University of Massachusetts Amherst

ScholarWorks@UMass Amherst

Doctoral Dissertations   Dissertations and Theses

November 2023

The Dialectical Virtue of Ideological Reduction

Keehyuk Nahm
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_2

Part of the Epistemology Commons, and the Metaphysics Commons

Recommended Citation

THE DIALECTICAL VIRTUE OF IDEOLOGICAL REDUCTION

A Dissertation Presented

by

KEEHYUK NAHM

Submitted to the Graduate School of the

University of Massachusetts Amherst in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

September 2023

Department of Philosophy

THE DIALECTICAL VIRTUE OF IDEOLOGICAL REDUCTION

A Dissertation Presented

by

KEEHYUK NAHM

Approved as to style and content by:

_____

Sophie Horowitz, Chair

_____

Ned Markosian, Chair

_____

Phillip Bricker, Member

_____

Maya Eddon, Member

_____

Ana Arregui, Member

_____

Kevin Klement, Department Chair

Department of Philosophy

DEDICATION

Dedicated to my parents, Nahm Jinwoo and Kim Yoonsook

ACKNOWLEDGEMENTS

ABSTRACT

THE DIALECTICAL VIRTUE OF IDEOLOGICAL REDUCTION

SEPTEMBER 2023

KEEHYUK NAHM, B.A., SUNGKYUNKWAN UNIVERSITY

M.A., SUNGKYUNKWAN UNIVERSITY

Ph.D, UNIVERSITY OF MASSACUSETTS AMHERST

Directed by: Professor Sophie Horowitz and Professor Ned Markosian

Many would agree that there is something generally appealing and attractive about reduction. By this, I do not mean that reductive theories are accepted across the board, nor do I mean that they should be. All I mean is that there is something recognizably "good" about the reductive method that may be outweighed by other considerations. For instance, it is extremely rare for one to reject the reductionist position of a given domain while conceding that the proposed reductive procedure is successful. Typically, the opposition consists in denying that the subject matter *can* be reduced. This suggests an unspoken rule of the dialectical that, if something can be reduced, then it should be reduced; or, all else being equal, a reductive theory is to be favored over a non-reductive theory.[1] In a similar light, conventional metaphysical wisdom tells us that primitivist positions should only be accepted as a last resort. Again, there may be cases where the

---

[1] The closest example of an exception is William D'Alessandro (2017) who argues that the ontological reduction of arithmetic to set theory is successful but unexplanatory. However, he does not fully take this to imply that the reduction should be *rejected*, and suggests (in a non-committal fashion) that the reduction may earn its keep for largely pragmatic reasons (e.g., providing a shared standard of provability across different mathematical disciplines).

situation is dire enough for us to resort to primitivism. Nevertheless, few will oppose the general idea that we ought to reach for primitivism *only when* all other accounts fail. Let us call these attitudes 'UNSPOKEN RULE' and 'LAST RESORT', respectively.

In this dissertation, I offer an explanation of how these attitudes apply to ideological reduction by appealing to what I call dialectical virtues and vices. For this end, the first half of the dissertation is dedicated to clarifying what ideology is and what ideological reduction involves. In the second half, I introduce the notion dialectical virtues and vices and argue for their epistemic relevance. Dialectical virtues are features of a theory that place its advocates at a comparative advantage in the context of a debate; dialectical vices are those that place them at a comparative disadvantage. I argue that there are certain dialectical virtues and vices that are associated with (ideological) reductionism and primitivism, and that the above attitudes are the results of our being sensitive to them. Furthermore, I put forth this explanation, not as a merely psychological or sociological explanation of these attitudes. Instead, I argue that dialectical virtues and vices are epistemically relevant and that have serious implications for our evaluation of theories.

PREFACE

Many would agree that there is something generally appealing and attractive about reduction. By this, I do not mean that reductive theories are accepted across the board, nor do I mean that they should be. All I mean is that there is something recognizably "good" about the reductive method that may be outweighed by other considerations. For instance, it is extremely rare for one to reject the reductionist position of a given domain while conceding that the proposed reductive procedure is successful. Typically, the opposition consists in denying that the subject matter *can* be reduced. This suggests an unspoken rule of the dialectical that, if something can be reduced, then it should be reduced; or, all else being equal, a reductive theory is to be favored over a non-reductive theory.[2] In a similar light, conventional metaphysical wisdom tells us that primitivist positions should only be accepted as a last resort. Again, there may be cases where the situation is dire enough for us to resort to primitivism. Nevertheless, few will oppose the general idea that we ought to reach for primitivism *only when* all other accounts fail. Let us call these attitudes 'UNSPOKEN RULE' and 'LAST RESORT', respectively.

In this dissertation, I offer an explanation of these attitudes by appealing to what I call dialectical virtues and vices. Dialectical virtues are features of a theory that place its advocates at a comparative advantage in the context of a debate; dialectical vices are those that place them at a comparative disadvantage. On my proposed account, the above attitudes are the result of our being sensitive to dialectical virtues and vices that are associated with reductionism and primitivism.

---

[2] The closest example of an exception is William D'Alessandro (2017) who argues that the ontological reduction of arithmetic to set theory is successful but unexplanatory. However, he does not fully take this to imply that the reduction should be *rejected*, and suggests (in a non-committal fashion) that the reduction may earn its keep for largely pragmatic reasons (e.g., providing a shared standard of provability across different mathematical disciplines).

This account is not meant to be a merely psychological or sociological explanation of these attitudes. Instead, I argue that dialectical virtues and vices are epistemically relevant and that should be relevant to our evaluation of theories.

A few clarifications are in order. First, one may think a much more straightforward account can be achieved by appealing to some principle of theoretical simplicity. However, this is not as straightforward as it may initially seem. On the one hand, notice the sociological fact that those who are sympathetic to the relevant attitudes easily outnumber those who accept simplicity as an epistemic theoretical virtue (as opposed to a merely pragmatic one) in philosophy. [3] Thus, considerations of simplicity hardly explain why these attitudes are so widespread among philosophers. On the other hand, even for those who do subscribe to simplicity, the way in which simplicity is methodologically prescribed does not match up with the aforementioned attitudes. At the end of the day, simplicity is just one theoretical virtue that should be balanced against many others and is often qualified with a *ceteris paribus* clause. In contrast, the attitudes above are not that sensitive to such delicate details. Instead, it often goes without saying that nothing more than the mere prospect of a reduction is enough to put significant pressure on the primitivist's position. Although these considerations may not be enough to fully refute a simplicity-based explanation, they do suggest that an alternative is worth exploring.

The second clarification is that the scope of this dissertation is restricted to what we may call *ideological reduction* as opposed to ontological reduction. (As leading examples, I mostly focus on counterfactual analyses of causation, the Block Universe Theory of time, and Lewisian Modal Realism.) A key reason for this is that the more ambitious, comprehensive project is

---

[3] A selective survey: for theorists in favor of theoretical simplicity / parsimony in philosophy, see Baker (2003), Bradley (2019), Nolan (2003; 2014), Sider (2013); for theorists against it, see Brenner (2016), Huemer (2009), Roche (2018), Willard (2014); for theorists who give a more nuanced answer, see Sober (2015, especially chapter 5).

hampered by the fact that the literature does not contain enough efforts to clarify the very notion of ideology. To combat this difficulty, I dedicate the first half of this dissertation to clarifying what ideology is and what ideological reduction involves. In the second half, I introduce the notion dialectical virtues and vices and argue for their epistemic relevance, and how an explanation in terms of dialectical virtues and vices can be given for the aforementioned attitudes about reductionism and primitivism.

TABLE OF CONTENTS

CHAPTER 1: ON IDEOLOGY

1.1 Introduction

According to Quine, the ideology of a theory consists of the totality of "ideas" that are expressible within the theory (Quine 1951). This gloss of ideology, however, is elusive in many ways. In this chapter, I aim to sharpen the notion of ideology by drawing some distinctions and discussing their implications. As a starting point, we may approach this task by critically analyzing a recent taxonomy put forth by Peter Finocchiaro's (2019b). The taxonomy involves different views about what bits of ideology are. He describes these views as: (i) the linguistic view that says bits of ideology are the unanalyzed linguistic terms that are used to state the theory, (ii) the semantic view that says they are the meanings (i.e., concepts) of the expressions that are used to state the theory, and (iii) the worldly view that says they are the immanent feature of the world that the theory characterizes. (I say more about how this description of the worldly view should be understood in Section 1.3.) While this taxonomy provides a useful outlook on how philosophers generally talk about ideology, it is wanting in a number of ways. First, while Finocchiaro suggests that we must choose at most one of the three views, this seems unnecessary. Instead, it seems to suffice that we are clear about which sense is in use at a given moment. Second, while he does not develop the worldly view in much detail, the few remarks he makes do not seem to point to a plausible view of what bits of ideology are. Third, his taxonomy follows the tendency of not sharply distinguishing a narrow sense of 'ideology' that only refers to primitive or fundamental ideology from a broader sense that also includes non-primitive, non-fundamental ideology. (Notice

that only his description of the linguistic view is explicitly restricted to unanalyzed terms.) In what follows, I consider each of these shortcomings in turn and discuss how to address them.

Using these as talking points, I sketch out a form of semantic pluralism about the term 'ideology' according to which there are multiple senses of the term that are equally deserving of the name. Specifically, I argue that the term is ambiguous across two dimensions. The first concerns what bits of ideology are said to be. As suggested above, the version of pluralism I defend here recognizes two senses of 'ideology' that applies to linguistic terms and concepts, respectively. I call these two senses *linguistic ideology* and *conceptual ideology*. The second dimension of the ambiguity involves a restricted and an unrestricted sense of ideology. In many cases, philosophers use 'ideology' to specifically refer to *primitive* ideology. But we may entertain a more general sense of the term that also encompasses non-primitive bits of ideology as well. I call these two senses *primitive* and *general* ideology. In sum, I argue that (i) 'ideology' involves ambiguity between primitive linguistic ideology, general linguistic ideology, primitive conceptual ideology, and general conceptual ideology, and that (ii) this ambiguity can be made harmless simply by being clear about which sense of the term is being used.

The plan of the chapter is as follows. In Section 1.2, I explain why linguistic and conceptual ideology are not competing conceptions of ideology and discuss how the two relate to each other. In Section 1.3, I explain why I reject that the worldly view characterizes a proper sense of 'ideology'. In Section 1.4, I discuss the distinction between primitive and general ideology. Specifically, I focus on (i) how attending to non-primitive ideology can be serviceable to non-fundamental metaphysics and (ii) developing a criterion for a theory's commitment to general ideology.

