exists. But, it's a mistake to draw the line between three- and four-dimensionalism here, and it's a mistake because this question is orthogonal to the question about the nature of persistence. Three- and four-dimensionalists alike *can* be eternalists; they can agree about the ontology of time while still disagreeing about what it is for objects to persist through time.

For this reason, we may as well assume eternalism; this will simplify our discussion. If you can't correctly distinguish three- and four-dimensionalism while assuming eternalism, then you aren't distinguishing the views in the right way more generally.

Let's now examine various proposals for how to draw the line between three- and four-dimensionalism in modest terms. While we won't be able to consider every possible proposal, I think what I say about these proposals will generalize to other modest characterizations of the dispute. We'll consider distinguishing the views in empirical terms; in terms of existence, mereology, modality, and identity; and in terms of epistemic or conceptual or psychological priority.

2.1.1 Empirical facts

As far as I can tell, no one thinks that the question disputed by three- and four-dimensionalists is a straightforwardly empirical question. Thus, we can't simply identify three- and four-dimensionalism with an empirical claim.

I am happy to grant that the dispute between three- and four-dimensionalists is empirical in the sense that empirical evidence can bear on the dispute. But to distinguish the views themselves, we should look beyond the empirical facts.

2.1.2 Existence

We might distinguish the views by what they say about the *existence* of certain entities: instantaneous temporal stages.

Instantaneous temporal stages are momentary objects which exist at each time when a persisting object exists. David Lewis and Ted Sider, both advocates of four-dimensionalism, draw the line between three- and four-dimensionalism in terms of the *existence* of such entities. This means that to be a three-dimensionalist, you would need to deny that there are instantaneous temporal stages.[6] Thus:

> **Existence** Four-dimensionalism is the thesis that every persisting thing has instantaneous temporal stages at every time at which it exists. Three-dimensionalism is the denial of this.

Three-dimensionalists should object. Kit Fine, a three-dimensionalist, tells us that "even the three-dimensionalist might be willing to admit that material things have temporal parts."[7] This is because adding extra things to your ontology—temporal stages, in this case—doesn't thereby commit you to a particular view about the nature of persistence.

[6] This is implicit in Lewis' discussion of temporary intrinsics, as his objection to three-dimensionalism only goes through if the three-dimensionalist must deny that there are instantaneous stages (*On the Plurality of Worlds*, pp. 203–4). Sider is explicit that "the more basic question of whether temporal parts exist at all" is his "primary concern" (*Four-Dimensionalism*, p. 60).

[7] "Guide to Ground", p. 42. See also Donnelly, "Endurantist and Perdurantist Accounts of Persistence", pp. 41–44.

What reason could the three-dimensionalist give to explain the existence of so many short-lived objects? One suggestion is that the instantaneous temporal stages are cheap additions to the three-dimensionalist's ontology. An illustration: Socrates sometimes sits and sometimes stands. Perhaps with no further cost, we can expand our ontology to include an object which only exists when Socrates sits: sitting-Socrates. Likewise, perhaps with no further cost the three-dimensionalist can expand her ontology to include an object which only exists at time *t*: existing-at-*t*-Socrates. And so, the three-dimensionalist can admit that there are instantaneous temporal stages.[8]

Of course, two parties are free to have a dispute about the existence of instantaneous temporal stages, and they can call their views "three-dimensionalism" and "four-dimensionalism". My point is that a dispute about the existence of instantaneous temporal stages doesn't have to be a dispute about the nature of persistence. We might have two parties who entirely agree about the metaphysics of persistence but disagree about whether our ontology should include momentary objects at each time when the persisting object exists.[9]

In order to set aside this orthogonal dispute, let's suppose the three-dimensionalist grants the existence of just as many instantaneous temporal stages as the four-dimensionalist admits.

It's tempting to say that there's an important difference between the temporal stages admitted by the four-dimensionalist and those which can be admitted by the three-dimensionalists. The three-dimensionalist's stages are—you might say—just cheap substitutes. Fair enough—I say—but what's the difference between the things which the four-dimensionalist calls "instantaneous temporal stages" and the things which the three-dimensionalist can admit? To distinguish three- and four-dimensionalism along these lines, we would need to look beyond the mere existence of instantaneous temporal stages and say something about the difference between the "genuine" and "cheap" stages.

How then are we to draw the line? One natural thought is that three- and four-dimensionalists disagree about the relation between stages and persisting objects, or the relation between facts

[8]For Fine's account, see his discussion of rigid embodiments in "Things and Their Parts", pp. 65–8.

[9]As Ofra Magidor argues (in "Endurantism vs. Perdurantism?: A Debate Reconsidered"), many of the arguments against three- or four-dimensionalism have in fact targeted an orthogonal thesis about how plenitudinous our ontology of stages is.

about stages and facts about persisting objects. In the next three sub-sections, we'll consider some increasingly sophisticated attempts of this sort.

2.1.3 Mereology

We might draw the line in terms of mereological relations. On this proposal, the three-dimensionalist denies that persisting objects are *mereological sums* of temporal stages:

> **Mereology** Four-dimensionalism is the thesis that persisting objects are mereological sums of instantaneous temporal stages. Three-dimensionalism is the denial of this.

What does this mean? Mereological sums are just collections or aggregates of parts without any metaphysical principle of unity—they are like heaps. Although the three-dimensionalist can grant the existence of temporal stages, they can't say that the relation between a persisting object and its temporal stages is just like the relation between a heap and its grains of sand.

Insofar as there is an orthodox view about the difference between three- and four-dimensionalism, this is it. For instance, in a recent paper arguing that the question about how many temporal stages there are is orthogonal to the dispute about persistence, Ofra Magidor simply takes it for granted that the dispute about persistence is a dispute about whether these stages are *parts*.[10]

One small problem for this proposal is that it incorrectly classifies positions which should fall on the four-dimensionalist side. While Lewis thinks that ordinary objects are mereological sums of temporal parts, stage-theorists like Ted Sider and Katherine Hawley deny this. They think tables and chairs and other ordinary objects are the momentary temporal stages—these things persist by standing in the appropriate relation to other stages at other times.[11] Further, the stage-theorist may even deny that there are any mereological sums—there are just simples and sets of simples.[12]

[10] Ibid.

[11] Sider thinks this relation is analogous to the Lewis' counterpart relation ("All the World's a Stage", p. 437); Hawley thinks it is a relation of counterfactual and causal dependence ("Persistence and Non-Supervenient Relations", p. 63; *How Things Persist*, p. 85).

[12] Sider is such a stage-theorist. He argues that once we have sets, there's no need to also have mereological sums ("Against Parthood", pp. 284–5). See also Mark Johnston's argument against accepting both sets and sums ("Hylomorphism", pp. 687–696).

What unifies the different versions of four-dimensionalism? What's the common thesis which is being rejected by three-dimensionalists?

For Lewis, as soon as there are instantaneous temporal stages with the appropriate intrinsic properties and spatio-temporal relations to each other, there are persisting objects.[13] Sider agrees with Lewis about what the world has to be like in order for there to be persisting objects—as soon as the appropriately arranged temporal stages are there, we've got persisting objects. Sider and Lewis chiefly disagree about how to label things; when asked to pick out the tables, Lewis will point to the mereological sums of stages and Sider will point to the individual stages. But, for both Lewis and Sider, once we are committed to the appropriately arranged temporal stages, the existence of the persisting objects—whatever they are—is not a *further* commitment.

As I see it, this is a key point of disagreement between Lewis and Sider on the one side and the three-dimensionalist on the other. They disagree about the extent to which persistence involves anything over and above the facts about instantaneous temporal stages. Four-dimensionalism is a "reductive" view on which persistence involves nothing—or at least, hardly anything—over and above facts about temporal stages. Three-dimensionalism denies this; persistence, according to the three-dimensionalist, involves either something further or something independent from the facts about temporal stages.

Let's consider one way to cash out talk of "nothing over and above" and draw the line between three- and four-dimensionalism.

2.1.4 Modality

As I've noted, three- and four-dimensionalists can agree about the existence of temporal stages, and they can agree about the intrinsic properties and spatio-temporal relations of these temporal stages. Let's call what they can agree about the *stage-facts*.

[13] Here I grant Lewis the claim that as soon as there are some individuals, there is also a mereological sum of those individuals: "Mereology is ontologically innocent…. given a prior commitment to cats, say, a commitment to cat-fusions is not a *further* commitment. The fusion is nothing over and above the cats that compose it" (*Parts of Classes*, p. 81).

Many four-dimensionalist claim that facts about persisting objects *supervene* on the stage-facts.[14] This means that when any two possible worlds agree about the stage-facts, they also agree about the persistence-facts. Thus:

> **Supervenience** Four-dimensionalism is the thesis that the intrinsic properties of persisting things supervene on the intrinsic properties and relations among their instantaneous temporal stages. Three-dimensionalism is the denial of this.

What motivates such a supervenience thesis? As Lewis puts it when discussing supervenience more generally:

> A supervenience thesis is, in a broad sense, reductionist. But it is a stripped-down form of reductionism, unencumbered by dubious denials of existence, claims of ontological priority, or claims of translatability.[15]

Now, Lewis and Sider both endorse something like this supervenience thesis. And many three-dimensionalists will deny it. But I don't think this is a good way to draw the line.

The dispute between three- and four-dimensionalists was supposed to be a dispute about the nature of persistence in the actual world. If there's a dispute about the range of other possible worlds, then this should be on account of an underlying dispute about the nature of persistence in the actual world.

We can see this more vividly by supposing that the three- and four-dimensionalist somehow both become convinced that the actual world is the only metaphysically possible world. This should be orthogonal to their dispute about persistence. If they accepted this claim about other possibilities, they would agree the persistence facts supervene on the stage-facts, since everything supervenes on everything if every truth is necessary. But I don't think they would be forced to drop their dispute about the nature of persistence.[16]

---

[14] E.g., Lewis, "Humean Supervenience Debugged", pp. 225–6. But Hawley, as I noted above, claims that we also need to include facts about causal and counterfactual relations between the stages in the supervenience base.

[15] "New Work", p. 29.

[16] Shamik Dasgupta employs this move to show that we shouldn't use a supervenience claim to distinguish substantivalism and relationalism about space—see his "Bare Necessities", pp. 117–8. Sider generalizes the move to other

Thus, even if the typical three- and four-dimensionalist do disagree about certain modal claims, there are good reasons for thinking that the nature of persistence is not essentially a modal matter. If we want to characterize a dispute about the nature of persistence, we need to supplement the modal question.

2.1.5 Identity

Until now, we've been considering ways to characterize three-dimensionalism as the negation of four-dimensionalism. But clearly a more satisfying approach would be to give a positive statement of the nature of persistence according to the three-dimensionalist. Let's consider one promising attempt we can make using modest notions.

The idea of this proposal is to distinguish three- and four-dimensionalism in terms of *identity*—identity in the strict and numerical sense. The three-dimensionalist says that the flower which is now faded and the bright flower of yesterday are the very same flower—there is only one flower. On the other hand, it's tempting to think that the four-dimensionalist must deny this numerical identity, since the stage which existed yesterday and the stage existing today are numerically distinct.

But this initial suggestion is too quick. Of course the four-dimensionalist will say that distinct temporal stages are distinct. Nevertheless, the four-dimensionalist can also say that there's just one flower in the strict and numerical sense. Lewis, for instance, will say that the flower is the mereological sum of the stages, and there's one such sum existing both yesterday and today.

This response provides only minor respite. The mereological sum may be numerically identical over time, but it is the sum of temporal stages which are not identical over time. As Lewis explains,

> Perdurance [i.e., four-dimensional persistence] corresponds to the way a road persists through space; part of it is here and part of it is there, and no part is wholly present

debates in his Locke Lectures, and gives further reasons for the shift to "postmodal" tools—see *Tools of Metaphysics and the Metaphysics of Science*, ch. 1.

> at two different places…. Endurance [i.e., three-dimensional persistence] involves overlap: the content of two different times has the enduring thing as a common part. Perdurance does not.[17]

Thus, setting aside sums, we might try to characterize three- and four-dimensionalism in terms of the way in which objects are numerically identical across time:

> **Identity** Three-dimensionalism is the thesis that there's something (not a mereological sum of temporal parts) which is numerically identical across distinct times. Four-dimensionalism is the denial of this.

Thomas Hofweber and David Velleman propose a distinction in this vicinity. Like me, they find the debate about parts unsatisfactory. They propose that we instead focus on how the identity of an object at any particular time is determined.[18] Does the identity of the object, at any time, supervene on properties which are *intrinsic* and *local* to that time? If the persisting object is the mereological sum of temporal parts, then its identity at any time is determined by the identities of the temporal parts existing at different times. But we can see that there's another way to persist by reflection on the idea that our personal identity is "fully determined within the context of each moment of its existence."[19]

I think these proposals, while promising, are still unsatisfactory. The four-dimensionalist can admit persisting objects which are identical across time—identical in the strict and numerical sense—and whose identity at each time does not depend on entities at other times. And, the four-dimensionalist can admit such entities while remaining a four-dimensionalist because these entities are supposed to be ontologically *cheap*, just like the instantaneous temporal stages admitted by Fine and the mereological sums admitted by Lewis.

Thus, the four-dimensionalist can mimic the three-dimensionalist by admitting yet another kind of entity which is "nothing over and above" the four-dimensionalist's temporal stages. Let me explain.

---

[17] *Plurality*, p. 202.
[18] "How to Endure", p. 55.
[19] Ibid., p. 51.

2.1.6 Three-dimensionalism on the cheap

Consider a sequence of individuals which occupy some role—say, the monarchs of the United Kingdom. Given our commitment to the individual monarchs, we can say that it is not a *further* commitment to admit an object for which the monarch-role is essential but the particular monarchs which play that role are not essential. Let's call this entity "the Crown", and say that the individual monarchs—the things which happen to play the role which is essential to the Crown—the "realizers" or "embodiments" of the Crown.

What's important for our purposes is that the Crown is not essentially embodied by any particular monarch, although it cannot exist without being embodied by some individual monarch or other. And, it is fully explained in terms of the individual monarchs and the monarch-role which they each play; to get the Crown from the monarchs, we simply shift which properties are essential and which are accidental.[20]

In much the same way, the four-dimensionalist can admit something which is numerically identical over time yet not a mereological sum of the temporal stages which "embody" it at distinct times. If you like, you can think of this as like a mereological sum but with the peculiar features of its parts "abstracted away".

Because these entities would be fully explained by the things which the four-dimensionalist accepts, the four-dimensionalist is free to add them to her ontology while remaining a four-dimensionalist. Like the mereological sum, the variable embodiment is supposed to be a cheap addition to the four-dimensionalist's ontology. Like the sum, it's fully explained in terms of things which the four-dimensionalist accepts, although its essence includes less than the essence of the corresponding sum.

Each variable embodiment is strictly and numerically identical across times. And because the temporal stages which embody the entity are not essential to it, its identity at each time is not

[20] Fine admits such entities as what he calls "variable embodiments"; see "Things and Their Parts", pp. 68–72. This is an instance of what Johnston and Leslie refer to as "ex-ing", or the operation of removing something from an essence; see their "Essence and Accident".

determined by the identities of these temporal stages.[21] Thus, we have yet again failed to draw a satisfactory line between three- and four-dimensionalism.

It's tempting to say that there's an important difference between the persisting objects admitted by the three-dimensionalist and variable embodiments which can be admitted by the four-dimensionalist. The four-dimensionalist's variable embodiments are just cheap substitutes. Fair enough—I say—but what's the difference? We've still failed to precisely characterize the nature of persistence according to the three- and four-dimensionalists.

2.1.7 Deflationary accounts

One possibility is that we're looking for a difference in the wrong place. Clearly, the three- and four-dimensionalist are describing the world in two different ways. But maybe there's no genuine disagreement about the nature of persistence.

If there's no genuine disagreement about the nature of persistence—just two different ways to describe the same state of affairs—then we might conclude that the three- and four-dimensionalist are talking past each other and not having a genuine dispute.

But we don't have to draw that conclusion. We might instead say that they're having a dispute about the epistemic, conceptual, or psychological features of these competing descriptions: which notions are easier for us to use, more deeply entrenched in our psychology or language, simpler to define using concepts which figure in the deliverances of our perceptual system, and so on.[22]

To take such an approach is to offer a *deflationary* account of the dispute; it locates the point of contention in us and our relation to the world rather than in the world as it is in itself.

I won't argue that the deflationist is wrong. Maybe there is no genuine dispute to be had about the nature of persistence, and in this case, maybe we should interpret three- and four-dimensionalism along deflationary lines. But, even if this is the best we can do, I think we should recognize that it's not giving us everything we wanted.

[21] For the same reason, it's also a counterexample to **Supervenience**.

[22] Famously, Nelson Goodman distinguishes projectable and non-projectable properties—green and grue—by citing a difference in entrenchment; see "Prospects for a Theory of Projection" in *Fact, Fiction, and Forecast*.

Disputes about concepts, human psychology, and natural language can be important and substantive. But I don't think the typical three- and four-dimensionalists take themselves to be having a dispute about these things. They intend to have a dispute about the world which doesn't turn on contingent facts about our concepts or our psychology or our language.

Or, at least some three- and four-dimensionalists would like to be having a deeper dispute than this, so we should look for another way of drawing the line between their views. Disputes which turn on contingent facts about us—our concepts, psychology, and language—are *shallow* disputes. For the rest of the paper, when I say "three- and four-dimensionalism", I shall mean *non*-shallow three- and four-dimensionalism.[23]

2.1.8 Immodest metaphysics

Our survey so far suggests that we can't give a satisfying account of the three- and four-dimensionalist positions using only terms which are sufficient for describing the facts of science and common sense. Of course, there are many specific proposals which we have ignored. But the problem with the proposals we considered is fairly general. These proposals fail because they shift the dispute. We wanted to characterize a dispute about the nature of persistence. But instead we get—for instance—a dispute about the ontological innocence of admitting variable embodiments. Or, we get a shallow dispute about epistemic, conceptual, psychological or linguistic facts.

My goal thus far has been to show that there's work for metaphysical posits to do in explaining the dispute between three- and four-dimensionalists. We can't get everything we want without positing metaphysical explainers.

It doesn't follow that we *can* get everything we want. It also doesn't follow that metaphysical posits will give us a more satisfactory account of three- and four-dimensionalism. But I think we have sufficient reason to consider immodest accounts of the dispute. This is what we shall do for the remainder of the paper.

[23]Some philosophers are happy to interpret most or all metaphysical disputes as shallow—see, e.g., Eli Hirsch ("Quantifier Variance and Realism") or Amie Thomasson (*Ontology Made Easy*). And although Kit Fine thinks that some metaphysical disputes are deep, he too thinks that a large part of the project of metaphysics is to describe "the world of appearances" rather than "the world of reality" ("Naive Metaphysics", p. 110).

2.2 Bald-faced primitivism

What if the full story is that there's some basic feature of the world—call it "endurance"—which the three-dimensionalist accepts and the four-dimensionalist denies? If "endurance" names a basic feature of the world, isn't the best—and perhaps only—way to draw the line of dispute in terms of this notion?

The proposal here is to posit that there is a *brute* difference between the world as described by the three-dimensionalist and the world as described by the four-dimensionalist. If there is a brute difference, then we may have a name for the difference—"endurance" in this case—but we aren't able to give any account of what constitutes the nature of the difference.

There are two sorts of worries about brute differences. One might worry about brute difference in general, or about positing this particular brute difference.

Even if there can be brute differences—and it's hard at the moment to see how they could be avoided completely in a full account of the world—it's plausible that we should give up on the possibility of explaining the difference between perduring objects and enduring ones only as a last resort once we've failed to explain what might constitute the difference.

Here's the more general worry. Suppose we fail to explain what constitutes the difference between three- and four-dimensionalism. One conclusion we might draw is that we failed because there is a brute difference—a difference which isn't constituted by anything further. But another possibility is that we failed because there is no difference—there is no genuine dispute about the nature of persistence. How do we tell the difference?

Given how easy it is to posit a brute difference, it's reasonable to worry that the would-be three-dimensionalist has retreated to an empty label and there is no such thing as endurance. Why should we believe that there is a genuine alternative to four-dimensionalism if all the three-dimensionalist can say is that there's something—endurance—and that it's not any of the things the four-dimensionalist can accomodate in her account of persistence.[24]

[24]This question is analogous to that pressed by Richard Rorty against the Cartesian picture of the mental (*Philosophy and the Mirror of Nature*, ch. 2).

These worries are not decisive. But I think there is a more decisive worry about admitting brute differences when we're drawing the line between three- and four-dimensionalism. Let me press that worry now.

2.3 Against positing brute differences

The goal in this section is to defend the core thesis of the paper:

> **Articulation** We shouldn't posit a metaphysical explainer if the resulting explanation would rely on an epistemic brute difference.

An *epistemic brute difference* is when you believe that there's a difference, and you have a label for the difference—e.g., being *F* vs. being *non-F*—but you don't have any beliefs about what it is to be *F* rather than *non-F*.[25] For instance, you're acquainted with two colors, you believe that there's a difference between two colors, but you lack the conceptual resources to form beliefs about any specific respects in which they're different.

Let me emphasize that I am talking about notions which are *epistemically* brute. Something can be epistemically brute *for you* even if it in fact has a metaphysical analysis. It's epistemically brute because *you* are entirely ignorant of this analysis.

The intuitive worry about positing brute difference is that we don't understand what we're positing; we have labels, but we can't tell what if anything these labels pick out. Epistemic brute difference are *unintelligible*, and unintelligible explanations aren't much good.

The obvious response to this worry is that while we lack knowledge by description of our posits, we can still understand them through knowledge by acquaintance.[26]

[25] You might worry that I'm smuggling in a metaphysical notion of "what it is to be". I'm not. By "what it is to be *F*" I mean whatever features you take to be relevant to explaining whatever it is that you're introducing the posit to explain. Maybe you know that Ariel prefers things which are *F* to things which are *non-F*. But if you're positing *F* to explain three- and four-dimensionalism, then Ariel's preferences are irrelevant to the explanation, so this doesn't count as knowing anything about what it is to be *F* in the case at hand.

[26] As Shamik Dasgupta suggests, knowledge by acquaintance is "a direct way to identify something by its nature, a way that does *not* involve" knowing a full *account* of what it is ("Inexpressible Ignorance", p. 463).

I think the intuitive worry should be taken seriously, and the obvious response is inadequate. Mere knowledge by acquaintance, if it didn't confer any articulate propositional knowledge, would not give us sufficient understanding of our posits—they would remain unintelligible. It's beyond the scope of this paper to address the general notion of intelligibility. Instead, I'll focus on a particular necessary condition on intelligibility and explain why it's bad to posit things which don't meet this condition.

Some people are forthright about positing epistemically brute differences ("every analysis needs to start somewhere!"). But most people don't *explicitly* posit epistemically brute differences. In the later sections of the paper I'll explain why more sophisticated posits—posits which invoke fundamentality, essence, ground, or partial explanation by theoretical role—still rely on epistemically brute differences.

2.3.1 Positing distinctions without a difference

Suppose you see a chair, you close your eyes, open them, and again you see a chair. Suppose you form the belief: I saw two distinct chairs. Why do you think you saw two distinct chairs?

In typical cases, if you think you've seen two objects rather than one object, it's because you think you've seen some difference between the two objects. For instance, you think there are two distinct chairs because you think one of the chairs you saw was a rocking chair and the other was not a rocking chair.

Of course, even if you can't tell the two experiences apart, it's still possible that you did see two distinct chairs. But you probably shouldn't *believe* that there are two distinct chairs unless you think that you have experienced some difference. Thus: don't posit a distinction between the chairs you experienced while also claiming there was no qualitative difference between the two experiences!

2.3.2 Why articulation?

Let's grant that it's bad to posit a metaphysical explainer while claiming that you have *absolutely* no idea what it is that you're positing. Don't say that the dispute between three- and four-dimensionalists is about whether objects are *foo* while disclaiming any idea whatsoever about what "foo" means!

But no one—as far as I know—says that. Here's what people might say instead:

> I do know what "foo" means—it means *foo*! Can I explain this in more basic terms? Maybe not, but every explanation has to start somewhere.

In the case of metaphysical posits, I don't think this response is satisfactory. For such metaphysical posits, I think you need to be able to *articulate* something about the nature of what you're positing.

To see why, let's consider another example:

> You see a chair, you close your eyes, open them, and again you see a chair. The second appearance is qualitatively different from the first. Should you believe that there are two distinct chairs?

The answer to this questions depends on the particular way in which the appearances were different. If the first appearance presented something shaped like a rocking chair and the second presented something shaped like a desk chair, then you should believe that you saw two chairs rather than one.

But suppose that when you open your eyes, you see a chair that looks brighter than what you saw before. Suppose that it looks exactly like the chair you saw before would look if someone turned the light on. Or, suppose you are presented with a different chair shape, but it's exactly the shape that you would see if the chair you saw before were rotated.

In these cases, you don't have a reason to posit distinct chairs. These differences in appearance can be fully explained by differences in things other than the chair: the lighting and the chair's position relative to your eyes.

If we know that there's a difference in our two experiences but we can't articulate the difference, why would we believe that this is a difference in the *object* of the experience—the chair, in this case—rather than a difference in other causes or elements of the experience? If all we have to go on is the appearances, then—I say—we shouldn't believe that there are two chairs.

Likewise, suppose you say that the dispute between three- and four-dimensionalists is about whether persisting objects are *foo*. Fair enough, I say, but what sort of difference is that? Is it a difference in the nature of persistence itself, or is it a difference in how we think about or experience persistence? When we can't articulate anything specific about the difference—we just know that it's *foo*—I don't see how we could be justified in believing that there are distinct ways for the nature of persistence to be.

But there's a worry for my argument. There are lots of familiar cases in which it's plausible that we are justified in positing distinctions when we can't articulate the difference. If we can be justified in positing distinctions without articulation in these cases, why not in the case of metaphysical posits?

As far as I can tell, there are two kinds of cases to be addressed: ordinary perceptual experiences of objects in the external world, and experiences of qualia—things like tastes and flavors, colors, pleasures and pains. My response to why articulation isn't required in these two kinds of cases will be different. Let's consider each kind in turn.

2.3.3 Perceptual faculties

We start with yet another instance of looking at chairs:

> You see a chair, you close your eyes, open them, and again you see a chair. *Something* looks different. But you have no idea how exactly the two experiences differed. In fact, let's suppose that you are unable to discern *any* particular difference between the two experiences. Still, it seems to you that you saw two distinct chairs.

Are you justified in believing that there are two chairs, even though you can't articulate the difference that you saw? And if so, how are you justified in believing that the difference in your experience indicates a difference in the chair rather than a difference in other causes or elements of the experience?

I think it's plausible that you are so justified. Why? Because your visual system is able to detect differences and "tell" you that there are distinct objects even when you yourself aren't aware of the features which your visual system is tracking. And your visual system is very good at determining whether changes in appearance can be explained by changes in lighting or relative position. Your visual system doesn't give you all the data it uses to determine this, but it does give you reliable beliefs about whether you are seeing two objects or one.

To vastly oversimplify matters, we can think of the perceptual faculty as a source of something like testimonial knowledge. Of course you can have testimonial knowledge that there are two chairs without having any idea about the differences between the chairs. Likewise, you can have knowledge of distinctness derived from your visual system without having any idea about the differences. When something or someone *intervenes* in this way between you and the basic evidence, then it's easy to explain how you could be justified in positing a distinction without any ability to articulate the difference.

In the case of metaphysical posits, I don't find it plausible that there's anything standing between us and the data. There's no subpersonal faculty which determines whether we should posit a distinction between three- and four-dimensionalism. Thus, setting aside testimonial evidence, if we're justified in believing that we should make a metaphysical posit, this must be on the basis of the data that's available to us.

But there's a further worry for my argument. There's a different kind of case—qualia—in which it's plausible that there's no intervening faculty to explain how we can have knowledge of distinctness without the ability to articulate differences. Let's now consider the case of qualia.