1.2 Linguistic and Conceptual Ideology

When faced with multiple options, the philosopher's natural inclination is to make a choice. But this is not necessary for linguistic and conceptual ideology. The fact is that we use linguistic terms to express the concepts we grasp and possess concepts that represent the world in certain ways. And depending on the context, we may freely choose to talk in terms of the one sense instead of the other. Take the case of modal ideology. Sometimes, we may speak of linguistic items such as the modal operators '□' and '◊' when, say, we need a rigorous and precise formulation of our theories before our eyes. But when this does not matter as much, we may speak of the concepts of *possibility* and *necessity* to explore our understanding of modality. Some may have reasons to actively reject one of the two senses of 'ideology'. For instance, as Finocchiaro notes, skeptics of meaning (like Quine himself) would be inclined to avoid talking in terms of conceptual ideology. Nevertheless, this is hardly a reason to think that there is a single sense of 'ideology' that best captures how philosophers talk about ideology. And for those who do not share such skeptical sentiments, all that matters is that we appreciate how a theory's linguistic ideology relates to its conceptual ideology.

A key clarification to make is that none of my above remarks are meant to suggest that the two senses of ideology are interchangeable, nor that there is a one-to-one correspondence between them. While terms and concepts serve a common purpose of representing the world in certain ways, they are nonetheless different things that may involve different sets of criteria for ideological commitment. For linguistic ideology, all we need to do is look at the page on which the theory is written; roughly speaking, if certain linguistic resources are used to state the theory, then the theory

is ideologically committed to those resources.[4] However, a theory may commit to concepts that are not expressible by its linguistic ideology, at least not in a straightforward way. For example, it seems possible for a theory to be couched in a finite language such as English but nonetheless be committed to a continuum of concepts. Suppose a theory ideologically recognizes a continuum of colors between red and yellow in the following sense: (i) the theory posits primitive color predicates (e.g., 'is red', 'is reddish', 'is yellowish', etc.) to express the various ways in which objects can be colored, (ii) it posits a primitive two-place predicate 'is redder than' that applies to objects, and (iii) the theory claims that, for any two objects o_1 and o_2, if o_1 is redder than o_2, then it is possible for there to be an object o_3 that is redder than o_2, but not than o_1. In this case, the theory conceptually recognizes more shades of red than it can name; while it recognizes a continuum of ways in which objects can be distinctly colored, there are not enough color terms in English to go around.[5,6]

The above remarks suggest that there may be no single, complete way to represent a theory's ideology. Thus, in every case, we need to specify whether we are talking about linguistic or conceptual ideology. As far as our description of a theory's ideology goes, this perhaps is not much of a concern. It would be nice if there were such a way. But so much for the worse if there is not. As long as we are keen on the relationship between the two, this discrepancy does not pose a significant obstacle to understanding the theory. However, it does raise a more pressing issue for

---

[4] Daniel Rubio (2020) makes note of the same "look at the page" procedure for identifying the linguistic ideology of a theory. But contrary to what I say immediately above, he assumes that the results of this method allow us to identify the conceptual ideology of the theory as well.

[5] For this reason, it is also a misleading to say conceptual ideology constitutes "the semantic view." For the color theory at hand, only a subset of its conceptual ideology stands in a semantic relationship to its linguistic ideology; thus, there is more to the role of conceptual ideology than its role in semantics.

[6] The point here is *not* that the theory is unable to state the proposition that there is a continuum of ways in which objects can be colored. As I just did, this proposition can easily be expressed in English. The point is that, because there is a mismatch between the number of color predicates and the number of color concepts the theory commits to, the theory's linguistic ideology and conceptual ideology are non-isomorphic.

measuring and comparing ideological simplicity. If the commitments to the different types of ideology do not fully overlap, we may have a theory that is ideologically simpler than a rival in terms of one type of ideology but not in the other. Thus, we may not be able to say that one theory is ideologically simpler than another *all things considered*.[7]

For those who do not accept simplicity as an epistemic theoretical virtue, this issue may also be brushed off to the side.[8] But for those who *do* accept simplicity as a theoretical virtue, there is fortunately a solution. A simple way around this issue is to take ideological simplicity to be measured solely in terms of (primitive) conceptual ideology. Consider again the previous theory of colors. In this case, the reason the theory commits to fewer color predicates than concepts is because of the expressive limitations of English. And insofar as simplicity is considered to be theoretical virtue, it seems plainly wrong to think that a theory is able to achieve this feat because of the limitations of the language it is couched in. In this light, it seems plausible to restrict the role of linguistic ideology to serve as a written record of our theories for communicative (and archival) purposes. For instance, if it weren't for our corporeal limitations, and if it were possible for us to directly communicate our thoughts without the aid of language, then there would be no need for linguistic ideology at all. (Or so it seems.)

Here is a more long-winded argument. One may object that the problem of expressive limitations can be overcome by recasting the theory in a much more expressive language. This line

---

[7] This issue of ideological simplicity is also the main reason Finocchiaro argues that philosophers should collectively choose one among the many senses of 'ideology'. (Note: Finocchiaro specifically expresses concerns about confusions regarding simplicity judgments regarding the number of ideological kinds a theory commits to. The notion of ideological kinds and their relationship to ideological simplicity was first proposed by Sam Cowling (2013); see Finocchiaro (2019a) for Finocchiaro's own view of ideological kinds. For reasons of space, I ignore these issues and treat ideological simplicity as only concerning individual primitives.)

[8] For theorists who reject simplicity as a theoretical virtue in philosophy, see Brenner (2017), Huemer (2009), Roche (2018), Willard (2014); for theorists in favor of simplicity in philosophy, see Baker (2003), Bradley (2019), Nolan (1997; 2014), Sider (2013); for a more nuanced view, see Sober (2015, especially chapter 5).

of thought follows in the footsteps of Quine. While Quine's discussion of ideology initially focused on linguistic terms (Quine 1951), he later expresses a more expansive understanding of ideology that, in a certain sense, brings concepts more to the forefront. He writes (Quine 1983, p.501):

> I now propose to extend ideology beyond the subject's own verbal limits, to cover inarticulate abilities to recognize and discriminate. If the subject is disposed to react to the presence of any and every cat in some manner in which he is not disposed to react to anything but cats, then I shall reckon the term 'cat' to his ideology, though it is strictly my term, with which he may or may not be conversant.

Given his well-known sympathy toward behaviorism, we may plausibly understand his talk of "behavioral dispositions" in terms of the subject's possession of certain concepts. Understood this way, the core idea here can be said to be to interpret one's theory to contain the term 'cat' in case the concept *cat* plays some non-eliminable role in the theory. In our terminology, this amounts to interpreting a theory's linguistic ideology in such a way that ensures that it contains a term for every bit of its conceptual ideology.

This interpretative method for ideology can be understood in tandem with Quine's method of determining ontological commitments. On the Quinean method, ontological commitments are to be "read off" not from *any* formulation of a theory, but instead from a canonical regimentation of it. There are of course substantive issues regarding what a canonical regimentation involves and

what it is to "read off" ontological commitments from it.[9] But at a general level, the underlying idea can be expressed as the claim that there is some privileged, ideal language that is suitable for determining the ontological commitments of a theory. In a similar spirit, it can be argued that the ideological commitments should be determined in the same way. Plausibly, the canonical regimentation of a theory should not only provide insight into its ontology, but into its ideology as well. Understood this way, the quoted passage above can be read as laying down a desideratum for how a theory should be regimented, namely that it should be regimented in a way that its conceptual ideology is expressed in its entirety. This ensures that linguistic and conceptual ideology never come apart, which resolves the worry about there being different ways to measure ideological simplicity.

The problem with this approach, however, is that the theory would need to be couched in a language that is practically unusable both as a language and as a means to measure ideological simplicity.[10] Consider once again the previous theory of colors. Because the theory contains non-denumerably many color predicates, it is impossible for *us* to write down each color predicate that can be read off from the canonical regimentation. Thus, to be able to communicate the theory, we would need to resort to a non-canonical regimentation of the theory and infer what its canonical regimentation would look like. But this seems no different from what I suggested: to ignore our linguistic expression of the theory and to instead focus on that which it expresses.

---

[9] See Bricker (2016).
[10] I thank Phil Bricker for pointing this out to me.

1.3 The Exclusion of Worldly Ideology

As previously mentioned, Finocchiaro suggests that the term 'ideology' may also be used to refer to the immanent features of the world that our theoretical terms and concepts characterize. Let us call this sense 'worldly ideology'. In this section, I explain why I exclude worldly ideology from my version of pluralism and reject that worldly ideology constitutes a proper sense of 'ideology'. To be clear, I am not arguing against the *metaphysical claim* that linguistic and conceptual ideology express certain features of the world. Let us call this metaphysical claim *the worldly view of ideology*. Although not much of what I say in this dissertation hangs on the worldly view of ideology, to show all my cards, I am largely sympathetic to it. What I oppose is the *semantic claim* that the term 'ideology' can be used in a way to directly refer to these "features of the world," whatever those turn out to be.

To fully appreciate my reasons for rejecting worldly ideology as a sense of 'ideology', we need to flesh out the idea of worldly ideology in some more detail. The key question to ask is what this sense of 'ideology' identifies bits of ideology with: i.e., what it is that serves as "the features of the world" that are characterized by linguistic and conceptual ideology. Let me first address how this question should *not* be answered. From the above gloss of worldly ideology, one may think that the task at hand is to search for a kind of entity that we can identify ideology with. However, this is mistaken. The problem with this approach to the question is that it blurs the distinction between ontology and ideology in a problematic way. For instance, consider the commonly accepted idea that there is a type of commerce that obtains between ontological and ideological commitments—that we can reduce our ontological commitments by positing primitive

ideology and *vice versa*.[11] If worldly ideology is understood as consisting of entities, contrary to this idea, no such commerce would be possible. On this conception, the postulation of primitive worldly ideology would simply amount to the postulation of more entities; thus, one would merely be trading in one kind of entity for another.

The deeper problem at hand concerns the issue of ontological correlation. The reason the above kind of commerce is possible is because, in general, it is not required that every bit of a theory's ideology corresponds to some element in its ontology.[12] Thus, the commerce proceeds by trading in an ontological commitment for an ideological commitment that has no ontological correlate. But the fact that some bits of ideology lack ontological correlates plays a much more important role in contexts outside of this kind of commerce. This is because taking on certain kinds of ideological commitments (particularly, those to primitive ideology) not only allows one to reduce the number of entities one commits to, but also allows one to avoid specific ontological commitments that are problematic for one's theory. This holds regardless of whether or not one takes predicates and concepts to generally have ontological correlates and applies to nominalists (broadly construed) and universal theorists alike. That is, even the universal theorist needs to posit primitive ideology that has no ontological import. For example, on the pain of Bradley's regress, it is commonly held that the universal theorist must take her notion of *instantiation* to be primitive. Here is how the general story goes. One way to characterize a theory of universals is that it accepts the following instantiation schema for a wide range of predicates:

(IS) '*a* is *F*' is true in virtue of *a*'s instantiation of the universal *F-ness*.