2.3.4 Qualia

What are qualia? They are—supposedly—the particular qualities which can be directly present to the mind: canary yellow, the flavor of turmeric, or that indescribable sensation in your ankle. Or, qualia are particular qualitative experiences which can be directly present to the mind: the experience as of canary yellow, the experience as of the flavor of turmeric, or the experience as of having a certain sensation in your ankle.

Consider the following:

> You see a patch of blue and a patch of red. You can tell—just by thinking about the experience of each color—that these are distinct colors. But you have no idea how to articulate the specific differences between the two colors. It seems to you that they are different without being different in particular, articulable respects.

In the same way, you might think that you know that various flavors and smells or pleasures and pains have distinct natures even if you can't articulate more specific features which are different. You know that the experience as of sweetness and the experience as of sourness are distinct experiences, even if you don't have the words or concepts to articulate the difference.[27]

For our purposes, the important thing about qualia is that it's not outlandish to think they are presented *directly* to the mind—there's no intervening faculty. And it's not outlandish to think we can tell that two qualia have distinct natures even when we can't articulate or conceptualize particular respects in which they are different. So, we have another example where articulation isn't required for us to be justified in believing that there's a distinction, and there's no intervening faculty to explain why articulation isn't required.

Here's why articulation isn't required for qualia. Recall, articulation was supposed to help us distinguish two kinds of experiential differences: differences in the object of the experience, and difference in other causes or elements of the experience. More concisely, articulation enables us to distinguish mere differences in appearance from differences in reality.

[27] As Mark Johnston puts something like this idea, "The intrinsic nature of canary yellow is fully revealed by a standard visual experience as of a canary yellow thing" ("How to Speak of the Colors", p. 223).

The reason why articulation isn't required in the case of qualia is that it's plausible that there's no appearance/reality distinction when it comes to our experiences of qualia. Let me explain.

If you think that qualia are presented directly to the mind, and if you think that we can just tell that we're experiencing distinct natures, then it's also natural to think that part of what we mean by "qualia" is to be the kind of thing which can't be experienced in qualitatively distinct ways. You can't have the very same taste which tastes two different ways—if it tastes differently, then it's a different taste. Likewise, you can't have the same color qualia which appears in two different qualitatively different ways.

When we had two chair-experiences, we needed to know if this difference between the experiences corresponded to a difference in the object—i.e., the chair—or just a difference in things other than the chair. This depends on the particular respects in which the chair-experiences were different. But when we have two qualia-experiences, we know that they correspond to two distinct qualia regardless of the particular respects of difference.

Put another way, there's no gap for qualia between how they appear and how they really are, because what they really are just is their appearances. That's why articulation isn't required for us to have a justified belief that the difference in our experience reveals a distinction between natures rather than appearances of these natures.

In the case of metaphysical posits, I don't find it plausible that there's no gap between appearance and reality. It's perfectly sensible to think that there are differences in our *experience* of persistence—how persistence presents itself to us—which don't correspond to differences in the *nature* of persistence. Thus, if we're justified in believing that we should make a metaphysical posit, this must be on the basis of experience which reveals a difference in the nature of the object rather than a difference in how the object appears to us. And I don't see how we could be directly aware of this difference without being able to articulate any specific features of the thing we're positing to explain the distinction.

2.3.5 When is articulation required?

The only cases in which you can be justified in positing a distinction which you can't articulate are when (i) there's some intervening faculty—as in the case of visual perception—or (ii) there's no appearance/reality distinction—as in the case of qualia. In the case of metaphysical posits, neither of the conditions obtain: metaphysical posits are not known through some intervening faculty, and they are often introduced to explain the nature of something which could present diverse appearances. Thus,

> We shouldn't posit a metaphysical explainer if the resulting explanation would rely on an epistemic brute difference.

2.4 The persistence of brute differences

The bald-faced primitivist posits a brute difference. Brute differences are bad. But you don't have to be a bald-faced primitivist about the difference between three- and four-dimensionalism. You can use other metaphysical posits to draw the line. And, you can say things about our understanding of these metaphysical posits. Is that enough to show that we can use these posits without relying on epistemic brute differences?

The goal of this section is to argue that it's harder than you might think to avoid epistemic brute differences. To see this, we'll examine various moves the immodest metaphysician can make to explain the dispute between three- and four-dimensionalists. I'll argue that although the proposed posits aren't epistemically brute—since lots of things can be said to describe them—there's still something epistemically brute doing important explanatory work in the resulting explanations.

2.4.1 Wholly present

Here's one proposal: the three- and four-dimensionalist disagree about whether persisting objects are *wholly present* at each moment of time when they exist.

What does this mean? Very loosely, we might say of an extended event—a performance of "Jerusalem", for instance—that half of it is in the past and half is yet to occur. But we wouldn't say that part of the choir is in the past. Rather, the natural thought is that the whole choir is present at each time when it exists.

There are different ways in which to interpret talk of being "wholly present". Let us consider two proposals: explanation by analogy, and explanation by account.

2.4.1.1 Analogy

We might use talk of being "wholly present" to explain the notion of whole presence by analogy to universals. As Lewis suggests,

> Endurance corresponds to the way a universal, if there are such things, would be wholly present wherever and whenever it is instantiated.[28]

The idea is that universals are common constituents of many individuals—the roundness of this mug and the roundness of that mug are one and the same universals. As Plato taught us, a universal is *one over many*. It is wholly present in each individual which bears it, and not present in the way a sail is present over the heads of many sailors by having a bit of sail which is over each head.[29]

But this way of drawing the line between three- and four-dimensionalism cannot be the full story. It merely shifts the explanatory burden onto a different debate. There must be something in common which makes the two debates analogous. How can we cash this out?

2.4.1.2 An account

Rather than using talk of being "wholly present" as an analogy, we might cash out such talk in terms of the *exact location* of persisting objects.

[28] *Plurality*, p. 202.
[29] *Parmenides* 131c.

What is exact location? The exact spatial location of my desk is the entire region it fills. Although there are many subregions which my desk pervades and weakly occupies—it is weakly located at both the region under that stack of books and the region under my computer—it has only one exact spatial location (at least, at the present time).

Universals, on the other hand, can be distinguished from particular property instances by the fact that they can have multiple exact locations at a time.[30] Roundness is exactly located in the region occupied by each mug.

We can now cash out the analogy between universals and endurance:

> **Exact Location** Three-dimensionalism is the thesis that persisting objects are exactly located at each of the many three-dimensional regions occupied by their stages. Four-dimensionalism is the denial of this.[31]

But this proposal is unsatisfactory. Here's why. The four-dimensionalist can admit a relation—call it "exact stage location"—which allows her to mimic what the three-dimensionalist says about exact location.[32] In particular, the four-dimensionalist can say that some region is the *exact stage location* of an object just in case the three-dimensionalist says the region is the *exact location* of the object.

What is the difference between the relation which the four-dimensionalist calls "exact stage location" and that which the three-dimensionalist calls "exact location"? If these relations were the same, then a dispute about the exact location of persisting objects would be merely verbal. Thus, the present proposal relies on there being a difference between *exact location* and *exact stage location*. But what is this difference?

[30] The "multi-location" view of universals is often attributed to Armstrong—see, e.g., Kris McDaniel's "No paradox of multi-location"—although Armstrong himself is not entirely happy with talk of location (*Universals*, pp. 98–9). Mark Johnston notes in his discussion of Armstrong's "Identity Through Time" that the "motivating analogy for Armstrong is between enduring individuals wholly present from one moment to the next and universals wholly present in many *individuals*" (*Particulars and Persistence*, p. 58).

[31] For a worm-theorist like Lewis, the exact location will be the four-dimensional region occupied by the sum of stages. For a stage-theorist like Sider or Katherine Hawley, the exact location will be the three-dimensional region occupied by the present stage. In either case, they will say that the exact location is the single location occupied by the entire object.

[32] *Exact stage location* is the relation between objects and regions such that a region is the exact stage location of an object just in case it is the exact location of an instantaneous stage of the object.

The three-dimensionalist is free to respond that these are just different relations and there's nothing to be said about what constitutes the difference. But this would be to posit an epistemic brute difference, much like the three-dimensionalist who posits a brute difference between endurance and perdurance.

Let's consider what more the three-dimensionalist could say to articulate the nature of difference instead of positing it as an epistemic brute difference.

2.4.2 Fundamentality

How can we avoid saying that the dispute about "exact location" relies on a brute or shallow difference? One proposal is that it's a dispute about which relation is more *fundamental*.

The three-dimensionalist who takes the fundamental relation to be the *exact location* relation will say that fundamentally speaking persisting objects are multiply-located, while the four-dimensionalist will presumably deny this.

This is one way to use the notion of fundamentality to draw a line between three- and four-dimensionalism. Let me mention another way. We might distinguish the positions by which facts they take to be fundamental. On this proposal, the three-dimensionalist would say that facts about persisting objects are fundamental while the four-dimensionalist would deny this, and say that only facts about instantaneous temporal stages are fundamental.

Drawing the line in this way allows the three- and four-dimensionalist to agree about what there is—persisting objects, temporal stages, and so on. The disagreement would concern which of these are fundamental.

In order for these proposals to give us what we want, there are two features which fundamentality would need to have:

i. Fundamentality would need to be *sparse*: it's possible for there to be things which aren't fundamental.

ii. Fundamentality would need to be *deep*: whether or not something is fundamental is objective and doesn't depend on contingent facts about our psychology or language.

If fundamentality isn't sparse, then we can't use it to distinguish three- and four-dimensionalism in terms of what's fundamental. If fundamentality isn't deep, then disputes about what's fundamental won't be deep and substantive disputes.

2.4.2.1 Primitive fundamentality

Many of those who invoke fundamentality will say that surely here is a reasonable place to posit a brute difference. If anything is a fundamental feature of the world, unable to be cashed out in other terms, then fundamentality itself is a plausible candidate. As Jessica Wilson puts it, "the fundamental should not be metaphysically characterized in negative terms—or indeed, in *any* other terms. The fundamental is, well, *fundamental*."[33] Thus—you might think—it would be entirely appropriate to take fundamentality as *primitive*.[34]

But this would be to posit yet another epistemic brute difference: the difference between *being fundamental* and *being non-fundamental*. On this proposal, the difference between three- and four-dimensionalism would lie entirely in which things are taken to be fundamental. And what does that mean? Well, if it's a brute difference, then it's hard to say. I have argued that if we take ourselves to be in the position that the only way to distinguish three- and four-dimensionalism is to posit an epistemic brute difference, then we shouldn't take ourselves to have good reason for believing there's a difference which calls for this kind of explanation.

Of course, maybe fundamentality is primitive, and maybe there are two possibilities about the nature of persistence which can only be distinguished in terms of fundamentality. Nothing I have said demonstrates otherwise. But if we take ourselves to have good reason to think there are two possible positions about the nature of persistence, then such brutely distinguished positions can't

[33] "No Work for a Theory of Ground", p. 560.

[34] For another argument that fundamentality is a fundamental notion, see Sider, *Writing the Book of the World*, pp. 137–141.

be the two possibilities we have in mind since these brutely distinguished possibilities wouldn't be the ones that we have good reason to posit.

2.4.2.2 Non-primitive fundamentality

Instead of taking fundamentality as a primitive notion, we might try to explain the difference between being fundamental and being non-fundamental. We might explain it using other metaphysical posits, or we might explain it in terms of its theoretical role. Let's now consider these options.

I shall argue that these alternatives are no better than taking it as primitive. Or, more precisely, I don't think they are genuine alternatives to a primitive notion of fundamentality. In each case, we'll still need to posit a brute difference, and this brute difference will turn out to be no different than the brute difference between the fundamental and the non-fundamental.

2.4.3 Essential dependence

Let's consider the idea that we use the notion of *essential dependence* to distinguish three- and four-dimensionalism, rather than positing a primitive notion of fundamentality.

Here's one way to spell out this proposal. Recall our earlier discussion of temporal stages. In the context of that discussion I argued that both three- and four-dimensionalists could agree that persisting objects are not mereological sums of temporal stages.

But one might instead distinguish the views by what they say about the *essences* of the temporal stages. In particular, when the three-dimensionalist admits the existence of temporal stages, the three-dimensionalist is admitting stages as things which essentially depend upon persisting objects, while the stages considered by the four-dimensionalist are not taken to essentially depend upon persisting objects.

Consider, for instance, the momentary stage of the flower at precisely four o'clock. The four-dimensionalist thinks this is something which could exist even if there were no persisting flower.

The three-dimensionalist, on the other hand, can and should claim that part of what it is to be this very temporal stage is to be a stage of the persisting flower.

With such a notion of essential dependence in hand, it appears we have no need for a primitive notion of fundamentality. We can distinguish the three- and four-dimensionalism by asking: Are there momentary objects which essentially depend upon persisting objects? Or, are there momentary objects which do not essentially depend upon persisting objects?

2.4.3.1 An abundance of essences

But there's a problem with this proposal. Here's the problem. Like the three-dimensionalist, the four-dimensionalist can also grant that there's a momentary object which essentially depends on the persisting object. And, like the four-dimensionalist, the three-dimensionalist can also grant that for every moment at which a persisting object exists, there's a momentary object which doesn't essentially depend upon the persisting object. Thus, we haven't yet put our finger on the difference between three- and four-dimensionalism.

How can the three- and four-dimensionalist admit such entities? In both cases, objects with such essences arise because we can arbitrarily add or subtract (independent) features from an essence to get a new essence. For instance, if there's a lump of clay shaped into a human figure, then we can also grant the existence of something which is essentially shaped thus—a statue.[35] This allows the four-dimensionalist to admit temporal stages which essentially depend on the persisting objects of which they are stages.

And, going in the other direction, our earlier discussion of the Crown provides an example of subtracting features from an essence. The Crown was described by starting with particular individuals who happened to be monarch. We then removed from their essences the particular identities and peculiar features of the individuals—while essentializing the fact that they were the monarch—so that we ended up with an object which was essentially the monarch but not essentially any particular monarch.

[35]For further discussion of such coincident entities, see Lewis in "Counterparts of Persons and Their Bodies" and Fine in "Coincidence and Form".

Thus, the three- and four-dimensionalist alike can agree that there are momentary objects which essentially depend upon persisting objects, and there are momentary objects which do not essentially depend upon persisting objects. Where's the disagreement?

The obvious response is that these counterexamples are employing a different notion of essence. There's a "cheap" notion of essence on which every consistent set of features corresponds to an essence. These "cheap" essences correspond to a plenitude of entities: Socrates insofar as he is sitting, the thing which is embodied by Elizabeth II insofar as she is the monarch, and so on.

If we want to use the notion of essential dependence to draw the line between three- and four-dimensionalism, we must ignore such "cheap" essences and their corresponding abundant entities.

2.4.3.2 Back to fundamentality

We thus say: some essences are privileged—the "genuine" ones—and some aren't. But this leads to the question: what is the difference between a "cheap" and a "genuine" essence?

What's happened is that we have in effect just retreated to the distinction between the fundamental and the non-fundamental. Some essences are privileged, some aren't—whether we label them "genuine" or "fundamental" makes no difference so long as they are sparse and disputes about which things have them are taken to be substantive and deep. The shift to essence has not helped us avoid a brute difference between the sparse and the abundant.

2.4.4 Metaphysical explanation

One gambit would be to explain the difference between "cheap" and "genuine" essences. Can we do this without invoking yet another brute difference between the sparse and the abundant? If so, then we will thereby have a way to articulate the difference between being fundamental and being non-fundamental.

2.4.4.1 Essence and explanation

Let's consider one attempt to state the difference in terms of *metaphysical explanation*. The idea is to distinguish genuine and cheap essences by the kind of explanations which they make true.

An example to illustrate the difference. Suppose I have two children: Roger and Beatrice. A person is the kind of thing we usually assume has a genuine essence, and let's assume this now. Let's also assume, reasonably enough, that neither child essentially depends upon the other.

Suppose that there's also something which is essentially the *younger* sibling. Let's suppose that this is an entity with a "cheap" essence: it is Beatrice insofar as she is the younger sibling of Roger.

It's necessary that if something with this essence exists, then Roger also exists. But does its existence provide the *metaphysical* explanation for the existence of my first child? No—at least, assuming that this is a cheap essence. Rather, it's natural to say that the metaphysical explanation goes in the other direction. The fact that Roger exists is part of the metaphysical explanation for why Beatrice is the younger sister of Roger, and this fact in turn metaphysically explains the fact that the thing which is essentially a younger sibling exists.

On the other hand, it's natural to think that when we have something with a "genuine" essence, the existence of this thing is what metaphysically explains the truth of the propositions which figure in the essence. Suppose that it's part of the "genuine" essence of an electron that it attracts other electrons. In this case, it's natural to think that the fact that these two objects are electrons is what explains why they are attracting each other. It's not that they are electrons because they attract—as it would be if electrons only had "cheap" essences—but they attract because they are electrons.

We now have a way to distinguish "genuine" and "cheap" essences: genuine essences back metaphysical explanations, cheap essences don't. Or, more precisely, the existence of something which has a cheap essences does not metaphysically explain the facts which are essential to the thing, while the existence of something which has a genuine essence does.

Applying this to the dispute between three- and four-dimensionalists, we would distinguish the views by what they say about the metaphysical explanation of momentary objects. Both the

three- and four-dimensionalist can grant that there's some momentary object which is essentially a stage of a persisting object. But they will disagree about the relevant metaphysical explanations. As Fine puts it, the three-dimensionalist "will take the existence of a temporal part at a given time to be grounded in [i.e., metaphysically explained by] the existence of the persisting object at that time" while the four-dimensionalist reverses the order of metaphysical explanation.[36]

2.4.4.2 Brute explanation

Our task isn't just to give an account of the difference between three- and four-dimensionalism by positing some metaphysical explainer; it's to do so without relying on an epistemic brute difference, and in particular, without relying on a brute difference which amounts to the difference between being fundamental and being non-fundamental.

By now you should realize what comes next. I ask: by talking about metaphysical explanation, have we avoided positing yet another brute difference?

Unfortunately, we have not. The current proposal relies on a brute difference between *metaphysical* explanation and other types of explanations. Here's what I mean.

We can all grant that facts about instantaneous temporal parts can figure in an explanation of facts about persisting objects, and that facts about persisting objects can figure in an explanation of facts about instantaneous temporal parts. The dispute between three- and four-dimensionalists, on the current proposal, concerns which of these explanations is *metaphysical*.

Let's now consider this difference. There are many types of explanations; only some explanations are *metaphysical* explanations. Thus, the property of being a *metaphysical* explanation is sparse (just like fundamentality). And if the notion of metaphysical explanation is to help us give a satisfactory account of the difference between three- and four-dimensionalism, the difference between metaphysical explanation and other sorts of explanation must be deep and substantive (just like fundamentality).

[36] Fine, "Guide to Ground", p. 42.

Unless we say more about what makes something a *metaphysical* explanation, we have in effect—yet again—just retreated to positing a brute difference between the fundamental and the non-fundamental.

When we posit an epistemic brute difference between metaphysical explanation and other types of explanation, the main worry is that we don't understand metaphysical explanation well enough to believe that there is such a thing.

2.4.4.3 Familiar explanations

But—you may object—we aren't positing metaphysical explanation as some mysterious "I know not what". Even if we can't give an account of what it is for one thing to be metaphysically explained by another, metaphysical explanation is still perfectly familiar.[37]

Further, we can point to instances of metaphysical explanation.[38] It's relatively uncontroversial that the fact that I have more than five books on my desk is metaphysically explained by the fact that I have twenty seven books on my desk. If we can understand such instances of metaphysical explanation, what's the worry about invoking metaphysical explanation to distinguish three- and four-dimensionalism?

There's an important difference between these cases. In the relatively uncontroversial instances of metaphysical explanation, it's relatively uncontroversial *what explains what*. And it's relatively uncontroversial because we can understand why the one fact explains the other. Of course the fact that there are twenty seven books would explain why there are more than five books; there's no possibility in which the fact that there are twenty seven books fails to explain why there are more than five books. The *direction* of the explanation is fixed by what we can say about the relata of the explanatory relation (without having to explicitly say anything about the explanatory relation).

[37] As Gideon Rosen puts it, "we are often tempted to invoke the idioms of metaphysical dependence, which suggests that we often take ourselves to understand them" ("Metaphysical Dependence: Grounding and Reduction", p. 112).

[38] Ibid., pp. 110–3.

When we invoke metaphysical explanation to distinguish three- and four-dimensionalism, we are using a notion of metaphysical explanation for which it's an open question what explains what: three-dimensionalists say that the existence of a persisting object metaphysically explains the existence of the instantaneous temporal stages, while four-dimensionalists deny this. The direction of the explanation is not fixed merely by what we can say about the relata of the explanatory relation; it must be settled by explicit stipulation.[39]

The direction of familiar metaphysical explanations is already fixed; the direction in the case at hand would need to be settled by explicit stipulation. I think that this difference makes an important difference. In particular, because we would have to settle the direction by explicit stipulation, there's room for a further brute difference: the difference between the world in which the explanation goes in one direction and the world in which it goes in the other.[40]

Unless we say something about what this difference amounts to—unless we can articulate something about the difference between having and not having *this* type of metaphysical explanation going in one direction rather than the other—then we're just positing a difference which we don't understand well enough to know why we'd want to posit such a difference in the first place. Such metaphysical explanation is not the perfectly familiar and unobjectionable type of metaphysical explanation; it is the objectionable type of metaphysical posit.[41]

[39] Another example is the dispute about whether wholes are metaphysically explained by their parts or parts are metaphysically explained by wholes; see Jonathan Schaffer, "On What Grounds What", p. 363. It's also worth noting that Fine and Schaffer both suggested that one job for their notion of "ground" is to stipulate the direction of explanation (Wilson reports that they press this as an objection to her; see "No Work", p. 564, n. 74).

[40] I think that this difference is connected with the dispute about what grounds the grounding facts—i.e., what metaphysically explains the fact that *p* metaphysically explains *q*? Those like Karen Bennett and Louis deRosset who say that *p* itself metaphysically explains the fact that *p* explains *q* needn't regard the direction of the explanation as standing in need of further explanation (Bennett, *Making Things Up*, ch. 7; deRosset, "Grounding Explanations"). I plan to say more about this in future work.

[41] Thus, I am happy to concede much of what Rosen says about the intelligibility of a notion of metaphysical explanation–what he calls "grounding". But I deny that this intelligible notion is able to do the metaphysical work to which he and many others want to put it. Along the same lines, Thomas Hofweber argues that using examples to show that there's a "primitive, but nonetheless egalitarian, notion" of metaphysical explanation doesn't work because "the cases generally rely on trading a substantial notion of priority for one of several insubstantial ones" (*Ontology and the Ambitions of Metaphysics*, p. 341).

2.4.5 Theoretical roles

We've now considered various attempts to explain fundamentality using other metaphysical posits. The problem with these explanations is that each relies on an epistemic brute difference—and not just any brute difference, but the very same difference that we were hoping to explain.

There's another strategy for attempting to explain what we mean by "fundamental". This strategy is the one adopted by Lewis and Sider. Instead of explaining fundamentality by positing yet more metaphysical explainers, they try to explain fundamentality by specifying the *theoretical role* which this metaphysical posit is intended to play. Let me explain.

The model for this strategy comes from the philosophy of science. Consider the term "mass". How does this term get its meaning? On one popular view, the term "mass" can be interpreted as picking out that property, whatever it is, which (best) satisfies the description which we get by replacing the term "mass" with a variable in our theory of physics.[42]

Supposing Newtonian physics were correct, we might understand "mass" as picking out whichever property satisfies the laws of Newtonian physics—roughly, mass is the unique *F* such (i) any body with *F* will change its velocity when a force is exerted on it in proportion to how much *F* it has and in proportion to the size of the force, and (ii) any object with *F* attracts every other body with *F* in proportion to how much *F* they have and inversely in proportion to the square of the distance between them.

And of course, what goes for "mass" goes more generally. What Lewis and Sider both propose is that we can understand a metaphysical term like "fundamental" by specifying the role that fundamentality plays in our theory, and how it relates to other notions.

However, because the metaphysical posit we're considering is supposed to be something new and different from what we were able to talk about in more familiar terms, we can't interpret the theoretical role as giving a full analysis or definition of the thing which satisfies the role. Otherwise, there would be no need to posit fundamentality in the first place. If we can already talk about

[42]For the general idea of defining theoretical terms using their role in the theory, see Lewis "How To Define Theoretical Terms". For the theoretical role of fundamentality, see his "New Work". See also Sider, *Writing*, p. 9.

being a regular quadrilateral, there's nothing which would be better explained were we to posit the notion of being a square. And if we can already define fundamentality, then there's no need to posit fundamentality as a metaphysical explainer.

How might one deny that what it is to be fundamentality *just is* to be the thing which plays such and such role (specified in familiar and modest terms), while still claiming that this role does explain the notion of fundamentality?

I see two basic strategies. But in each case, there's still a brute difference. To see this, let's consider an instance of each strategy.

2.4.5.1 Restricting the candidates

One strategy is to implicitly restrict the candidates for satisfying the role. The immodest metaphysician says that the notion being explained is not simply *whatever* plays the role; her theoretical term picks out the thing which plays the role in the right way.

For instance, she might say her term picks out the *best* candidate for playing the role—where "best" here means something not definable in ordinary terms.[43] Or, it is the candidate which *explains* the role—where "explains" again means something not definable in ordinary terms. For instance, if part of the role of fundamentality is that it's easier to form beliefs about more fundamental things, then we might add that fundamentality is what *explains* why it's easier to form beliefs about these things.

In a nutshell, fundamentality isn't *whatever* satisfies the description given in the role; it's the most *fundamental* thing which satisfies the description.

2.4.5.2 Relying on contested notions

Another strategy is to explain fundamentality in terms of *contested* descriptions. These are descriptions which might pick out fundamentality—but only if certain unspecified facts obtain and if we interpret the description in the right way.

[43]So one might read Lewis' proposal in "Putnam's Paradox" that *naturalness* (fundamentality) is a constraint on reference which isn't "just more theory".

Suppose, for instance, that we define fundamentality in terms of the aim of physics.[44] On this proposal we would say—and this claim may even be true—that part of the aim of physics is to discover truths about the objectively privileged structure of reality. Can we use this fact about physics to describe what we mean by "fundamental"?

Perhaps. But—unfortunately—this still doesn't help us understand what we mean by "fundamental" well enough for us to know that we should posit such a thing. Even if physics does in fact aim at discovering fundamental truths, this would only help us understand fundamentality if we could distinguish physics from something just like physics except in this crucial respect.

Clearly, physics aims at discovering an especially interesting class of truths. But can we take it for granted that these are the *fundamental* truths, rather than truths which are interesting in some less metaphysically contentious way? In other words, can we take it for granted that giving a full account of what it is to be physics would require us to posit fundamentality?

We're trying to explain what it is to be fundamental, but to do so, we already need to understand how physics is different from faux-physics which doesn't aim at discovering fundamental truths. The difference between physics and faux-physics turns on the notion of fundamentality. Thus, when we single out physics from faux-physics in our explanation of fundamentality, we're thereby relying on the very difference which we were trying to explain.

2.4.5.3 Brute differences remain

On each way of using the theoretical role to explain fundamentality while also avoiding a reduction of fundamentality to what can be said in more familiar terms, we fail to completely pin down what we mean by fundamentality in a way that's sufficiently intelligible to someone who doesn't already understand fundamentality.

Thus, all of these ways of trying to explain the nature of fundamentality leave a gap between what the immodest metaphysician wants to posit and what we're able to understand.

[44] Sider claims that fundamentality, like truth, is "a constitutive aim of the practice of forming beliefs" (*Writing*, p. 61).

The explanatory gap may be explicit—there's just an extra ingredient of reality which we must take as primitive. Or the gap may be implicit—we can explain fundamentality only by relying on the notion in our explanation. Either way our explanation of fundamentality leave some residual difference which is completely unexplained—a brute difference. And it is precisely this residual difference which we rely on when we try to explain what the deep and substantive difference between three- and four-dimensionalism is supposed to be.

In a nutshell: once we start positing new notions, it's hard to satisfy **Articulation**!