---

[11] See Oliver (1996) for a general discussion.
[12] See Quine (1951).

Different versions of the theory take the schema to apply to different ranges of predicates. For instance, a *sparse view* of universals restricts (IS) to only those predicates that express fundamental properties whereas an *abundant view* applies it much more broadly. But even on an abundant view, the general consensus is that (IS) does not apply to the instantiation predicate itself. If such an application is allowed, the universal theorist must posit a universal *I* that corresponds to the two-place predicate 'instantiates' and take instantiation facts to hold in virtue of the instantiation of *I*. That is, it requires the universal theorist to accept:

(1) '*a* instantiates *F-ness*' is true in virtue of *a*'s and *F-ness*'s instantiation of *I*.

The acceptance of (1) opens the gates to Bradley's regress. On the pain of inconsistency, affirming (1) requires the universal theorist to posit a further universal, call it '*I**', to account for the instantiation of *I*. In turn, consistency requires her to posit yet another universal, call it '*I***', in order to account for the instantiation of *I**, so on and so forth *ad infinitum*.[13]

One popular response to this regress is to block the regress by rejecting (1). And this is often done by insisting that 'instantiates' does not stand for a universal but instead expresses a fundamental, "non-relational tie" that unites a particular with its universals. Plausibly, this is simply a long-winded way of saying that 'instantiates' is a primitive bit of the theory's ideology.[14] On the one hand, the appeal to a non-relational tie hardly qualifies as an analysis. If anything, it is a metaphor that sheds indirect light on the metaphysical nature of instantiation. Hence, 'instantiates' remains to be an unanalyzed predicate, and thus is a bit of the theory's primitive ideology. On the

---

[13] There is some disagreement over whether the Bradleyan regress is in fact vicious, however. (See Maurin (2012) and Perovic (2017) for a comprehensive overview of the issue.) The point of this example, however, is simply to illustrate that the need for primitive ideology that has no ontological correlate is not specific to nominalists and that it can arise in dialectical contexts outside of nominalism about properties.

[14] A similar diagnosis is provided in Bricker (2018, ch1) and Cowling (2020).

other hand, it is important to note that it is not enough to resist positing a *universal* of instantiation. What generates the regress is the idea that the union between a particular and a universal must be accounted for by positing a further *entity*, and whether the posited entity counts as a universal is inconsequential.[15] Thus, for the solution to have any force, this talk of a "non-relational tie" must not be cashed out in ontological terms at all. By bringing the two hands together, it simply follows that, according to this solution, the instantiation predicate is a bit of primitive ideology that has no ontological correlate.

The lesson of the discussion so far is that Finocchiaro's talk of "features of the world" should not be taken to be referring to entities of any sort. Accordingly, worldly ideology should not be understood as consisting of entities, either. If this is correct, then it seems that the only viable interpretation is that worldly ideology describes a sense of 'ideology' that identifies ideology with what Ted Sider (2011) describes with his notion of non-ontic, metaphysical structure.

For Sider, the notion of metaphysical structure plays an important role in articulating the idea that *ideology matters*. The idea he articulates has two parts. The first is that "ideology matters" in the sense that judgments about ideology are not merely "psychological" or "conceptual," but concern objective matters about the world. To this end, he argues that there is a fact of the matter as to the kinds of structure that the world contains and takes the ideology of a theory to be responsible for capturing this. The second part is that "ideology matters" in the sense that it serves a theoretical role that is distinct from ontology. For reasons not unrelated to our previous discussion, Sider argues that judgments about ideology concern an *aspect* or *dimension* of the world that cannot be captured by ontological terms.[16] To this end, he emphasizes that talk of metaphysical

---

[15] See Perovic (2017), especially section 3.2.
[16] See Sider (2011, especially ch6.3).

structure should not be understood to be about some kind of entity but nevertheless should be understood as saying "something true, something objective, something about reality" (Sider 2011, p.109).

Although I am largely sympathetic to Sider's overall meta-metaphysical framework, I still think it is a bad idea to use 'ideology' to directly refer to metaphysical structure. I provide two main reasons. The first reason is that worldly ideology does not seem to be a sense of 'ideology' that is actually in circulation. For starters, observe that Sider himself does not use the term 'ideology' in this way. In every instance, ideology is said to *express* or *capture* the world's structure, and nowhere in his writings does he ever identify the two. Furthermore, while Finocchiaro associates the notion of worldly ideology with two authors, neither of these authors say enough to warrant this association. The two authors he discusses are Guiliano Torrengo (2014), who is said to identify ideology with (fundamental) properties, and Stephan Krämer (2010), who is said to identify ideology with facts. (Note: Since properties and facts are commonly thought as *entities*, this interpretation technically associates them with the conception of worldly ideology I object to above. But for the reasons I provide below, we need not associate them with that conception of worldly ideology, either.) It is difficult to fully address the point without providing an extensive exegesis of each of their discussions. But in a nutshell, what seems to motivate Finocchiaro's interpretation in both cases is that the authors each claim that the postulation of fundamental properties and of facts leads to an increase in ideological complexity.[17] From this, he seems to infer that the two authors are respectively taking ideological complexity to be a measurement of the number of fundamental properties and of facts that a theory accepts. In turn, he concludes that the two authors are respectively identifying ideology with properties and with facts. However, their

---

[17] As it happens, both authors make these remarks as part of their criticism of presentism, though of different varieties.

12

point about ideological complexity does not in any way require them to *identify* ideology with either of the two. Instead, an equally reasonable interpretation is that they take ideological complexity to be merely correlated to the postulation of these entities. For instance, for all they have said, it is possible that they maintain that the correlation exists because of the primitive terms and concepts that need to be posited in order for these properties or facts to be expressed within the theory.

The second reason I reject worldly ideology is that it describes a sense of 'ideology' that is virtually unusable. This stems from the fact that, due to its non-ontic nature, it is particularly difficult for us to talk about structure in a direct and non-metaphorical way. In order to avoid reifying structure, we need a way to talk about structure that does not treat them as entities. This means that, unless we are talking metaphorically, we shouldn't really be using any referential expressions or quantificational idioms to talk about structure. In this light, the expressions 'features of the world' or 'structure' that I have been using so far should be understood to be metaphorical. Then how can we talk about structure non-metaphorically? According to Sider, our best bet is to ascend to the semantic plane (See Sider 2011, pp.108-109). That is, instead of directly making claims about the kinds of structure that the world contains, he suggests that we make claims about which of our expressions are structural. For this purpose, he introduces an operator '$\mathscr{S}$' that, when attached to an expression of any grammatical category, expresses that the expression in question is a structural term. For instance, suppose we wish to express the idea that the world contains a kind of structure that corresponds to some bit of (linguistic) ideology *I*. On Sider's framework, this idea should be expressed by '$\mathscr{S}(I)$', which in English, loosely translates to 'The expression '*I*' is structural.'.

While this strategy allows us to make various first-order metaphysical claims about what structure the world contains, it has limited utility for making meta-theoretical claims about the nature of structure. For instance, it seems desirable for us to be able to say that two distinct bits of primitive ideology *I* and *I** express distinct *kinds* of metaphysical structure, e.g., that primitive tensed operators express a kind of *temporal* structure whereas primitive modal operators express a kind of *modal* structure. However, it is hard to see how this idea can be expressed using '$\mathscr{S}$' without the aid of any referential or quantificational idioms.[18] In this sense, there is also no obvious way for us to even distinguish between different bits of worldly ideology. Second, using '$\mathscr{S}$' to talk about worldly ideology makes our talk of worldly ideology unavoidably indirect. Since the strategy is to talk about structure indirectly by talking of expressions, it also has us talk about worldly ideology indirectly by talking about linguistic ideology. This means that we can never actually adopt worldly ideology in our discourse; anything we want to say about worldly ideology must be regimented in terms of linguistic ideology instead, which makes it unsurprising that nobody actually uses 'ideology' in this way.

In sum, I conclude from these considerations that the notion of worldly ideology describes a novel sense of 'ideology' that we don't really need. For this reason, I exclude it from my pluralism about 'ideology'.

---

[18] In chapter 2, I discuss that a similar reason suggests against taking the relata of reduction to be Siderian structure as well.

1.4 General Ideology

1.4.1 The Need for Non-Fundamental Ideology

Let us now consider general ideology. I do not expect the idea of general ideology itself to be controversial. When Quine coined the term 'ideology', he explicitly took the question of "what ideas are fundamental or primitive for a theory" to be a "subdivision of ideology" (Quine 1951, p.14.). So, from the start, 'ideology' was meant to extend to the non-fundamental and to the non-primitive. However, because of the tendency to ignore non-primitive ideology, the notion of general ideology has been poorly explored.[19] For reasons of space, I only briefly consider two issues. First, I explore what we can achieve by attending to general ideology. To this end, I consider its role in non-fundamental metaphysics. Second, I sketch out a criterion for a theory's commitment to general ideology.

An obvious reason to attend to non-primitive ideology is that we need it in order to do non-fundamental metaphysics. Over the past few decades, an increasing amount of attention has been given to metaphysical inquiries about non-fundamental topics such as social ontology—the study of social kinds and entities such as race, gender, etc. As social kinds are often thought to be non-fundamental, several authors (especially those working on feminist metaphysics) have issued a sharp, meta-metaphysical reminder that we should not take metaphysics to be solely the matter of the fundamental and that there are important and substantive matters to be resolved regarding the non-fundamental as well.[20] The general lesson from this is that that our canonical toolkit for metaphysics (and philosophy in general) should include ways to investigate non-fundamental

---

[19] To be fair, the tendency to ignore non-primitive ideology is not without good reason. Plausibly, it is only the primitives that count against ideological simplicity, and hence, there are many contexts to which non-primitive ideology is irrelevant.
[20] See (Barnes 2014; Mikkola 2015).

phenomena. And one of the minimal requirements for this is to endorse some form and degree of permissivism regarding the non-fundamental; after all, if our theories cannot even recognize that putatively non-fundamental phenomena concerning genders and races *obtain*, our investigations of them would not be able to get off the ground.

One way to meet this requirement is to accept some form of ontological permissivism regarding non-fundamental entities. For instance, following Jonathan Schaffer (2009; 2010), one may adopt a notion of fundamentality that attaches directly to entities, and demand parsimony for fundamental entities while maintaining permissivism for non-fundamental entities.[21] This way, one may acknowledge the occurrence of some non-fundamental phenomenon by way of their ontology. E.g., one may acknowledge phenomena involving social kinds such as the above by identifying them with non-fundamental entities that are distinct from their grounds. However, this is not the only way. As an alternative, I put forth what we may call 'ideological permissivism' which is a methodological principle according to which one is free to posit non-fundamental, non-primitive ideology as one sees fit. By accepting ideological permissivism, one may acknowledge the occurrence of non-fundamental phenomena by taking it to be describable solely in terms of non-fundamental ideology.

'Ideological permissivism' is a new term for an old idea, and the acknowledgement of non-fundamental phenomena by way of ideology is nothing new. In this light, I am merely pointing out that we already have an effective means to engage in non-fundamental metaphysics at our disposal. Consider the classic debate between the universal theorist and the nominalist. Both sides of the debate agree that property-phenomena obtain; i.e., they both affirm various sentences that express

---

[21] See Schaffer (2017) for a fuller picture of how the framework helps with non-fundamental metaphysics. Also see Griffith (2017) for how grounding is serviceable to non-fundamental metaphysics, especially social ontology.

cases of attribute agreement (e.g., "There are two objects that are both red."), along with those that involve abstract singular terms (e.g., "Red is a color."). But they disagree about the nature of property-phenomena. One aspect of this disagreement is obviously ontological, as they disagree over the existence of universals. But another aspect of it is over whether property-phenomena are fundamental. Because universals are supposedly posited as fundamental entities, it is appropriate to interpret their view as taking property-phenomena to obtain at the fundamental level of reality. In opposition, the nominalist (traditionally construed) denies this, the rationale being that property-phenomena can be accounted for without positing anything that is ontologically or ideologically fundamental. To this end, she maintains that property-phenomena can be fully described without introducing any new entities and with only employing terms and concepts that are analyzable in some way. A key aspect of this line of reasoning is that property-phenomena are said to be non-fundamental by virtue of the fact that they can be accounted for purely in terms of non-fundamental ideology.

(Sidenote for Lewisians: It is important to stress that the above is not meant as an exhaustive description of *the* debate over properties, and the landscape is much more nuanced. The above description follows the tradition of associating ontological posits with claims of fundamentality and ideological posits with claims of non-fundamentality. But this association does not hold across the board, neither in the properties debate nor in metaphysics in general. If one's meta-metaphysical framework allows for non-fundamental entities as distinct existents, one may take a phenomenon to be non-fundamental by accommodating it through non-fundamental ontological posits. Also, one may take a phenomenon to be fundamental by making primitive ideological posits. For the latter, consider A-theories of time that posit primitive tensed operators, or primitivist theories of modality that posit primitive modal operators. For many, positing

17

primitive ideology for the purpose of accounting for phenomena regarding temporality and modality counts as admitting that the relevant phenomena to concern the fundamental level of the world.[22] In this light, even the nominalist about properties may agree that (at least some) property-phenomena are fundamental in case she posits primitive ideology for it. And this is the case for Lewis's natural class nominalism (Lewis 1983; 1986) according to which (i) properties are identified with sets of possibilia and (ii) there is a primitive distinction between natural and non-natural properties. On this view, the properties themselves are "non-fundamental" in the sense that we need not posit entities whose sole purpose is to be identified with them. Nevertheless, because he posits a primitive notion of naturalness, property-phenomena that involve natural properties are accepted as fundamental. That is, the theory accepts that there is something special and deeply profound about sharing a natural property such that its recognition requires positing fundamental, primitive ideology.)

For the above kind of nominalist about properties, the claim of non-fundamentality is often made as an "evasive" maneuver in the sense that its main purpose is to *deny* that the instantiation of properties is a fundamental phenomenon. That is, the key point that is being made is that, at the fundamental level, the world is devoid of any property-phenomena. But the same kind of claim can be made for a different dialectical purpose. In the case of non-fundamental metaphysics, one can claim that a given phenomenon is non-fundamental to actively *affirm* that it is a part of reality by claiming that it obtains at the non-fundamental level. In this context, the claim of non-fundamentality is made not as a response to an opponent, but as a description of one's position. In this fashion, the social ontologist (or rather, the social metaphysician) may account for social kinds by way of non-fundamental ideology. That is, instead of accommodating social kinds into her

---

[22] This is so on the previously discussed, objective conception of ideology.

framework by positing them as non-fundamental entities, she may do so by taking the ideological posits we need to talk about the social domain to be non-fundamental and analyzable.

By saying all this, I am not suggesting that ideological permissivism and ontological permissivism are competing methodological principles. On the face of it, it seems possible that certain kinds of non-fundamental phenomena call for non-fundamental ontological posits whereas others call for non-fundamental ideological posits. Or perhaps in certain cases, the subject matter may require that we make both kinds of posits. But for what it's worth, there are two advantages that ideological permissivism has over its ontological counterpart.

First, ideological permissivism is supported by theoretical conservativism in that it is already integral to how theorists typically think about ideology. As previously mentioned, it is widely held that ideological simplicity is restricted to primitive ideology, and theorists are typically unfazed by the employment of new terms and concepts that come with an analysis. In various cases, it may of course be contested whether the newly introduced bits of ideology can play the role that is envisioned for them and whether there are any interesting truths that can be expressed with them. But this does not undermine ideological permissivism as a methodological principle. If anything, it is matter of how the dispute in question should be resolved. And this is something that should be settled within the arena of non-fundamental metaphysics. For instance, consider how it helps characterize a debate Barnes (2014, p.340) alludes to, namely Sally Haslanger's and Ásta's debate over *what gender is*.[23] According to Haslanger, a person's gender is determined by their place in hierarchical social structures based on "normative assumptions of perceived sex characteristics and their assumed role in reproduction." In contrast, Ásta argues that a person's gender is

---

[23] At the time Barnes's article was published, Ásta's work has been published under "Ásta Sveinsdóttir."

determined by "patterns of belief about social role and status." The meta-theoretical approach I am suggesting is to not worry about whether gender, along with Haslanger's hierarchical social structures or Ásta's patterns of belief (and the social role and status that these beliefs are about) *exist as entities*. Instead, it is to treat the relevant terms as non-fundamental ideological posits that are used to describe facts about entities we accept independently from the debate, namely people. Understood this way, the issue at hand can be understood as a dispute over what non-fundamental ideology we need to posit in order to give an adequate explanation of what gender is. The role of ideological permissivism is to provide a meta-theoretical characterization of this debate as being over what non-fundamental ideology is required for this task. And determining *which* ideological posits are fit for this task is to be determined within the debate thus characterized.

The second advantage is that ideological permissivism evades a problem that ontological permissivism faces regarding certain kinds of issues. Schaffer's ontological permissivism is part of his broader meta-metaphysical framework that takes the focus of metaphysics away from existence questions and places it on questions of what grounds what. E.g., instead of inquiring whether numbers or composite objects exist, we are supposed to inquire whether these entities are grounded by other entities, and if so, what those grounds are. But because questions about what grounds what presuppose the existence of entities that stand in the grounding relation, a meta-metaphysical framework that focuses on grounding questions presupposes the existence of a vast array of entities that these inquiries are about. In this vein, the role of ontological permissivism is to provide justifications for this presupposition, and thus supply the entities that serve as the subject matters of our investigations.

A byproduct of this framework is that it rules out the possibility of there being substantive questions about what non-fundamental entities exist. From the ontological permissivist's

perspective, any existence question regarding a non-fundamental domain is to be trivially answered in the affirmative. But this seems wrong. For instance, as noted by Barnes (2014, pp.342-343), there are many substantive existence questions we may ask about the social domain, e.g., questions about *which* specific genders exist (e.g., whether genderqueer, non-binary, genderfluid, etc. all exist as distinct social categories), and relatedly, about *how many* of them there are. The problem here is not about the veracity of the ontological permissivist's answers to these questions. Rather, the problem is that it fails to account for their substantivity. That is, while the above kinds of questions should be answered through serious metaphysical investigation, the ontological permissivist's framework wrongfully prejudges them at the meta-theoretical level.

In contrast to this, ideological permissivism helps us give a simple yet plausible diagnosis of the kinds of issues Barnes alludes to. For the questions about which genders exist, the ideological permissivist gives a similar answer as the ontological permissivist and says that each (non-synonomous) gender term picks out a distinct social category. However, this answer is entirely unproblematic in that it does not entail any substantive judgments about the subject matter. All that is implied by this answer is that there are different kinds of *intelligible* social categories, which is something that one must be presupposed for one to even ask whether the relevant kinds of genders exists. The substantive issue is whether any of these gender terms have *instances*, and probably in many contexts, whether any of these instances are located in the actual world. These are substantive questions that ideological permissivism cannot (and should not) answer. As a methodological principle, its role stops at providing a framework to understand substantive metaphysical issues. And for these cases, it is able to do so in the above way.

1.4.2 The Criterion for General Ideological Commitment

Let us now turn to the question of what a theory's general ideology consists of: i.e., what the criterion for the commitment to general ideology is. The most natural way to think about general ideology is to start with a theory's primitive ideology and then to "build up" its non-primitive ideology from it. In effect, the approach is to develop two separate criteria for primitive ideology and for non-primitive ideology that are individually sufficient and jointly necessary for the commitment to general ideology. The criterion for primitive ideology is mostly straightforward: a theory is ideologically committed to the primitive terms and concepts that are required for the theory to be stated in language and to be represented in thought. For non-primitive ideology, we may say a theory is committed to any term or concept that is definable in terms of its primitive ideology. By combining the two, we get:

GENERAL IDEOLOGY$_0$: Theory T is committed to a bit of (general) ideology i iff i is (i) a primitive bit of ideology that is required for the theory to be represented or (ii) definable by the primitives of T.

An issue we immediately run into is that there are two different ways in which a bit of ideology can be said to be "primitive." On the one hand, a bit of ideology may be *absolutely primitive* in the sense that it cannot be analyzed at all. On the other hand, it may be *primitive relative to a theory* in the sense that it cannot be analyzed within a theory in which it is employed.[24]

---

[24] The terminology for this distinction was suggested to me by Maya Eddon.

For instance, suppose a physicalist picture of the sciences is true and suppose that all high-level disciplines (e.g., chemistry, biology, etc.) are reducible to physics. On this picture, none of the "primitives of chemistry" are absolutely primitive because, *ex hypothesi*, any chemical term can be defined by physical terms. Instead, they are only primitive relative to chemistry.