2.5 Immodesty without brute differences

At this point, things may look bad for run-of-the-mill immodest metaphysics. I have argued that there is a serious challenge for **Immodesty**, and that the typical moves made by immodest metaphysicians are inadequate to address this challenge.

Some of my readers will surely take this challenge as (yet another) reason to reject **Immodesty**. But that's not how I take the challenge. I take it as a reason to look beyond the run-of-the-mill. If we're going to endorse **Immodesty**, then we need to look for a new kind of immodest view—one which doesn't rely on epistemic brute differences.

In this concluding section of the paper, I shall describe the kind of view I have in mind. It's beyond the scope of this chapter to show how the view I describe really does avoid epistemic brute differences; this is a task I take up in chapter 3. Still, I think it's best to end on a hopeful (albeit inconclusive!) note.

To motivate my new kind of view, let me first sketch a more general explanation for why our earlier attempts to avoid brute differences didn't work. Then, I'll describe a view which—I think—doesn't do the thing which got other proposals into trouble. Finally, I'll show that we could use the view to distinguish three- and four-dimensionalism.

2.5.1 Partial explanation

When we posit a metaphysical explainer—as **Immodesty** says we should—then we are positing a difference. The thing we posit must be different from the things that we fully describe without the posit. Otherwise, our posit wouldn't be a *new* metaphysical explainer.

When we posit something which—as far we can tell—is *entirely* different from familiar things, then we have posited an epistemic brute difference. This is what **Articulation** forbids. We have considered various instances of what **Articulation** forbids: to posit an unexplained distinction between the fundamental and the non-fundamental, an unexplained distinction between "cheap" and "genuine" essences, or a brute notion of metaphysical explanation which can go in different directions. Because we couldn't explain these posits in terms we already understood, they resulted in epistemic brute differences.

To avoid positing a brute difference, we would need to posit something which is partially explained by familiar things—it *is* explained, but it's not *fully* explained. Unless it is *at least* partially explained, it violates **Articulation**. Unless it is *at most* partially explained, it is not an instance of **Immodesty**.

We considered one way to partially explain our posits: characterize them using their theoretical roles. When we posit something as the satisfier of a theoretical role stated in modest and familiar terms, the theoretical role tells us how our posit is similar to what we already understand.

But—unfortunately—this strategy still led to a brute difference. The problem was that **Immodesty** required us to say that the posit was not simply *whatever* satisfied some particular role stated in perfectly precise and familiar terms. Either the posit is the *privileged* satisfier of the role, or it's what satisfies the role when contested notions in the role are *suitably interpreted*. In both cases, these extra conditions are ways of saying how the posit is *different* from other things which might satisfy the role—i.e., why it isn't *fully* explained by the theoretical role. And, this difference is in effect precisely the difference that we want explained.

In a nutshell: we can't reconcile **Immodesty** with **Articulation** unless we find a way to say "but it's not *fully* explained" about our posit without just introducing an epistemic brute difference.

2.5.2 Three- and four-dimensionalism

Here's one way to think about the type of posit which would allow us to reconcile **Immodesty** and **Articulation**.

Consider some changing object: a leaf changing from green to brown. Everyone can admit that the leaf is green at one time and brown at another. The four-dimensionalist thinks that these kind of facts about what qualities things have at distinct times will give us the full story of persistence; there's just a bunch of discrete events.

But the three-dimensionalist thinks there's something more than this. The intuitive idea behind three-dimensionalism is that there's something unified which runs through this complex sequence of events. We considered various attempts to cash out this intuitive idea—talk of identity in the strict and numerical sense, talk of fundamentality or "genuine" essences, and so on.

I have argued that these attempts to posit something unified are unsatisfactory. They are unsatisfactory because they ultimately give us no handle on the difference between the sense in which the events are supposed to be unified and the kinds of unity which can be admitted by the four-dimensionalist. They simply amount to a stipulation that there is some unification which is different from the kinds of unity which the four-dimensionalist can fully explain. And—I have argued—if this is the best we can do to explain the posit, then we aren't epistemically justified in positing such a thing.

But what else could the three-dimensionalist add to the neutral description of change as a sequence of events events? The three-dimensionalist must add something which makes a difference from how the four-dimensionalist understands the particular events, and does so without simply stipulating a difference from what can be fully explained by the four-dimensionalist.

Let me use an analogy to show the sort of thing the three-dimensionalist could propose in order to do this. Imagine that you're reading a novel. When you read a novel, you don't read the chapters as if they were describing entirely separate and distinct events. You read them in light of each other. You think the chapters can be understood by their place in the overall narrative. In fact, you may think that the reason why certain things happen in this chapter is related to what

happens both in early and later chapters. The events in the book occur, in part, because of the way they fit together into a larger narrative.

In a similar way, we can think of the three-dimensionalist as claiming that the events in the life of a persisting object are to be seen *as* episodes in the life of a persisting object, and not as completely discrete events. There's something which unifies the events, and isn't just the sum of the events.

Of course, the four-dimensionalist can say that the events stand in causal relations to each other. But the three-dimensionalist says they're unified in a more robust sense than this.

In the case of a novel, this unifying explanation can be found in the artist's vision of the work. Pity and mercy kept Bilbo from killing Gollum when he had a chance in *The Hobbit*, but his action is also partially explained by the fact that there are things for Gollum to do in *The Lord of the Rings*.[45]

In the case of a persisting person, here's one proposal of the sort of thing a three-dimensionalist could posit. She might say that the unifying explanation can be found in *who* the person is—their "personality", if you like. If we posit a unifying explanation of a person's actions, then we would deny that the person is just the sum of their particular action (both the actions which they actually do and the actions which they would have done in various circumstances). Yet we still see glimpses of who they are in what they do; we see the actions as expressions of some imperfectly revealed personality.

Of course, the three-dimensionalist doesn't have to posit personalities as the unifying explanations of our actions. But I think this possibility nevertheless helps us to see how, in principle, there could be a way of making sense of the difference between three- and four-dimensionalism without relying on an epistemic brute difference.

[45] As Gandalf says: "I have not much hope that Gollum can be cured before he dies, but there is a chance of it. And he is bound up with the fate of the Ring. My heart tells me that he has some part to play yet, for good or ill, before the end; and when that comes, the pity of Bilbo may rule the fate of many—yours not least" (*The Fellowship of the Ring*, book one, chapter ii).

It's harder to say what this unifying explanation would be in the case of non-living objects we encounter in the external world. And even if we can get a grasp on what such a thing would be, we haven't yet considered any reason to believe that there actually are such unifying explainers.[46]

That said, this analogy gives us a way to make sense of the difference between three- and four-dimensionalism without relying on an epistemic brute difference. The thing that we're positing isn't *entirely* unfamiliar. We're familiar with sequences of events; three- and four-dimensionalist alike can agree that the leaf is green at one time and brown at another. And, the unifying explainers posited by the three-dimensionalist are partially similar to the sequences of events which they are supposed to unify.

Further, the relevant notion of partial similarity can itself be partially explained by analogy. We're familiar with the way in which the narrative of a novel is partially explained by the sequence of individual events in the narrative. And—I suggest—the three-dimensionalist's unifying explainer stands in an analogous relation: the relation between the unifying explainer posited by the three-dimensionalist and the sequence of events which it's posited to explain is analogous to the relation between the episodes and the narrative in a novel.

Now, I don't think that the only available interpretation of these analogies is incompatible with four-dimensionalism. The four-dimensionalist could offer a more deflationary interpretation of these analogies, explaining them in a way that's perfectly consistent with four-dimensionalism. But this deflationary interpretation isn't the only one available to us. Even if these analogies don't explicitly pin down the difference between three- and four-dimensionalism, the analogies can point toward the sort of thing that the three-dimensionalist could posit to distinguish her view from four-dimensionalism.

[46]Perhaps three-dimensionalism makes the most sense if you're already a theist who thinks that there is a being who created the universe with the whole story already in mind. For example, a central idea of the Christian interpretation of history is that certain episodes in the past should should be interpreted in light of the future. The Gospel of Luke depicts the risen Messiah practicing this method of hermeneutics: "And beginning with Moses and all the Prophets, he interpreted to them in all the Scriptures the things concerning himself" (Luke 24:27 in the *English Standard Version*). Of course, I haven't argued that there's no way for the three-dimensionalist to explain this unity without positing a divine author; it's just harder for me to imagine what this unifying explanation would be.

What's important is that, given this explanation of the posit, there's no need to add that this isn't a full explanation. In particular, analogies clearly don't give us a way to *define* our posit in terms of familiar things; explanations by analogies are always *partial*.

The key gambit for harmonizing **Immodesty** with **Articulation** is to say that that our posits are supposed to be the kinds of thing which can only be explained by such analogies; there's no more basic way to cash out the relation between the familiar facts and the posited explainer.

At this stage, you might worry that the moves I've just made to show that we can understand the difference between three- and four-dimensionalism could also be used to show we understand the posits I rejected earlier. Can we use analogies like this to explain the proposed posits we considered before—i.e., the posits which were rejected as introducing epistemic brute differences?

Let me say first: if someone wants to use "fundamental" for the sort of unifying explainer which these analogies are partially explaining, I have no objection. But I think that there's an important difference between the posits I have just introduced and the way in which fundamentality (and other problematic posits) are typically understood.

The unifying explanation here posited by the three-dimensionalist is not something extra simply added to our description of the sequence of events. It is instead supposed to show up as a difference in the events themselves, much as the larger narrative makes a difference in what is happening in each chapter.[47] And I think that when we posit this kind of unity, it's not the case that there must be some epistemic brute difference doing explanatory work.

There's still a residual worry that this proposal isn't as radical as I suggest, and the typical immodest posits can be explained through analogies. Much more needs to be said about how such explanations work, and why they are as radical as I claim. I shall address these matters in more detail in the next chapter.

[47] Against this, one might say that the larger narrative doesn't change what's happening in each chapter—the events are the same—but it adds something extra which we might call "significance". But I think it's plausible that there isn't a sharp line to be drawn between "what's happening" and its "significance". In the same way, I want to push back on the idea that there's a sharp line between each event in the life of a persisting object considered in isolation, and the event understood in light of something larger (e.g., the personality partially revealed in the event).

What has been argued thus far? I have argued that there's a problem with the typical immodest posits used to explain the difference between three- and four-dimensionalism. And, I have argued that there's a gambit we must make if we are to avoid this problem. In a nutshell, the gambit is to take analogies seriously as giving us a basic form of metaphysical explanation—a form of explanation which is irreducibly partial.

If this gambit works, we now have a less problematic account of two possible positions about the nature of persistence. Let me note: it's unclear to me whether we should endorse three- or four-dimensionalism, so construed. It's not obvious how we should settle the dispute. But I think this is an acceptable result. Deep and substantive disputes can be difficult to settle.

Of course, this all turns on a big *if*: the gambit allows us to distinguish three- and four-dimensionalism only if we can indeed make sense of these posits without any epistemic brute differences, and for that, we need to be able to make sense of these metaphysical analogies. But at any rate, I hope to have motivated the idea that *if* there's metaphysical immodesty to be had, it will be different from the kinds of immodesty posited by views that we're currently familiar with.

Chapter 3

Holism and Metaphysical Explanation

In previous chapters, I have defended two theses about the nature of metaphysics:

> **Immodesty** There are things we can talk about only by positing them as metaphysical explainers.
>
> **Articulation** We shouldn't posit a metaphysical explainer if the resulting explanation would rely on an epistemic brute difference.

The task of the current chapter is to explore the metaphysical picture which results when we put these two claims together. I'll argue that if we want to harmonize **Immodesty** with **Articulation**, we'll need to reject certain dichotomies. We'll start with the dichotomy which can be expressed as

> **Atomism** There are no irreducibly partial explanations.

This thesis states a dichotomy about explanation. If **Atomism** is true, then everything is either fully explained or fully unexplained—or, it can be divided into logical parts such that each part is either fully explained or fully unexplained.

I shall argue that unless we reject this dichotomy, it will be impossible to posit something which is both new—thus immodest—and yet doesn't introduce any epistemic brute differences—thus articulate.

It's one thing to say that there could be an explanation which is irreducibly partial, it's another to provide such explanations or to give a positive account of them. I'll argue that the best way to make sense of irreducibly partial explanation is to reject another dichotomy: that between things which are logically complex—they have multiple, logically distinct parts or constituents or properties—and things which are logically simple.

What we get once we reject these dichotomies is a kind of metaphysical explainer which doesn't fit into our familiar categories. It is the kind of thing which I call a *holistic* metaphysical explainer. And it's the kind of things which can reconcile **Immodesty** with **Articulation**—at least, it can if anything can.

Thus, I'll argue that given that we have reason to accept both **Immodesty** and **Articulation**, we should also endorse

Holism There are holistic metaphysical explainers.

3.1 Preliminary clarifications

Before we proceed to the defense of **Holism**, there are a few respects in which it's important to clarify the scope of the project.

3.1.1 Modesty about our immodesty

A moment ago, I said that holistic explainers are the kind of things which can reconcile **Immodesty** with **Articulation** *if anything can*. Let me emphasize this caveat. I won't try to *prove* that it's possible to posit things which are both immodest and articulate. My aim is something more modest.

My aim is to show *how* it's possible—supposing that it is—for holistic metaphysical explainers to be posited without introducing brute differences. The kinds of reasons I provide will not be dialectically effective against the skeptic who denies this possibility. But they are the kinds of reasons which enable me to explain how my own belief is rational.

I think that such arguments are the best we can give when we're addressing questions about the nature of metaphysical explanation. There just isn't enough that we can take for granted as a neutral starting point; too much is up for grabs.[1]

It is important to note that just as I think there's no proof that **Holism** allows us to harmonize **Immodesty** and **Articulation**, I also think there's no proof against this claim. Neither side has a decisive argument; the arguments on both sides will have an air of circularity.