Among the two notions, it is the latter that we should attend to. With the notion of absolute primitiveness, we may be able to define the general ideology of total theories—i.e., the conjunction of every theory one accepts in their attempt to represent reality in its entirety.[25] However, if we are to say that the commitments of one's total theory depends on the commitments of the individual theories it comprises (as the good Quinean should), we are primarily in need of a criterion that applies to non-total theories. And since we cannot plausibly expect that all theories have absolutely primitive ideologies (as suggested by the above physicalist's conception of chemistry), the criterion under development should be formulated in terms of relative primitiveness.

It is worth noting that absolutely primitive ideology and relatively primitive ideology bear different relationships to fundamentality. So far, I have treated fundamentality as an all-or-nothing matter and have not sharply distinguished between fundamental ideology and primitive ideology. But the introduction of relative primitiveness requires us to entertain a more nuanced picture of fundamentality that admits of degrees. It is evident that absolutely primitive ideology goes hand-in-hand with a notion of absolute fundamentality. But this is not so for relatively primitive ideology, since the physical terms that define the relatively primitive ideology of chemistry supposedly

---

[25] There are further issues regarding whether such a complete representation of reality is possible. See Dasgupta (2015) and also Hofweber (2005). But at the least, a total theory can be understood as an attempt to represent as much of reality as we are capable of representing.

exhibit a greater degree of fundamentality. For this reason, the relative primitive ideology of a theory can only be said to be more fundamental than the rest of its ideology.[26]

To avoid confusion, let us say that a bit of ideology I is *foundational* for a theory T iff (i) I is primitive relative to T, and (ii) I is required for T to be stated. With this notion, the previous first-pass criterion can be refined as:

GENERAL IDEOLOGY: Theory T is committed to a bit of general ideology I iff I is either (i) foundational for T or (ii) is definable by what is foundational for T (in conjunction with any term that is more fundamental than I).[27]

I submit GENERAL IDEOLOGY as the criterion for the commitment to general ideology. A consequence of this criterion, though, is that a theory's general ideology encompasses *anything* that is definable in terms of its foundational ideology regardless of whether or not they serve a real purpose. And some may worry that this over-generates its non-foundational ideology. After all, something being defin*able* is one thing; something actually being defined is another. For instance, as opposed to 'tiger', we may define 'schmiger' so that it applies to all tigers and Fido the dog; as opposed to 'electron', we may define 'schmelectron' so that it applies to all electrons and Priscilla

---

[26] A further question concerns how relative primitiveness relates to non-ontic structure. See Sider (2017, especially section 5) for further discussion.

[27] The parentheses are included to allow for supplementary terms to be pulled from outside of the discipline. Returning to the example of chemistry, these could be terms that describe various physical processes (e.g., the losing or gaining of an electron) that are involved in the definition of a chemical term.

the proton, etc. And it seems absurd to say that these terms are part of the vocabularies of biology and physics.

To be clear, the issue here has nothing to do with ideological simplicity. As previously mentioned, ideological simplicity is only a matter of primitives, or in light of our previous discussion, to what is foundational to a theory. So, at the end of the day, the inclusion of frivolously defined terms does not make our theories any more complex. Then what *is* the problem? One way of understanding the issue is that the criterion generates a vast amount of non-foundational ideology that are never relevant to our theoretical discourse. Understood this way, the worry is not so much about a defect of the criterion itself, but instead reflects the need of a supplementary account of when it is relevant to our theoretical discourse.[28]

One tempting approach to the task is to double down on graded fundamentality. Because of their arbitrary and disjunctive nature, frivolously defined terms are grossly non-fundamental. Thus, if it can be argued that a term needs to be at least "reasonably fundamental" to be relevant to our discourse, then we would have a simple way to account for their irrelevance.[29] For instance, one may argue that 'schmiger' and 'schelectron' are excluded from our biological and physical discourse because they are too disjunctive and thus are too non-fundamental to play any significant theoretical role; thus, they are excluded from our discourse precisely because they are frivolously defined.

---

[28] The suggestion here is similar to how counterfactual theorists of causation often respond to alleged cases of spurious causation. See Schaffer (2012) for a detailed, critical exegesis of the overall approach.

[29] The general approach here is related to Sider's (2017) response to the previously discussed criticism from Barnes and Mikkola. In his terminology, the question of "reasonable fundamentality" can be rephrased as "being sufficiently joint-carving." (Also see Sider 2011, ch.7, sect.11.) I thank Phil Bricker for pressing me on this line of response.

However, this approach cannot be made to work. An obvious question to ask is when a term fails to be "reasonably fundamental" to be part of our theoretical discourse. Right off the bat, there seems to be no single answer that applies across the board. If the standard is made too lax, then it will open the door to the frivolous terms we wish to exclude; but if it is made too strict, it will close the door to terms that are useful or even required by theories about non-fundamental matters. For example, compare 'schmelectron' to, say, 'cubism' (the art movement). While we may plausibly say that 'schmelectron' is unreasonably non-fundamental for physics, it not too far-fetched to think that it is nonetheless more fundamental than 'cubism', or any other term that is useful in disciplines higher up the ladder. Hence, we face a dilemma. Any strict standard that excludes 'schmelectron' from our discourse about physics would also have us exclude 'cubism' from our discourse about art; in the opposite direction, any lax standard on which 'cubism' is reasonably fundamental would not be able to exclude 'schmelectron' from our discourse about physics.

So, for the appeal to reasonable fundamentality to be fruitful, it must be argued that there are different standards for reasonable fundamentality that applies to different disciplines. But what could possibly determine what level of fundamentality is appropriate for which theory? One may hope to answer this question by appealing to intuitions about the fundamentality of entire theories, as opposed to the bits of ideology they posit. For example, consider the commonly held layered view of the sciences, with physics at the bottom, and chemistry, biology, psychology, etc. on top, in that order. Supposedly, this picture captures a sense in which the different scientific disciplines are fundamental to different degrees, with physics being the most fundamental of the lot, and the other disciplines being less fundamental than what comes before but more fundamental than what comes after. To avoid confusion, let us reserve 'fundamentality' for bits of ideology and use 'deeper

than' to refer to the comparative fundamentality of entire theories. I.e., on this terminology, we may say that chemistry is *deeper than* biology, but *less deep than* physics, and so forth. In this light, one may argue that what is "reasonably fundamental" for a given theory depends on how deep the theory supposedly is—the deeper the theory, the more fundamental its ideology needs to be. That is, one may hope to develop an account to the effect that:

REASONABLE FUNDAMENTALITY: A bit of ideology I is relevant to the discourse involving theory T iff (i) I is definable by the foundational terms of T and (ii) it is reasonably fundamental for T's depth.[30]

This much is mere terminology. To determine when a term is and is not reasonably fundamental for a theory, more needs to be said about what the 'deeper than' relation consists in. A way that comes immediately to mind is by way of how the terms of one theory are definable by the terms of another. For instance, we may define 'deeper than' as follows:

DEFINITIONAL DEPTH: Theory X is *definitionally deeper than* theory $Y =_{\text{def}}$ every term of Y is definable by the terms of X (or in conjunction with terms from a theory that can be used to define the terms of X) but not *vice versa*.

---

[30] For the discourse-exclusion problem, the suggestion is to take the depth of a theory to create a context that fixes how fundamental a term needs to be in order to be relevant to our discourse about it.

The intuitive idea here is that physics is "deeper" than chemistry in the sense that every chemical term is definable in physical terms, but not *vice versa*. (The purpose of the parenthetical clause is to allow for definiens being pulled from still deeper theories.) And this is indeed one sense in which some philosophers suggest that physics is deeper than (i.e., more fundamental than) chemistry, and so forth. However, this approach gives us no more guidance than before. What we need is a conception of depth that has useful implications for how fundamental a term needs to be for it to be relevant to a theory. And <u>Definitional Depth</u> remains silent on this.

For any progress to be made, there needs to be an ordering of the theories that reflects the comparative fundamentality of their terms. As a first pass, consider:

ORDERED DEPTH$_0$: Theory X is *orderly deeper than* theory $Y =_{\text{def}}$ every term of X is more fundamental than any term of Y.

Applied to the previous example of the stratification of the sciences, this entails that any chemical term is more fundamental than any biological term, and that no chemical term is more fundamental than any physical term, so on and so forth. For the sake of simplicity, suppose that a given term's level of fundamentality is simply a matter of definitional complexity.[31] The rough idea is that,

---

[31] This is in the spirit of how Lewis (1983) defines comparative naturalness. But Sider (2011, ch.7 sect.11) argues that we need to take more into consideration than definitional complexity. For instance, he argues that we also need to consider a term's connections to law-like generalizations, probability, unification, causality, etc. These issues are irrelevant to the point that I am making.

when defined in absolutely fundamental terms, the definition of any chemical term is more complex than any physical term, but less complex than any biological term, and so forth. However, this is not how we typically think the sciences are stratified. Take the set of all of our theoretical terms (i.e., the general ideology of one's total theory) and consider a partitioning P of it in terms of the *equally complex* relation. Next, place these partitions in an ordered sequence of increasing degrees of complexity. E.g. (say $P = (S_n)_{n=\mathbb{N}}$):

(*) $(S_1, S_2, S_3, S_4 \ldots)$

For the sciences to satisfy <u>Ordered Depth$_0$</u>, there need to be cut-off points along this sequence that demarcate the different disciplines. But it is implausible that there are any natural breaks to be found. We may of course plausibly expect that a biological term will have a more complex definition than *its chemical definiens*. But why think that the definitions of biological terms are always more complex than that of *any* chemical term? For instance, it seems possible for there to be a highly variegated chemical phenomenon the definition of which (when put in purely absolutely fundamental terms) is much more complex than such a definition of, say 'digestion'. If this is indeed possible, then it is hopeless to define the depths of our theories in this fashion.

To get around this problem, we may add restrictions to the terms of a theory that determine its depth. For instance, suppose we triple down on fundamentality and say that the depth of theories is only a matter of the relationship between their foundational terms. That is:

ORDERED DEPTH$_1$: Theory X is *orderly deeper than* theory $Y =_{\text{def}}$ the foundational terms of X are more fundamental than the foundational terms of Y.

The restriction to foundational terms avoids the aforementioned problem. But this time, we lose sight of the prize. Since this conception of depth only concerns the relationship between the foundational terms of theories, it implies nothing about a theory's non-foundational ideology. All this tells us is that the foundational ideology of physics is more fundamental than that of aesthetics, and this fails to have any implications for when a term is fundamental enough to be relevant to our discourse.

1.4.3 Toward a Dialectical Account of Relevance

I propose a different approach to the problem. Instead of taking the relevancy of terms to be a matter of some "internal" feature such as their degree of fundamentality, I suggest that we consider the roles that a term may play in the dialectic more generally. My proposed solution draws inspiration from a meta-theoretical point made by David Armstrong (1978; 1980). The point in question is made when Armstrong sets the stage for his argument for immanent universals. He writes: (Armstrong 1978, p.11, boldface mine):

> G. E. Moore never tired of emphasizing that, in the case of many of the great metaphysical disputes, the gross facts are not in dispute. What is in dispute, he contended, is the account or analysis to be given of the gross facts. This appears to be the situation in the dispute between Nominalism and Realism. Both can agree

30

that the paper is white and rests upon a table. It is **an adequacy-condition** of their analyses that such statements come out true.

What is interesting here is that Armstrong is deriving an adequacy condition that applies to *theories* by considering what is and what is not disputed by *theorists*. That is, on the grounds that the dispute between nominalism and realism about properties is not over whether objects have attributes or stand in relations, he argues that a theory of properties is adequate only if it is capable of expressing truths regarding such matters. This point is made more salient in his "Against 'ostrich' nominalism", in which he writes (re-quoted from Lewis 1983, p.351, boldface mine):

> I suggest that the fact of sameness of type is a Moorean fact: one of the many facts which even philosophers should not deny, whatever philosophical account or analysis they give of such facts. Any comprehensive philosophy must try to give some account of Moorean facts. They constitute **the compulsory questions** in the philosophical examination paper.

Although his language went through some changes, the overall spirit remains the same: on the grounds that *philosophers* should not deny facts about attribute agreement, he claims that it is "compulsory" for a *theory* that the philosophical account it provides is compatible with these facts.

The insight I take from Armstrong is that there are things we may demand from theories on the basis of what we may demand from theorists. There are interesting questions to ask about what to think about theories that fail to meet these kinds of demands. Most notably, one may ask whether a theory's failure to meet such a demand constitutes a reason to believe that that theory is false. I address this question in chapters 3 and 4 of this dissertation and argue that it does. But the point I am currently making does not require there being such a robust connection between dialectical matters and truth. For now, we only need to attend to the simple fact that there are features of the dialectical situation of theorists that shape the theories that they put forth. In the case of the properties debate, the relevant feature of the dialectic is that theorists are required to acknowledge Moorean facts about attribute agreement; because of this, the theories they put forth are expected to be able to produce and express true statements that report Moorean facts.

This suggests that, apart from the role ideology plays for *theories*, there is an additional dialectical role that it plays for *theorists*. And my solution to the over-generation problem is to define a relevance condition in terms of this dialectical role. Let us say a *dialectical obligation* is an obligation one takes up by virtue of participating in a given discourse. To paraphrase Armstrong's insight in these terms, we may say that philosophers have the dialectical obligation to acknowledge Moorean facts—that, if a given fact is deemed as "Moorean," then one's correct participation in a philosophical discourse requires one to acknowledge that that fact obtains. I add to this the following observation: one way to discharge the relevant obligation is to ensure that the ideology (and/or ontology) of one's theory is capable of stating these facts within the theory. This point can be generalized by noticing the fact that the acknowledgement of Moorean facts is only one dialectical obligation among many that can be addressed through the ideology of one's theory. Because it is likely that different kinds of discourses involve different kinds of dialectical

obligations, it is difficult to give a comprehensive survey of the matter while staying within the purview of this dissertation. But the general idea can be expressed as follows. Let us say that a theorist *dialectically commits* to a bit of ideology in the context of a discourse *D* just in case one employs it in order to discharge some dialectical obligation or other that is relevant to *D*. For example, one dialectical obligation that is particularly relevant to this dissertation (especially to chapter 4) is the obligation to defend the claims one makes in the context of a debate. In this case, one can be said to dialectically commit to the terms and concepts one employs in her arguments and responses in the course of defending her claims. Using this notion of dialectical commitment, we can take the terms relevant to a discourse to consist of the terms that its participants are dialectically committed to. That is:

DIALECTICAL RELEVANCE: A bit of ideology is relevant to a theoretical discourse *D* iff there is a participant of *D* that (i) participates in *D* as an advocate of a theory *T* and (ii) dialectically commits to *I* by virtue of advocating *T*.

This condition solves the over-generation problem by filtering out frivolously defined terms on the grounds that they are frivolous in a somewhat different sense from before. On my account, terms like 'schmiger' or 'schemelectron' are excluded from discourses about biology and about physics not because they are disjunctively or otherwise gruesomely defined, but because there is no dialectical purpose that we can reasonably expect them to serve regarding our discourse about the biological and physical nature of the world —because it is difficult to think of a dialectical obligation that may push the standard biologist or physicist to dialectically commit to

these kinds of terms. (They admittedly serve a dialectical purpose regarding our current philosophical discourse about the general ideology of biology and physics, but that is a different story altogether that they *do* happen to be relevant to.)

1.5 Summary

In this chapter, I argued for a pluralist view of 'ideology' that takes the term to be ambiguous between four different senses. On the one hand, it is ambiguous between linguistic and conceptual ideology. Although the two senses are not interchangeable, I argued that their relationship is intimate enough that, insofar as we are clear about which sense we are using, no serious harm ensues from the ambiguity. Regarding this distinction, I argued that the term 'ideology' should be reserved for linguistic terms and concepts and that it should not be used to refer to the "features of the world" (i.e., Siderian structure) that these terms and concepts characterize. On the other hand, the term is ambiguous between primitive and general ideology. While the notion of general ideology is not entirely new, the tendency to focus on primitive ideology has left it unexplored. To shed light on this notion, I explored two issues. First, I argued that the way we typically think about ideology can be repurposed into a methodological principle that is serviceable to doing non-fundamental metaphysics. To this end, I argued for ideological permissivism that bestows us with the freedom to introduce non-fundamental ideology as we see fit. Second, to cultivate a more general perspective of general ideology, I made efforts to develop a criterion for the commitment to general ideology. One obstacle that arose was that the criterion generates bits of non-fundamental ideology the majority of which are seemingly irrelevant to our

theoretical discourse. To address this issue, I argued for a relevance condition that appeals to

ideology's role in providing theorists the resources to discharge their dialectical obligations.

CHAPTER 2: IDEOLOGICAL REDUCTION

2.1 Introduction

It is common to distinguish between *ontological reduction* on the one hand and *conceptual* or *analytic* reduction on the other. The primary concern of this dissertation is the latter type, which I shall call *ideological reduction* in what follows. The paradigmatic examples include the various counterfactual analyses of causation, the Block Universe Theory of time, and Lewisian modal realism.[32] Ideological reduction differs from ontological reduction in two main ways. First, the main goal of ideological reduction is not necessarily to shed light on *things*—entities that exist in the world. Instead, the goal is to shed light on the terms and concepts we use to express the various aspects of the world such as causality, temporality, and modality. Admittedly, there is a sense in which terms and concepts themselves are entities and there we may at times speak of the ontology of terms and concepts. But this is obviously not what the above theories are concerned with. Rather, the reduction concerns the content of these terms and concepts—that which they express. And as discussed in the previous chapter, the content of ideology need not be something ontic. Second, ideological reduction has a distinctive relationship to theoretical simplicity. Any claim of ideological reduction entails the rejection of primitivism; the paradigmatic examples respectively purport that all bits of ideology concerning causality, temporality, and modality are ultimately analyzable. Thus, other things being equal, we may expect a successful ideological reduction to lead to greater ideological simplicity. And according to many, ideological simplicity is largely an independent matter from ontological simplicity. In fact, it is often the case that ideological

---

[32] For causation, see Lewis (1973; 2000); for temporality, see Skow (2015); for modality, see Lewis (1986).

reduction commits theories to more expansive ontologies than their non-reductive alternatives, e.g., the Block Universe Theory in comparison to presentism and Lewisian modal realism in comparison to modalism.

Clarification: I do not take these remarks to entail that the distinction between ontological and ideological reductions is substantive and fundamental (in some sense or other), nor that it is clear-cut. Perhaps it can be shown that one type of reduction ultimately collapses into the other. Or perhaps, for various reasons, a given reductive procedure can be said to be ontological in one light, but to be ideological in another; or maybe there are various ways in which the one type of reduction is always accompanied by the other in some fashion. I leave these matters open. What I do assume is that there are some commonalities between the paradigmatic examples of ideological reduction, which may or may not be shared by ontologically reductive theories, and that there are interesting and valuable lessons to be learned by investigating them.[33] To this end, I take the apparent differences between ideological and ontological reduction at face value.

Traditionally, reduction in general has been closely associated with identity. In this chapter, I explore what this entails specifically for ideological reduction. In Section 2.2, I discuss a puzzle that arises for traditional, identity-based views of reduction and consider how it applies to ideological reduction. In Section 2.3, I argue against the attempt to solve this puzzle by disassociating reduction from identity and instead characterizing it in terms of metaphysical grounding. In section 2.4, I provide my proposed solution to the puzzle.

---

[33] Another orthogonal issue concerns how all this relates to the reduction of *theories*, e.g., the reduction of Newtonian mechanics to Special Relativity, and/or the reduction of chemistry to physics.

2.2 Ideological reduction and identity

Let us call the traditional, identity-based view of reduction *the identity view*. For our purposes, any view that takes reduction to involve an identity claim in some form or other counts as an identity view. In the case of ontological reduction, the identity claim is typically made with respect to the relata of reduction. Consider some common examples: it is often suggested that the physicalist reduction of the mind entails that the mental state of pain *just is* the physical state of one's c-fibers firing. Other examples in the same vein are plentiful: many take the reduction of water to H_2O to entail that water $= H_2O$, the reduction of the natural numbers to sets (either in terms of Zermelo ordinals or von Neumann ordinals) to entail that $0 = \emptyset$, and so forth. However, this does not appear to carry over to ideological reduction. Applied to ideological reduction, the thought would be (let the following be ambiguous between linguistic and conceptual ideology):

IDENTITY OF RELATA: If A ideologically reduces to B, then $A = B$.

We can immediately see that the thesis is grossly implausible for the reduction of linguistic ideology. For any serious lexical analysis, we do not expect the definientia and definiendum to involve identical terms. If anything, it should be the *meanings* of these terms that are identical, not the terms themselves. Thus, we cannot plausibly say that the reduction of linguistic ideology entails the identity of its relata.

Granted that successful lexical analyses involve terms with the same meaning, one may hope that this wrinkle to be ironed out at the level of concepts: if one takes concepts to serve as

meanings, then one may maintain that the analysans and analysanda of the corresponding conceptual analyses involve identical concepts. Accordingly, one may be able to explain away the oddity regarding linguistic ideology by taking the proper relata of ideological reduction to be concepts instead of terms and by taking a term to reduce to another only in some derivative sense. However, there are worries that remain even if we retreat to the level of concepts. To maintain IDENTITY OF RELATA, we need a sufficiently coarse-grained view of concepts that individuates them by their extensions. But it seems desirable for our theory of concepts to take them to be more fine-grained than that. Consider a (broadly) Fregean view of concepts that identifies concepts with Fregean senses.[34] It is well-known that, for senses to provide a solution to Frege's puzzle of co-referring terms, they must be individuated by their cognitive significance. The stock example of the puzzle is: while it is true that Mark Twain = Samuel Clemens, the names that flank '=' license different patterns of inferences. For instance, unless the identity claim is given as a premise, one cannot directly infer (1) from (2):

(1) Mark Twain wrote *The Adventures of Tom Sawyer*.