Sometimes those who argue against metaphysical realism are forthright on this point. For instance, after presenting his "argument" against (heavyweight) metaphysical realism, David Chalmers concedes that

> the main intuition in this argument is very close to a simple denial of heavyweight realism…. But still, it is an intuition that many people (especially outside the field of ontology) will share.[2]

More generally, Richard Rorty in *Philosophy and the Mirror of Nature* has suggested that there's a connection between his antirealism about metaphysics and his antifoundationalism about epistemology: these both arise from his rejection of the "Cartesian" picture of the mind and its relation to the world.[3]

While I don't endorse the picture that Rorty would put in its place, I do think he's right to reject this "Cartesian" picture and its commitment "to the construction of a permanent, neutral framework for inquiry… which can be isolated prior to the conclusion of inquiry".[4] And I agree with him that once we reject the picture, we both give up a certain sort of metaphysical realism and a certain sort of epistemic foundationalism. In particular, I think he's right that we should not be looking for probative reasons in this domain.

The primary focus in this paper is not our epistemic limits. But I do want to emphasize this limit so that the reader knows what to expect in the present paper. Don't expect a *proof* that **Holism**

[1] I have a bit more to say about this below; I plan to develop the claim more carefully in future work.

[2] "Ontological Antirealism" in *Metametaphysics*, p. 103.

[3] Let me note: I don't think Descartes endorses the so-called "Cartesian" picture. But this is irrelevant to our present concerns.

[4] *Mirror*, p. 8.

allows us to harmonize **Immodesty** with **Articulation**! My more modest aim in this paper is to explain why it's *perfectly reasonable* to believe that it does.

3.1.2 Epistemic and metaphysical bruteness

Something is an epistemic brute difference *for you* when you have a label for the difference—you have the word "fundamental", for instance—but you can't cash out any of the features of fundamentality (or whatever else you posited) which help it do the explanatory work that it was posited to do.

Now, we can imagine types of things which couldn't be posited without introducing brute differences. These would be things that just don't have any explanatory features to be articulated—they are "basic" or "primitive" explainers.

For instance, suppose you're a primitivist about goodness who thinks that goodness doesn't have the kind of intrinsic nature which can be articulated into distinct features: it just is what it is. And suppose you think that the facts we're trying to explain have to be explained by the kind of thing that goodness itself is, and not by any facts about how it's related to other things—for instance, by the fact that it's something we desire. In this case, it seems like anyone who posits such a notion of goodness is bound to introduce an epistemic brute difference. The kind of explainer that goodness is supposed to be be in this case makes it impossible to satisfy **Articulation**.

When something does its explanatory work in a way that depends on such a primitive nature, it isn't just epistemically brute *for you*. The fact that the thing is epistemically brute for you isn't explained by some peculiar features of you and your relation to the thing. What explains the epistemic bruteness is a fact about the thing itself and the type of explanatory work which that thing is supposed to do. This is the kind of thing which *cannot* satisfy **Articulation**. Let's say that such things are *metaphysically* brute (as well as epistemically brute).

The concern in this paper is with *metaphysical* bruteness. The fact that something is epistemically brute *for you* may not tell us much about the thing itself; plausibly, anything can be posited in a way that makes it epistemically brute. The interesting question for us is whether it's *possible*—

given the kind of thing that we're positing and the kind of explanatory work that it's posited to do—to posit the thing in a way that doesn't introduce an epistemic brute difference.

Thus, my aim is to argue that holistic explainers *can* be posited without introducing epistemic brute differences. I will not show that whenever we posit something holistic, there isn't an epistemic brute difference.

The strategy of the paper is to argue that certain dichotomies entail that we *can't* posit immodest explainers without there being epistemic brute differences. Thus, in order to harmonize **Immodesty** with **Articulation** we would need to say there are things which violate these dichotomies. In other words, to say that it's possible for there to be such things, we need to reject certain dichotomies.

3.1.3 Brute differences and explanatory tasks

Just because something is an epistemic brute difference *for you* doesn't mean it has to be an epistemic brute difference for everyone. But there's another way in which epistemic bruteness is relative. Just because you're able to posit something without introducing an epistemic brute difference in one explanation doesn't mean that you're able to posit something—maybe even the same posit—without introducing an epistemic brute difference into the explanation of something else. Let me address the significance of such relativity to explanatory task.

Suppose that we're able to posit the sparse properties and distinguish them without any epistemic brute difference from arbitrary classes. Then, we might posit the property of *being good* to explain what it is for (e.g.) acts of giving to charity and groves of stately beeches to be objectively similar, and we can do this—by hypothesis—without introducing any epistemic brute difference into the explanation.

But if we want to explain what it is for giving to charity and stately beeches to be things that ought to be appreciated—we want to explain the distinctively *normative* similarity—it may be that our posit of *being good* does introduce an epistemic brute difference. This is because we may be able to articulate the difference between *being good* and being in some arbitrary class but not able

to articulate the difference between *being good* and various non-normative, objective respects of similarity—say, *being lime green*. The fact that something is lime green doesn't explain why it ought to be appreciated!

Even if we can't articulate the difference between *being good* and *being lime green* without relying on an epistemic brute difference, this epistemic brute difference is irrelevant to the explanation of why various good things are objectively similar. As long as you can articulate the difference between *being good* and an arbitrary class, then you can posit goodness to explain the objective similarity between two things, because the difference you can't articulate—the one between *being good* and *being lime green*—isn't relevant to the explanation.

But this difference is relevant to the explanation of why you ought to appreciate those things, and it's relevant to the explanation of what it is for two things to be *normatively* similar. This means that if you're trying to explain what it is for two things to be normatively similar, and you aren't able to articulate the difference between *being good* and *being lime green*, then positing goodness in this case would result in an explanation which relies on an epistemic brute difference. And this is what **Articulation** tells us not to do.

Let me emphasize, then, that the goal in this paper is to explain the possibility of harmonizing **Immodesty** with **Articulation**, and this only requires that *in some cases* we can posit an explainer so that the explanation doesn't rely on an epistemic brute difference. I won't try to address how (or whether) this is possible in every case.

In particular, we'll focus on a "minimal" case: positing something to explain the difference between reductive and non-reductive views about *predication*. To do this without relying on an epistemic brute difference, we'll have to be able to articulate the difference between what the non-reductionist posits and what can be constructed out of arbitrary classes of individuals. In other words, we'll have to be able to make sense of what it would be for there to be "something more" than the things the reductionist—an extreme nominalist in this case—will accept.

A much harder task would be to show how we can posit specific and distinctive respects of similarity and use them to explain, say, non-reductive views about the normative or the mental.

To do this, we would need to be able to articulate not just the *general* difference between a genuine respect of similarity and an arbitrary class, but also the *specific* difference between (say) goodness and other genuine but non-normative respects of similarity. I do believe that this can be done, but it is (far!) beyond the scope of the present paper.

To sum up: the task of this paper is *how* it could be possible *in general* to posit something new as an explainer without it being *necessary* that the explanation relies on an epistemic brute difference. The way to do this involves the rejection of various dichotomies; let's now turn to this.

3.2 Dichotomies and brute differences

A dichotomy is a division into two exclusive categories. In the simplest case, a predicate and its complement form a dichotomy—say, things that are blue and things that aren't, or things that are fundamental and things that aren't. But in other cases, we might have two categories which aren't simply defined as complements yet are still taken to form a dichotomy. These are *non-logical* dichotomies, and it's with such dichotomies or putative dichotomies that we shall be concerned.

In this section, my goal is to show how certain (non-logical) dichotomies would make it impossible to harmonize **Immodesty** and **Articulation**. But—I think—we can reject these dichotomies, and in doing so, we open up the possibility of positing immodest yet articulate explainers. And—I'll show—we can use the rejected dichotomy to help us characterize the nature of the things which would have been ruled out by the dichotomy.

Before I address the particular dichotomies and what we get by rejecting them, let me first say more about dichotomies and their rejection.

3.2.1 What does it mean to reject a dichotomy?

Descartes famously supposed that *being a thinking thing* and *being spatially extended* formed a non-logical dichotomy—exclusive, yet neither the simple negation of the other.[5] According to

[5]See *Principles of Philosophy*, I.53.

Descartes, there couldn't be something which was both a thinking thing and an extended thing, and there couldn't be something which was "kind of" a thinking thing and "kind of" an extended thing.[6]

To say there's a dichotomy is to say there's a sharp line dividing the two categories without any middle ground between them. When the categories aren't simply defined as complements, it can be a substantive question whether they amount to a dichotomy. Descartes believed that the mental and the physical formed a dichotomy. But lots of people—panpsychists, for instance—deny that there's a sharp divide between the mental and the physical.

Rejecting a dichotomy is very similar to rejecting some ideology—e.g., the ideology of fundamentality. The antirealist—someone like Hilary Putnam, Nelson Goodman, or Richard Rorty—claims that there simply is no such thing as fundamentality. By rejecting this notion, they are also rejecting the putative dichotomy between things that are fundamental and things that aren't.

In this paper, our concern is with rejecting *non-logical* dichotomies. Consider Descartes' dichotomy between thought and extension. When you reject this dichotomy, you aren't necessarily saying that there's no such thing as thought or extension. You might be happy to grant that there are thinking things and there are extended things. But if you deny that thought and extension form a dichotomy, what you're rejecting is Descartes' claim that there's no middle ground; you're saying that something could combine features of thought and extension.

Or, to take a simpler case, when you deny the dichotomy between black and white, you aren't denying that there's anything entirely black or anything entirely white. Rather, you're allowing the possibility that some things are in between—i.e., grey. Of course, in the cases of black and white or thought and extension, most of us aren't tempted by the idea that these form dichotomies. And in the case of black and white, we already have a familiar vocabulary for the things that are in between.

The dichotomies which are relevant to the project in this paper are dichotomies which are much harder to deny. They are dichotomies in which the exclusiveness really does seem to be built into

[6] At least, this is what textbook Descartes thinks. For our present purposes, it doesn't matter if this was Descartes' actual view.

our idea of the two categories, so it's hard to see how there could be something which cuts across the categories.

3.2.2 Why reject a dichotomy?

Let me illustrate the role that rejecting such a dichotomy can play. The illustration comes from Plato's *Meno*. Early in the dialogue, Socrates asks Meno to say what virtue is.[7] After various failed attempts Meno and Socrates both find themselves in a state of aporia: they thought they knew what virtue was, but now it seems that they don't know.[8] They are looking for an account of something—virtue—but it turns out they don't seem to know what it is that they're looking for. At this point, Meno raises an important challenge:

> How will you look for it, Socrates, when you do not know at all what it is? How will you aim to search for something you do not know at all? If you should meet with it, how will you know that this is the thing that you did not know?[9]

This challenge presupposes a dichotomy. Either you already have something—you know what it is—in which case there's no need to look for it, or you don't have it—you don't know what it is—in which case you wouldn't be able to take it as the object of your search. And Socrates responds precisely by rejecting this dichotomy. His solution is to posit something between knowledge and ignorance: memory.[10] When we are looking for something which is already in our memory, the object of inquiry isn't directly present to us—we don't already know what it is—but it isn't completely absent either.

Here there are two things to note. First, the paradox of inquiry presupposes a dichotomy which Socrates rejects. And second, Socrates does not merely reject the dichotomy, but presents a positive story—developed further in other dialogues—of the new epistemic state which we get by rejecting the dichotomy between knowledge and ignorance.

[7] *Meno* 71d.
[8] 80a–d.
[9] 80d (trans. G. M. A. Grube).
[10] 81d.

Let me emphasize: rejecting this kind of dichotomy is just as immodest as positing a new notion as an explainer. In fact, the rejection is immodest in exactly the same way: it says there's something which stands in need of explanation by positing new notions.

3.2.3 What replaces a dichotomy?

When you reject a dichotomy like this, part of what you're doing is to shift how you understand the categories in the binary cases. And, part of what you're doing when you reject a dichotomy like this is to introduce the possibility of a new kind of thing which cuts across the categories. Thus, when you reject a dichotomy, you also have to posit something new to put in its place—categories which are similar to the ones we started with, but without being exclusive.

But it's important that we aren't just positing new and entirely unfamiliar categories in place of the ones that we're rejecting. Our grasp on such categories would then involve an epistemic brute difference!

In this paper, my strategy is to use the exclusive categories of the rejected dichotomy in order to partially explain the new categories which we're putting in their place. In particular, by identifying the dichotomies which would make it impossible to harmonize **Immodesty** and **Articulation** we can use these rejected dichotomies to help us understand the kind of thing which is ruled out by these dichotomies—i.e., the kind of thing which is both immodest yet articulate.

What this means is that my method of showing how to harmonize **Immodesty** and **Articulation** itself depends on the possibility of harmonizing these two theses, because the method involves positing something new. This is one of the reasons why I cannot *prove* that it's possible: the proof would be circular because it would presuppose the thing we're trying to prove.

Nonetheless, I think the kind of account that we can give of this possibility is still valuable. First, I think it is both true and illuminating. Second, I think there really is a worry which this account—albeit circular—helps us to address: if we couldn't give a consistent account of the possibility of harmonizing **Immodesty** and **Articulation**, we would have good reason to believe that it's not

possible to harmonize them. And, while the account cannot prove *that* it's possible to harmonize the two claims, it does show us *how* it could be possible—at least, if it is in fact possible.

With this general discussion of dichotomies and their rejection in hand, let us turn now to the particular dichotomies which must be rejected.

3.2.4 Atomism and partial explanation

Once we've seen what **Immodesty** and **Articulation** require, it's fairly clear that any posited explainer that satisfies both is going to be *partially* explained. If it's *fully* explained by notions we already accept and understand, then it's not a *new* posit. And if it's not explained *at all* by such notions, then there would be a brute difference between the thing we posit and the things we already understand. Thus, the obvious first move if we want to harmonize **Immodesty** and **Articulation** is to posit something which is *partially* explained.

Let's consider the case of predication. On the reductive view, predicational facts are fully explained by logical facts about collections of individuals. Suppose that Alice and Bob are all the *F*'s. Then—according to the reductionist about predication—what it is for Alice and Bob to be *F* is fully explained by Alice and Bob (and logical facts, if we think these need to be included), since the fact that Alice and Bob are *F* is just the fact that Alice and Bob are both in the collection of Alice and Bob.[11]

On the non-reductive view of predication, the fact that two individuals have a genuine respect of similarity in common is not fully explained by those individuals (and logical facts). Predicational facts involve "something more" than what's fully explained by the reductionist's account of predication. Thus, while agreeing with the reductionist that Alice and Bob figure in explanation of the fact that Alice and Bob are *F*, the non-reductionist says that this is only a *partial* explanation.