(2) Samuel Clemens wrote *The Adventures of Tom Sawyer*.

The Fregean account of this phenomenon is that, although the terms 'Mark Twain' and 'Samuel Clemens' are co-extensional, they nonetheless have different meanings by virtue of being associated with different senses. And for this account to work, senses clearly need to be individuated in terms of their cognitive significance: i.e., in terms of their "impact on the mind."[35]

---

[34] Terminological note: because Frege used the term 'concept' to refer to functions from objects to truth values, the "broadly Fregean view of concepts" should be sharply distinguished from Frege's own view of concepts.

[35] Among the plethora of resources of the topic, see Margolis & Laurence (2007, p.566) for a summary of the (broadly) Fregean view (which they call the 'Semantic View of Concepts') that contrasts it to its major rival, namely the Psychological View that takes concepts to be mental representations. Also, see Nelson (2008) for how the Fregean view individuates concepts raises tension with Frege's solution to the paradox of analysis. Nelson's discussion is particularly relevant to our discussion immediately below.

Thus, if one identifies concepts with Fregean senses, the same goes for concepts as well. Now consider what this implies for IDENTITY OF RELATA. For a given conceptual analysis to be informative and explanatory, there must be some cognitive difference between the analysans and analysandum. But on the (broadly) Fregean view, this entails that, contrary to IDENTITY OF REALTA, the expressions that appear in the analysans and the analysandum express distinct concepts.

The point here generalizes beyond the Fregean view. For our theory of concepts to be able to solve Frege's puzzle, we must take concepts to be at least as finely-grained as Fregean senses.[36] This suggests that the difficulty here cuts deeper than what we had started with. Initially, we started with the question of whether ideological reduction entails the identity of its relata. But the considerations above have brought us to something much like the paradox of analysis.[37] A key step in the above line of reasoning is that an analysis is informative and explanatory only in case the analysans and the analysandum exhibit some difference in cognitive significance. This can be seen as an instance of a more general prohibition of circular explanation, i.e.:

NON-CIRCULARITY: Nothing explains itself.

---

[36] The situation may even be worse on a view that takes concepts to be mental representations (i.e., the Psychological View), as mental representations are often suggested to be even more fine-grained than Fregean senses; on the face of it, different people can have different mental representations of the same Fregean sense. See Margolis & Laurence (2007, p.567-568) for how this observation is situated in the dialectic between defenders of the Fregean view and the Psychological View.

[37] The issue here is, in some sense, narrower than the paradox of analysis, since, as opposed to reduction, not all analyses purport to provide metaphysical explanations. E.g., the analysis "*x* is a vixen iff *x* is a female fox" does not appear to reveal any profound truths about the world, as opposed to our conceptual scheme of zoology. In this light, various kinds of metalinguistic solutions to the paradox may not easily carry over to our case. (See Daly (2010, ch2.6) for a brief overview of the metalinguistic approach to the paradox of analysis.)

And NON-CIRCULARITY seemingly raises tension with any version of the identity view. To see this, let us not worry about concept identity and leave open what it is that the relevant identity relation holds between. In any case, it is reasonable to expect any version of the identity view to accept the following, open-ended claim about the connection between reduction and identity:

GENERAL REDUCTIVE IDENTITY: If concept A reduces to concept B, then there is some X and some Y such that the reduction of A to B consists in the fact that $X = Y$.

On the aforementioned, coarse-grained view of concepts, we may take the variables 'X' and 'Y' to stand for the concepts A and B themselves. But other options are available. For instance, one may take the identity relation to hold not at the level of concepts, but at the level of propositions that "contain" (in some suitable sense) the concepts in question. This is how Gideon Rosen sets the identity view up for evaluation (Rosen 2009, p.122, p.124, italics in original, note: Rosen uses 'fact' to refer to true propositions):

[W]ithout significant distortion we may think of reduction and real definition as a relation among *propositions* that contain the target items as constituents […] To be a square just is to be an equilateral rectangle, let us suppose. […] [I]f our definition

41

of square is correct, then surely the fact that *ABCD* is a square and the fact that

ABCD is an equilateral rectangle are not different facts: they are one and the same.

It is also how Ted Sider describes what he takes the reduction of modality to entail (2003, p.7, all

emphases in original):

> A reduction of modality is an analysis of modal propositions in certain other terms.
> By analysis I mean identity: if the proposition that **it is necessary that everything**
> **colored is extended** is analyzed as the proposition that *p*, then the proposition that
> **it is necessary that everything colored is extended** *just is* the proposition that *p*.

Whether we take the proper relata of ideological reduction to be propositions instead of concepts

is beside the point. Whatever you say about this, the association with identity raises tension with

the common suggestion that reduction relations are explanatory. On the identity view, the

explanatory work is supposedly carried out by the revelation of identity relations. For instance, the

identification of pain states with certain physical states by itself is meant to reveal deep and

profound truths regarding *what pain is*, etc. (To put the point in terms of tokens, we may also say

that the identification of *that* pain state with *that* physical state reveals *what that pain state is*.) In

this light, we may say:

REDUCTIVE EXPLANATION (IDENTITY): If concept *A* is reduced to concept *B*, then, for any *X* and *Y* the identity of which is implied by the reduction, *X* is explained by *Y*.

The problem is that the three theses above appear to be jointly inconsistent. Suppose concept *A* is reduced to *B*. By REDUCTIVE IDENTITY, this entails that there is some *X* and *Y* such that the reduction consists in *X* = *Y*. By REDUCTIVE EXPLANATION (IDENTITY), this entails that *X* is explained by *Y*. But since *X* = *Y*, it seems to follow that *X* is explained by itself, which contradicts NON-CIRCULARITY. In short, in order to uphold REDUCTIVE IDENTITY, it seems that we must either allow for circular explanations or deny that reduction is explanatory, neither of which is palatable. Let us call this *the puzzle of reductive identity*. In this chapter, I do not explore every possible solution to this puzzle. Instead, I put forth what strikes me as a plausible solution and address what I take to be its major rival.[38]

Gideon Rosen argues that, on the pain of this puzzle, we should give up on the identity view. Instead, he argues that reduction should be understood in terms of metaphysical grounding.[39,40] Thus, he proposes the "grounding-reduction link" which, roughly speaking, says:

---

[38] Most notably, for reasons of space I do not consider solutions that reject NON-CIRCULARITY. While very few openly accept instances of blatantly circular explanations of the form '*X* explains *X*', there are theorists who deny that explanation is asymmetric. For instance, Mari Mikkola (2015) and Naomi Thompson (2016) argue for the non-symmetry of grounding and allows for cases of metaphysical interdependence. In the philosophy of science literature, Baas van Fraassen (1987) argues for a pragmatic theory of explanation that also (in some sense) rejects the asymmetry of explanation. (The relevant sense is that, for van Fraassen, whether *X* explains *Y* depends on the context of inquiry, and while *A* may explain *B* in some contexts, the direction of explanation may be reversed in other contexts.)

[39] For what it's worth, this distorts Rosen's narrative to some degree. Rosen's primary goal is to survey various ways in which grounding is service*able*; thus, his case is mostly presented as a means of preserving an intuitive connection between reduction and grounding whereas I describe him as reacting to a puzzle from a more neutral position. I take this to be inconsequential.

[40] In effect, Rosen rejects GENERAL REDUCTIVE IDENTITY and reformulates the thesis for reductive explanation in a way that drops the identity claim.

THE GROUNDING-REDUCTION LINK: If *A* reduces to *B*, then *A* is grounded by *B*.[41]

Let us call any view that accepts this link in some form or other *the grounding view*. (I spell this out in a bit more detail below.). Although the grounding view is an interesting avenue to explore, I believe that Rosen is too quick to dismiss the identity view. Furthermore, I argue that there are important ways in which reductive claims differ from grounding claims, and so that we should not take reduction and grounding to be so connected. In what follows, I raise some criticisms for the grounding view and later suggest a way for the identity view to resolve the apparent tension.

Before going forward, there are two verbal issues that are worth mentioning. In this chapter, I treat the grounding view and the identity view as incompatible, rival views of reduction. But it is noteworthy that the terms 'grounding' and 'reduction' are terms of art that can mean different things in different contexts. For instance, on a sufficiently broad sense of the term 'grounding', it is not too implausible to take the identity view to fall under the grounding view. That is, if one follows Carrie Jenkins (2011) and rejects that grounding is irreflexive, then there may be no real obstacle for taking identity to be a specific *kind* of grounding relation; one may simply take reduction to be a case of reflexive grounding.[42] Similarly, on a sufficiently broad sense of 'reduction', the grounding view and the identity view can be understood as expressing different

---

[41] All of the theses for the grounding view I present should be understood as involving *full grounding* rather than *merely partial grounding*, as anything short of it would appear inadequate to capture claims of reduction. This becomes especially relevant in Section 3.2.

[42] In a similar vein, one may follow Jessica Wilson (2014) and take identity to be one among many specific "small-g" grounding relations.

conceptions of reduction.[43] In this light, my concerns are not so much about associating reduction with grounding, but instead pertains to disassociating reduction with identity. That is, my key claim is that there are important differences between putative cases of reduction (e.g., our leading examples) and of grounding (e.g., that conjunctions are grounded by their conjuncts, that determinables are grounded by determinates, etc.). To maintain a sharp contrast, I assume that grounding is irreflexive, and that it has the other standardly assumed features, namely that it is a primitive, explanatory relation that is also asymmetric, and transitive. For expository purposes, I follow Rosen and treat the grounding relation to hold between Russellian propositions. But this is mostly for the sake of exposition, and the conclusion I arrive at can mostly be translated into other modes of grounding-talk in the usual way. (I make note of when the relevant choices matter.) [44]

A second point is that, despite what I say immediately above, the transitivity of grounding is not set in stone. The main drive behind the assumption of transitivity is that it allows us to use grounding to describe the metaphysical structure of the world. However, even if you think grounding is intransitive, the same can be achieved by employing its transitive closure.[45] Because determining whether grounding itself is transitive goes beyond the purview of this dissertation, I do not directly engage with this matter. Instead, I take the transitivity assumption with a grain of salt: if a given issue can be resolved by employing the transitive closure of grounding, I consider it not to raise tension with the transitivity assumption. (This will become especially relevant near the end of section 2.3.2.)