[11]On more sophisticated reductive views, we might also need to include the fact that the collection of Alice and Bob is in some other collection of collections (if we think that *F* is a "sparse" property). Or, we might also need to include the fact that other possible collections are "counterparts" of this collection (if we want to say that other things could have been *F*). But we can ignore these complications, because whatever the sophisticated reductionist adds, the non-reductionist is still going to say there's something more to predication than this. Our concern is to explain that "something more".

Let me make two points clear. First, I do think this is what the non-reductionist should say, and I do think partial explanation is what we need in order to harmonize **Immodesty** and **Articulation**. But, second, this can't be the full story. In particular, we still need to show that we aren't introducing yet another brute difference with our talk of "partial" explanation.

It's easy enough to see how a partial explanation is different from no explanation at all. In the one case, we don't have anything explanatory that we can say; in the other case, we do have some things we can say.

What's harder is to make sense of the difference between partial explanation and full explanation. And this is harder because we have to make sure that we aren't just introducing yet another brute difference.

Suppose you say that there are just different types of explanation: reductive explanation is one type, and non-reductive explanation is another. More precisely, suppose you posit a "brute" or "primitive" notion of partial explanation. Such a notion of partial explanation is no help to our project; our explanation of the difference between reductive and non-reductive positions would then rely on a brute difference between reductive (full) and non-reductive (partial) explanation.

A more promising proposal is to say that partial explanation involves the same type of explanation as is given in a full explanation, but it isn't full because it doesn't explain everything that was to be explained. Partial explanations fall short of full explanations. More directly, if we can understand the idea that full explanations have parts, then we can understand a partial explanation as part of a full explanation.

I think many people—and here I include my former self—take it for granted that this is how we must understand partial explanation. For instance, a recent paper on grounding simple takes "partial ground to be *defined* in terms of full ground"—i.e., as part of a full ground.[12]

I think that this account of partial explanation is often taken for granted because we implicitly accept a dichotomy about explanation. We think that every thing is either fully explained or fully

[12]Jon Erling Litland, "Could the grounds's grounding the grounded ground the grounded?" (p. 3 of page proofs). But note that Kit Fine—never one to make hasty assumptions—explicitly allows that "a [strict] partial ground need not always be a part of a full ground" ("Guide to Ground", p. 53).

unexplained—or, it can be divided into logical parts such that each part is either fully explained or fully unexplained. This dichotomy can be expressed as

Atomism There are no irreducibly partial explanations.

Unfortunately, partial explanation understood in the way required by this dichotomy—as part of a full explanation—does not help us avoid brute differences. Thus, if **Atomism** is true, it will be impossible to harmonize **Immodesty** and **Articulation**. Let me explain.

Suppose you're a non-reductionist about predication, and you say that Alice and Bob partially explain what it is for Alice and Bob to be *F*, but they aren't the full explanation. If **Atomism** is true, then this means that there's something about Alice and Bob's being *F* which has been left unexplained. And, more to the point, the thing that's left unexplained is precisely the thing which makes the difference between the thing that's fully explained—i.e., the class of Alice and Bob—and the thing that the non-reductionist is positing as the "something more".

Thus, if **Atomism** is true, any attempt to explain the difference between the reductive and non-reductive positions would still rely on a brute difference, because the difference between what's given in the explanation and what's left unexplained would be fully unexplained.

Note that you can't avoid this problem simply by denying that there's some independent thing that could be added to the partial explanation to make it a full explanation. Let me put the situation abstractly. Suppose that *A* and *B* partially explain *C*. The worry arises as soon as we admit that there's something—call it *C-minus*—which is fully explained by *A* and *B*. As soon as we countenance such a thing, we need to explain the difference between *C* and *C-minus*. Just denying that there's something which is this difference—i.e., something that we could add to *C-minus* to get *C*—doesn't help, because there's still a difference between *C* and *C-minus* that needs to be explained.

What **Atomism** entails is that if we have only a partial explanation of *C*, then we will have no explanation of the difference between *C* and *C-minus*. Thus, **Atomism** entails that there will be a brute difference.

But it's not enough to simply reject the dichotomy between the (fully) explained and the (fully) unexplained. We also need to do this without simply positing in its place a new "brute" or "primitive" notion of partial explanation. How do we do this?

3.2.5 Irreducible wholes

Before I answer this question, let me introduce another. Some people think that there's a difference between "genuine" wholes and mere aggregates. "Genuine" wholes are supposed be more unified than aggregates—there's something more to the "genuine" whole than just its parts. But how can this difference be explained?

Obviously, if we want to explain how parts are unified, we can't just add some extra independent part which is supposed to unify the other parts, because then we'd need to explain how that extra part is itself unified with the original parts. This is a version of Bradley's regress.[13]

One response to this worry is to hold on to the thought that unified wholes are made up of parts plus something which unifies them, but deny that the explanatory regress is vicious. The way to do this is to say that there's no need to explain how the unifier—let's call it *glue*—unifies the parts. I ask: what makes something a genuine unity? You respond: its parts are glued together. So I ask: what unifies the glue to the parts? And you respond: that's just the sort of thing that glue is! This is, in effect, Russell's response to Bradley.[14]

This response is a non-starter for our purposes, because it is simply to posit a "primitive" explainer: glue. The difference between "genuine" unities and aggregates is supposed to be that the parts in "genuine" unities are glued together. And what is glue? It's just the thing that makes the difference between "genuine" unities and aggregates. If this is the full story, then positing glue results in a circular and spurious explanation of the difference between "genuine" unities and aggregates.

[13] F. H. Bradley, *Appearance and Reality*, ch. 2–3.

[14] See the discussion by Jonathan Schaffer, "It is the Business of Laws to Govern", pp. 581–2.

An alternate response is to take a more radical approach by denying that the unities are made up from parts, and to say instead that the "parts" are derived from the unities.[15] The idea is that if we take the unities as basic, then there's no need to explain what unifies the parts. Rather, parts are things which we extracted out of the basic unity.

This response is also a non-starter for our purposes, because we still don't have an explanation of the difference between "genuine" unities and aggregates. Of course, we might say that "genuine" unities are *fundamental*—but as I've argued elsewhere, this runs afoul of **Articulation**.

What this suggests is that as soon as we think of "genuine" wholes as having parts—whether those parts make up the whole or are derived from it—then we will have a problem explaining the difference between the aggregate of these parts and the "genuine" whole. In particular, we won't be able to explain this difference without relying on a brute difference.

Of course, one way to "solve" the problem of explaining unity is to say that the only "genuine" unities are trivial ones which have no complexity, so there's nothing to unify! In other words, one might admit defeat from the problem of explaining unity and say that there are only simples with nothing to unify and complex aggregates with no special unity to be explained.

But if we aren't ready to admit defeat, the alternative left open is to say that there are things which are neither entirely simple (with nothing to unify) nor wholes composed by parts (whose unity cannot be explained). If we want there to be "genuine" unities, it seems we must reject the dichotomy between simples and wholes, as we ordinarily understand them. "Genuine" unities would be things which aren't quite simples or wholes—at least, not as we ordinary conceive of simples and wholes.

[15] This is supposed to be one advantage of the view that the world is a world of states of affairs, rather than objects, properties and relations; see David Armstrong, *A World of States of Affairs*, pp. 28–9, 113–23. For a further argument for the view that objects and properties are derived from states of affairs rather than ingredients which make up the states of affairs, see Kit Fine's "Neutral Relations".

3.2.6 Logical parts

How does this help us with our original problem of explaining partial explanation? It helps because the problem of explaining "genuine" unity and the problem of explaining irreducibly partial explanation are analogous.

We can see this analogy by introducing the notion of a *logical part*. Talk of logical parts is a way of talking about the aspects of something which correspond to the parts of the explanation of that thing. If something is logically simple, then it can only be fully explained or fully unexplained—and if we posit such a thing, it would result in a brute difference. If something is the explanatory sum of its logical parts, then it is fully explained just in case each of its logical parts is fully explained, and it is partially explained just in case some logical parts are explained and some are not.

The problem of explaining irreducibly partial explanation then just is the problem of explaining how there could be something which is neither logically simple nor logically non-simple by having logical parts. And, the solution is to reject the dichotomy between things which are logical simples and things which are "composed" of logical simples—i.e., things that would be fully explained just in case each part of the full explanation was given.

But this solution is quite puzzling. How can there be something which *kind of* has logical parts and *kind of* doesn't? How could something only *kind of* be a part, or only *kind of* figure in the explanation of something else? This really does seem like an "all or nothing" type of thing.

3.2.7 Parthood and identity

Let's think a bit more generally about parthood. In the simplest case, parthood is a matter of identity. If we have a whole which is just a collection of things, then to be a part of the whole is to be identical to something in this collection.[16] And we can think of logical parts in a similar way.

[16] In other work, I argue that this is the best way to think about mereology for wholes which are "nothing over and above" their parts.

In the simplest case, *A* is a logical part of *B* just in case we can carve up what it is to be *B* and find something contained which is identical to *A*.

It's perhaps hard at first to see how there's any possibility between identity and non-identity. But it's easier if we replace identity with a nearby notion: *sameness*. There is conceptual space between perfect sameness—i.e., identity—and perfect difference. The thing which lives between identity and difference is *similarity*.

Consider now the sort of thing that would be neither perfectly simple nor complex in a way that involves parts. Such a thing could not be divided in a way that results in logical parts which are *identical* to other things—i.e., the things that would explain it. But because it's not perfectly simple either, what you would be able to do is to find ways in which it's *similar* to other things. This would be a type of similarity which involves no underlying respects of perfect sameness.[17]

What we have now are three dichotomies which—I have argued—must be rejected in order to make sense of something which harmonizes **Immodesty** and **Articulation**. We must reject the dichotomy between the explained and the unexplained, expressed by **Atomism**. We must reject the dichotomy between the simple and the complex. And we must reject the dichotomy between identity and difference.

But what we do not yet have is a positive picture of what we get when we reject these three dichotomies. Let us turn to this now.

3.3 Two metaphysical pictures

Before I present the positive picture, it may be helpful to make explicit the picture with which it is to be seen in contrast. This is the picture which embeds the dichotomies I have argued we must reject. It is the *constitutive* picture of explanation.

[17] Thus, I think John Duns Scotus is making an important mistake by thinking that—as Richard Cross puts it—"we can talk of a word's having two (or more) *similar* sense only if there is *something in common* between the sense", and "that analogy is [not] possible without univocity" (*Duns Scotus*, pp. 33–4).

3.3.1 Constitutive explanation

On the constitutive picture of metaphysical explanation, there is a stock of basic ingredients from which everything else in the world is "constructed". Every metaphysical explanation is a recipe which tells us which basic ingredients we need and how to put them together to yield the thing to be explained. The aim of metaphysics, on this picture, is to determine the list of basic ingredients, and to determine the "recipes" by which particular non-basic things are put together.

This picture embeds **Atomism**, since at the bottom level, everything is made up of explanatory simples, and these are the sorts of things which are either explained or unexplained *simpliciter*. This picture embeds the dichotomy between simples and complexes, since everything is either simple—one of the basic ingredients—or "constructed" out of such simples. And, this picture embeds the dichotomy between identity and difference. If A figures in the explanation of B, it's because A just is (identical to) one of the ingredients which make up B.

In rejecting these dichotomies, what we reject is the constitutive picture. I think it's easy to see that on the constitutive picture—as I presented it—there's a problem making sense of non-reductive views if we don't want to simply add the item to our list of basic ingredients. For instance, there's a problem making sense of emergence on this picture.

This might prompt us to add something to the constitutive picture in order to make it more sophisticated. But when we reject the dichotomies which underly the picture, we aren't just rejected the simple version of the constitutive picture—we're undermining any attempt to make sense of the world along these lines.

Thus, what we need is a new picture, and one which allows us to make sense of a world in which there is no strict dichotomy between the explained and the unexplained, the simple and the complex, or the same and the different. Let me now sketch such a picture.

3.3.2 Holistic explanation

Suppose that an architect comes up with the idea of a house, and then, based on this idea, has the house constructed. Houses are complex objects—they have lots of parts which have to be put

together in the right way. But—it seems to me—ideas needn't be complex in the way that houses are complex. I don't think ideas are all like little drawings which correspond—part to part—to the things of which they are ideas. Thus, the idea and the house are different in an important respect: the house is complex, the idea is simple—or at least, it's simple relative to the manner in which the house is complex.

But there's also an important sense in which the idea and the house resemble each other. The way the house was put together depends on the idea. Had the architect's idea been different, the house would have been built differently. We might say that the simple idea is similar to the house in that the house "expresses" what was already in the idea, and the idea "encodes" the complexity that shows up in the house.

Maybe you don't think ideas are like this—maybe you think they are more like little drawings in your head. But you get the idea, so to speak, that something simple could resemble something complex.

Let me offer another illustration. Consider the function $f(x) = x^2$, thought of something like a rule or algorithm.[18] This function generates a graph, which we can represent as a set of coordinates: (1,1), (2,4), (3,9), and so on. The graph is complex: it's made up of a bunch of points. The graph corresponds to the function: they are in some sense similar. But the function doesn't have parts which correspond to the parts of the graph; the function is simple with respect to the kind of complexity of the graph. The function "encodes" the complexity which is "expressed" by the graph.

The idea and the function are (relatively) simple objects which resemble complex objects. Let's call such resemblance *holistic* similarity, as opposed to *part-to-part* similarity.

My suggestion—my conjecture, if you like—is that by positing something which is holistically similar to complex things which we already understand, we are positing something which is

(i) partially explained by its similarity to the complex things, and—more importantly—

[18] It's important for this example that we don't simply identify a function with the set-theoretic construction used to model it, i.e., the set of tuples. For most purposes relevant to the use of functions in mathematics, this identification is harmless. But Frege famously rejects it, and it is not a harmless assumption when working with a higher-order or intensional logic.

(ii) partially explained in a way that also explains why this explanation could not be a full explanation.

3.3.3 Explanation without proof

Now—as I have explained already—I don't think that I can *prove* that there's such a thing as holistic similarity which satisfies these conditions. But there's an important question which I think can and should be addressed. Grant that there is such a thing as holistic similarity. How can we talk about holistic similarity without positing it as something entirely new—i.e., something that we can only explain by relying on a brute difference?

To answer this question, I must make use of the assumption that there is in fact such a thing as holistic similarity. The reason that holistic similarity is not *entirely* new is that it is itself holistically similar to more familiar things. In particular, what we get when we reject a dichotomy is holistically similar to the exclusive categories we are rejecting—something in which the exclusive categories are *kind of* present, but not perfectly, since they are unified in a way that doesn't exclude each other and which doesn't allow them to be perfectly divided.

Let me be clear: this isn't something which I think we can *fully* understand. But, if it is possible to harmonize **Immodesty** with **Articulation**, I think we can only do it by positing metaphysical explainers which are holistically similar to what we already fully understand. And I think that this notion of holistic similarity is itself something which we can partially explain because it is holistically similar to the dichotomies which would stand in its way.

Still, you might worry that I have cheated by calling my posit "holistic similarity". In particular, you might worry that I haven't shown that my posit really does deserve the name "similarity", and by calling it that, I've made it seem more familiar—you might worry—than it really is.

Could my posit count as a notion of similarity? No—you might say—because similarity is a matter of sharing properties and relations! Or, if you're more formally inclined, you might say that similarity is just a matter of isomorphism with respect to certain properties and relations.

Let me call this notion *strict similarity*. I agree that strict similarity deserves the name "similarity". And, if you want to stipulate that you're using that word "similarity" only for strict similarity, I have no objection; you're free to use your words as you please.

What's important to my argument is that we really can get a grip on the notion of my posit—holistic similarity—by the way in which it's *kind of like* similarity in the more familiar sense. And if you think similarity in the familiar sense just is strict similarity, then you'll say that holistic similarity isn't similar to strict similarity, so we don't have as much of a grip on it as I have claimed.

What exactly is at stake here? As I said, you are free to use the word "similarity" however you please. But the more serious worry which underlies our dispute about the word "similarity" is that there's no way to use strict similarity—something we're familiar with—to partially explain my posited notion of holistic similarity. For this partial explanation to work, the two notions would have to be similar. But, if "similarity" just means *strict* similarity, then they aren't.

Putting the objection like this begs the question. I agree that my posit isn't *strictly* similar to strict similarity—it isn't suppose to be. But this doesn't show that my posit isn't similar to strict similarity. My posit is supposed to be *holistically* similar to strict similarity. And what's at dispute is whether this fact—the fact that my posited relation itself bears itself to strict similarity—gives us a way to partially explain the posit.

Now, if I were to argue that holistic similarity is a form of similarity because it's holistically similar to our more familiar notion of strict similarity, then this too would be begging the question. The points I have made so far are only meant to be defensive. I have responded to the objection that this can't count as a form of similarity because "similarity" just means strict similarity. And, I have pointed out that we could consistently use "similarity" for a class of relations that would, by this notion of similarity, count as similar to each other—it wouldn't have to be a disjunctive notion of similarity.

Are there any positive reasons to think that our familiar notion of similarity can be extended to include something broader than strict similarity? It would be nice if I could do something more

than show the consistency of my position. And, I think I can do more. With some imagination, I think we can find examples in which a broader notion of similarity is at play.

The simplest cases involve comparisons across different sensory modalities. We think it makes sense to say that some flavor is smooth (like a texture) or that some smell is bright (like a color). Of course, you might think that flavors and textures are comparable because they share certain highly abstract features. But in the absense of an argument otherwise, I think it's natural to think that these shared abstract features don't exhaust the comparisons that we intend to make. Thus, it's natural to think that there's a more general notion of similarity at play in these comparisons.

We also compare sensible qualities to affective states: blue is mellow, Holst's "Thaxted" is strong and triumphant. And, we may think it makes sense to compare works of art in very different media: Perry's anthem seems to match Blake's "And did those feet in ancient time"—they "harmonize" with each other—while Gene Roddenberry's lyrics to the Star Trek theme do not.[19]

Of course, there's another way to interpret these cases. You might say that these aren't genuine cases of similarity because we're projecting features which aren't really in the objects. Blue isn't mellow—you might say—it just causes us to have mellow feelings. I don't think I can say anything to prove that this interpretation is incorrect. But I think this interpretation isn't the only one that's available to us. I think it's natural for us to think that we're picking up on genuine similarities between (say) colors and textures, or pieces of music and emotional qualities. And it seems like a stretch to say that these genuine similarities can be fully explained as cases of strict similarity. If they are genuine cases of similarity, then they involve some broader notion of similarity.

None of this proves that the holistic relation should count as a form of similarity. I don't think that a non-question-begging demonstration is possible. But it does suggest that with a little imagination, we can get some grip on the possibility of a more general notion of similarity than strict (or part-to-part) similarity. If we can get some grip on this more general notion, then I think we can partially understand our posited notion of holistic similarity through the way it falls under

[19]Plausibly, this last type of case can be explained by the previous type of case: the lyrics and music both correspond to affective states, and so the similarity or lack of similarity between words and music can be explained in terms of ordinary similarity or dissimilarity of the affective states to each other.

the more general notion of similarity and stands in this more general notion to strict (or part-to-part) similarity.

With this notion of holistic similarity in hand, I think we can finally explain the difference between the reductive and non-reductive positions about predication. On the reductive view, there are just individuals—simples—and things which can be "constructed" out of these individuals—classes. On the non-reductive view, there is something more than this. But not just something *more*—it is also something which doesn't quite fit into the category of a simple or the category of a complex whole made up of simples. What we posit is something which stands in a holistic similarity relation to the many individuals—it isn't perfectly identical in any respect to these individuals, but it is similar to them all.

3.3.4 Images of Reality

This is, as far as I can tell, the central feature of Plato's Forms. On Plato's picture, it is supposed to be the case that the many particular things we call chairs, to take one example, are imperfect "imitations" which "resemble" the ideal Form of the chair. Sensible particulars lead us to recollect the Forms to which they are similar, but sensible particulars—because they are sensible particulars—fall short of perfect similarity to the Form.[20]

Or, taking two examples from the creation myth of *Timaeus*, time is "a moving image of eternity" and the entire cosmos is "an imitation of the [changeless] model, something that possesses becoming and is visible."[21] At the core of Plato's metaphysics is this claim: to explain the particular, familiar, mutable objects—to give an account of what they are—we must posit something which is similar to them all but in no way perfectly instantiated by any of the sensible particulars with which we are familiar.

And this is more or less what I have argued we too should say if we reject the dichotomies between the explained and the unexplained, the simple and the complex, and the same and the different.

[20] See *Phaedo* 74.

[21] *Timaeus* 37d, 48e–49a.

Again—let me be clear—I don't think that I have given a *proof* we should believe in Platonic Forms and a metaphysical picture built upon imitation. But I do think that I've shown we have reason to endorse both **Immodesty** and **Articulation**, and to make sense of the view which results, we would need to posit things very much like Platonic Forms and the relation of irreducibly imperfect similarity. This isn't an argument that we ought to be Platonists. But it is an argument that it makes sense to be Platonists—there's explanatory work that could only be done by positing holistic explainers.

Let me conclude by summing up in theological terms. The reductionist is someone who takes familiar things as *idols*—things that we can fully grasp, and yet which are taken to be part of the grounds for reality. Against this, I have argued that we should understand non-reductionism as the view that familiar things are appropriately seen as *icons* and *images* which point to some reality beyond what we can fully grasp.

Bibliography

Armstrong, David. 1980. "Identity Through Time." In *Time and Cause: Essays Presented to Richard Taylor*, edited by Peter van Inwagen, 67–78. Reidel.

———. 1989. *Universals: An Opinionated Introduction*. Westview Press.

Bennett, Karen. 2017. *Making Things Up*. Oxford University Press.

Bradley, F. H. 1897. *Appearance and Reality*. Second edition. Oxford University Press.

Chalmers, David. 2009. "Ontological Anti-Realism." In *Metametaphysics: New Essays on the Foundations of Ontology*, 77–129. Oxford University Press.

Cross, Richard. 1999. *Duns Scotus*. Oxford University Press.

Dasgupta, Shamik. 2011. "The Bare Necessities." *Philosophical Perspectives* 25 (1): 115–160.

———. 2015. "Inexpressible Ignorance." *Philosophical Review* 124 (4): 441–480.

———. 2017. "Constitutive Explanation." *Philosophical Issues* 27 (1): 74–97.

deRosset, Louis. 2013. "Grounding Explanations." *Philosophers' Imprint* 13 (7).

Donnelly, Maureen. 2011. "Endurantist and Perdurantist Accounts of Persistence." *Philosophical Studies* 154 (1): 27–51.

Fine, Kit. 1994. "Essence and Modality." *Philosophical Perspectives* 8: 1–16.

———. 1999. "Things and Their Parts." *Midwest Studies in Philosophy* 23 (1): 61–74.

———. 2000. "Neutral Relations." *Philosophical Review* 109 (1): 1–33.

———. 2001. "The Question of Realism." *Philosophers' Imprint* 1 (1) (June).

———. 2006. "In Defense of Three-Dimensionalism." *The Journal of Philosophy*: 699–714.

———. 2008. "Coincidence and Form." *Aristotelian Society Supplementary Volume* 82 (1): 101–118.

———. 2009. "The Question of Ontology." In *Metametaphysics: New Essays on the Foundations of Ontology*, 157–177. Oxford University Press.

———. 2012. "Guide to Ground." In *Metaphysical Grounding: Understanding the Structure of Reality*, 37–80. Cambridge University Press.

———. 2017. "Naive Metaphysics." *Philosophical Issues* 27 (1): 98–113.

Goodman, Nelson. 1978. *Ways of Worldmaking*. Hackett Publishing.

———. 1983. *Fact, Fiction, and Forecast*. Harvard University Press.

Hawley, Katherine. 1999. "Persistence and Non-Supervenient Relations." *Mind* 108 (429): 53–67.

———. 2001. *How Things Persist*. Oxford University Press.

Hirsch, Eli. 2002. "Quantifier Variance and Realism." *Noûs* 36 (s1): 51–73.

Hofweber, Thomas. 2016. *Ontology and the Ambitions of Metaphysics*. Oxford University Press.

Johnston, Mark. 1984. "Particulars and Persistence." PhD thesis, Princeton University.

———. 1992. "How to Speak of the Colors." *Philosophical Studies* 68 (3): 221–263.

———. 1993. "Objectivity Refigured: Pragmatism Without Verificationism." In *Reality, Representation, and Projection*, edited by John Haldane and Crispin Wright, 85–130. Oxford University Press.

———. 2006. "Hylomorphism." *The Journal of Philosophy* 103 (12): 652–698.

Johnston, Mark, and Sarah Jane Leslie. unpublished. "Essence and Accident: Against Limited Variety."

Kraut, Robert. 2010. "Universals, Metaphysical Explanations, and Pragmatism." *Journal of Philosophy* 107 (11): 590–609.

Lewis, David. 1970a. "Counterparts of Persons and Their Bodies." In *Philosophical Papers: Volume 1*.

———. 1970b. "How to Define Theoretical Terms." In *Philosophical Papers: Volume 1*.

———. 1983a. "New Work for a Theory of Universals." In *Papers in Metaphysics and Epistemology*.

———. 1983b. *Philosophical Papers: Volume 1*. Oxford University Press.

———. 1984. "Putnam's Paradox." In *Papers in Metaphysics and Epistemology*.

———. 1986. *On the Plurality of Worlds*. Blackwell.

———. 1991. *Parts of Classes*. Blackwell.

———. 1994. "Humean Supervenience Debugged." In *Papers in Metaphysics and Epistemology*.

———. 1999. *Papers in Metaphysics and Epistemology*. Cambridge University Press.

Litland, Jon Erling. forthcoming. "Could the Grounds's Grounding the Grounded Ground the Grounded?" *Analysis*.

Magidor, Ofra. 2016. "Endurantism vs. Perdurantism?: A Debate Reconsidered." *Noûs* 50 (3): 509–532.

McDaniel, Kris. 2003. "No Paradox of Multi-Location." *Analysis* 63 (4): 309–311.

Plato. *Parmenides*. trans. Mary Louise Gill and Paul Ryan. Hackett.

———. *Meno*. trans. G. M. A. Grube. Hackett.

———. *Phaedo*. trans. G. M. A. Grube. Hackett.

———. *Timaeus*. trans. Donald J. Zeyl. Hackett.

Putnam, Hilary. 1977. "Realism and Reason." *Proceedings and Addresses of the American Philosophical Association* 50 (6): 483–498.

———. 1992. "Replies." *Philosophical Topics* 20 (1): 347–408.

Rorty, Richard. 1979. *Philosophy and the Mirror of Nature*. Princeton University Press.

Rosen, Gideon. 2010. "Metaphysical Dependence: Grounding and Reduction." In *Modality: Metaphysics, Logic, and Epistemology*, 109–36. Oxford University Press.

Schaffer, Jonathan. 2009. "On What Grounds What." In *Metametaphysics: New Essays on the Foundations of Ontology*, 347–383. Oxford University Press.

———. 2016. "It Is the Business of Laws to Govern." *Dialectica* 70 (4): 577–588.

Sider, Theodore. 1996. "All the World's a Stage." *Australasian Journal of Philosophy* 74 (3): 433–453.

———. 2001. *Four Dimensionalism: An Ontology of Persistence and Time*. Oxford University Press.

———. 2011. *Writing the Book of the World*. Oxford University Press.

———. 2013. "Against Parthood." *Oxford Studies in Metaphysics* 8: 237–293.

———. unpublished. *The Tools of Metaphysics and the Metaphysics of Science*.

Thomasson, Amie L. 2015. *Ontology Made Easy*. Oxford University Press.

Turner, Denys. 1995. *The Darkness of God: Negativity in Christian Mysticism*. Cambridge University Press.

Velleman, J. David, and Thomas Hofweber. 2011. "How to Endure." *Philosophical Quarterly* 61 (242): 37–57.

Wilson, Jessica. 2014. "No Work for a Theory of Grounding." *Inquiry* 57 (5-6): 535–579.