---

[43] The possibility of this equivocal view of reduction was suggested to me by Alejandro Pérez Carballo and Ned Markosian on separate occasions.
[44] See Correia and Schnieder (2012).
[45] This point is made by Schaffer (2012, p.125). Similar remarks are found in Rosen (2010, p.116).

2.3 Against the Grounding View

2.3.1 The Grounding View

Let us first flesh out the grounding view in more detail. As mentioned before, I only consider how ideological reduction is to be understood on the grounding view and leave open what it entails for ontological reduction. For our purposes, we may define the grounding view as (let 'X-proposition' stand for 'proposition that contains concept X'):

PROPOSITIONAL REDUCTIVE GROUNDING: If concept A reduces to concept B, then all A-propositions are grounded by some B-proposition.

But similar to before, I take any view that accepts the following to be a version of the grounding view:

GENERAL REDUCTIVE GROUNDING: If concept A reduces to concept B, then there is some X and some Y such that the reduction of A to B consists in the fact that X is grounded by Y.

Because grounding itself is an explanatory relation, this trivially entails the following revision of reductive explanation:

REDUCTIVE EXPLANATION (GROUNDING): If concept *A* is reduced to concept *B*, then, for any *X* and *Y* the grounding relationship of which is implied by the reduction, *X* is explained by *Y*.

Due to the irreflexivity of grounding, the view easily avoids the puzzle of reductive identity. Since nothing grounds itself, it entails that the non-identity of the relevant propositions. This is made explicit by Rosen. Soon after the previously quoted passage, he continues to say (italics his, boldface mine):

> [The violation of irreflexivity] comes from our commitment to the thesis that facts and propositions are individuated by their worldly constituents and the manner of their combination. For surely, the *property* of being a bachelor just is the *property* of being an unmarried male (if the analysis is correct). […] We can resist this line of thought by insisting that the operation of replacing a worldly item in a fact with its real definition never yields the same fact again. **It yields a new fact that 'unpacks' or 'analyzes' the original**.

One prominent obstacle for the grounding view I shall ignore are those rooted in skepticism about grounding itself. For instance, those who are generally skeptical about the philosophical

value of grounding will obviously be doubtful about the value of characterizing *anything* in terms of grounding (e.g., Wilson 2014). I set these kinds of worries aside and instead argue that even the friends of grounding should be wary about blurring the line between reduction and grounding.

2.3.2 Explanatory Deference

On the face of it, there are indeed several features that grounding has in common with reduction. Superficially, they appear to share many formal features. It is desirable that our notion of reduction comes out as irreflexive, asymmetric, and transitive. And as previously mentioned, grounding is often said to have these very features. But the association is probably more deeply rooted in each relation's connection to explanation. The reduction of a given subject matter supposedly yields a complete explanation of it. And we expect nothing less from a complete description of its grounds. That is, if the grounding-reduction link is understood in terms of *full grounding* as opposed to *mere partial grounding*, the view appears to be able to capture the idea that the reduction base yields a complete explanation of the subject matter. However, a closer look reveals some important differences between the explanatory structures that each relation implies. My central claim is that reduction exhibits a "deference of explanation" whereas grounding does not. And because of this, I argue that it is best to distinguish reduction claims from grounding claims more sharply.

Here is what I mean by explanatory deference. When a reduction is successful, it seems utterly wrong to maintain that the reduced facts nonetheless fully retain their explanatory value. For instance, if the reduction of modality is successful, it would seem that all of the explanatory

work we had previously assigned to the modal facts is carried out by the non-modal facts that make up the reduction base. Thus, we lose nothing metaphysically significant by directly employing the requisite non-modal facts to do the work of the modal facts: if the reduction of modality is successful, then there is nothing that the modal facts *and only the modal facts* bring to the table. To say otherwise is to insist that there is an aspect of modality that is not fully captured by what is put forth as the reduction base. And this is tantamount to denying that the proposed reduction is successful.

In contrast, the grounding relation itself does not exhibit this kind of behavior. One way to appreciate the point is to consider an infinite series of grounding. Consider (take '$\leftarrow$' to mean 'is grounded by'):

(S) $F_1 \leftarrow F_2 \leftarrow F_3 \leftarrow F_4 \leftarrow \ldots$

At first glance, grounding structures like (S) may seem to involve a vicious infinite regress. The general idea is that, unless the series can be shown to come to an end, no explanation is ever achieved. For a concrete example, consider a grounding view of composition according to which composite objects are grounded by their proper parts. Call this view 'BUMP' (short for 'bottom-up'). It is commonly suggested the possibility of "gunky" worlds—worlds in which every object has a proper part—spells trouble for BUMP. This is because BUMP says the grounding structure of gunky worlds is like (S), with every compositional fact starting off an infinite chain of grounding. And since there is no fundamental layer from which composite objects are ultimately built, the thought goes, none of the compositional facts are properly grounded.

However, this simple suggestion turns out to be mistaken. The above line of reasoning appears to conflate a dependency between facts with a dependency between explanations. But we

can have the one without the other. For each of the explanations in (S) to be in jeopardy, the explanation at each step must be deferred to or be dependent on the next explanation in the series. That is, there is trouble only in case the explanatory value of F_1 is held hostage by F_2, which is in turn held hostage by F_3, and so forth. But it is unclear that this is so. The point is well put by Ricki Bliss (2013). Bliss argues that an infinite series of explanations is vicious only in case each step demands an answer for the very same question. In case each step demands an answer for a different question, the explanations at each step are independent from the explanations down the line and we simply have a benign infinite series of satisfying explanations.[46] (We'll call this 'Bliss's Point' from hereafter.) And whether the same question arises at each step depends on the question we're asking. For instance, if the question is "What explains F_1?", we are presented with a complete answer, namely "It is explained by F_2." For there to be any trouble, it must be that the question we are asking cannot be fully answered by referring to the facts that appear within (S), e.g., why we have the entire grounding chain at all.[47]

Applied to BUMP, gunky worlds cause trouble for the view only in case the explanation for the composition of each object depends on the explanations for the composition of their proper parts. And whether this is the case depends on the kind of explanation we're seeking. Suppose (S) is the result of asking what explains the composition of each object in gunky worlds. In this case, each step in (S) gives an answer to a different question. We first ask "What explains F_1?" and get the answer "F_2 does."; we then ask "What explains F_2?" and get the answer "F_3 does.", and so forth. And if BUMP only requires that every composite object is explained by their proper parts,

---

[46] A similar observation is made by Ross Cameron (2015, Chapter 2) in the context of the metaphysics of time and McTaggart's regress.

[47] Bliss's point has interesting implications for metaphysical foundationalism—the thesis that there is an absolutely fundamental layer on which the rest of reality is grounded. A central motivation for metaphysical foundationalism is the (supposed) absurdity of infinite chains of grounding. But Bliss points out this line of thought assumes that *being* or *existence* itself is in need of explanation, which she argues is akin to appealing to the Principle of Sufficient Reason.

then gunky worlds do not raise any problems for BUMP. For there to be a problem, (S) must be the result of asking a much more general question, namely *how there's any composition at all*. But this is not something that BUMP sets out to answer. BUMP is a view about the grounds of composite objects, not the grounds of *the operation of composition*, or *composition as such*.

The take-away of all this is that grounding does not inherently defer explanations: there is no apparent reason to think that the explanatory value of a given fact is dependent on its grounds. This makes grounding unsuitable for characterizing reduction. To further illustrate the point, consider what the grounding view implies for the transitivity of reduction. Suppose that some biological concept B reduces to some (complex) chemical concept C, and that C reduces to some (complex) physical concept P. Because reduction is transitive, this implies that B reduces to P. In this case, I argue that the explanatory value C has for B is deferred to P. In terms of propositions, I mean that there is nothing about the B-propositions that can only be explained by the C-propositions but not the P-propositions. Once we run the explanation in terms of the P-propositions, there should be nothing new that the C-propositions add to the explanation of the B-propositions. But consider what the grounding view implies for the case at hand. As previously mentioned, the grounding view is best understood as characterizing reduction in terms of full grounding. Thus, on the grounding view, our suppositions so far imply that:

(3) The B-propositions are fully grounded by the C-propositions.

(4) The C-propositions are fully grounded by the P-propositions.

(5) The B-propositions are fully grounded by the P-propositions.

Because grounding is assumed to be transitive, the grounding view has little trouble capturing the formal transitivity of reduction: just like how the reduction of B to C and the reduction of C to P

entail the reduction of *B* to *P*, (3) and (4) entail (5). However, (3)-(5) do not seem to exhibit a deference in explanation. Instead, we instead have a case of metaphysical overdetermination. Plausibly, (3) and (5) respectively imply that:

(6) The *B*-propositions are completely explained by the *C*-propositions.

(7) The *B*-propositions are completely explained by the *P*-propositions.

This suggests that the explanations that run from the *C*-propositions to the *B*-propositions are independent of the explanations that run from the *P*-propositions to the *B*-propositions. Since the *B*-propositions are *completely* explained by the *C*-propositions, the explanation from the *C*-propositions is not in need of any aid or supplementation from the *P*-propositions. Thus, whatever it is that the *P*-propositions are doing for the *B*-propositions, it has nothing to do with what the *C*-propositions are doing for the *B*-propositions. The same point holds in the opposition direction as well: given that the *P*-propositions completely explain the *B*-propositions, there is nothing left for the *C*-propositions to explain. In this fashion, it follows that the *C*-propositions and the *P*-propositions overdetermine the *B*-propositions in the sense that they serve as two independent sources of explanation for them.

To be clear, metaphysical overdetermination may not itself be a serious problem for a theory of grounding. For instance, it is often suggested that a disjunction that only has true disjuncts is fully grounded by each of its disjuncts. E.g., if *P* and *Q* are both true, it seems plausible that each of them is a full ground of (*P* v *Q*). And perhaps there is room to argue that the case at hand is no different from this. What *is* a problem, however, is that metaphysical overdetermination is nothing like explanatory deference. What we need is a story according to which (5) takes priority over (3)—that what's "really" doing the work is the *P*-propositions rather than the *C*-propositions.

But the story we arrived at is that there are special explanatory relationships that run between the *B*-propositions and the *C*-propositions that are independent of the *P*-propositions. Thus, (3)-(5) fails to tell the correct story of what is going on.

To avoid this result, the grounding theorist might try an alternative approach. In order for the grounding theorist to deny that (3) and (5) constitute a case of metaphysical overdetermination, she must deny that the *C*-propositions and the *P*-propositions both "fully ground" the *B*-propositions in the same sense. To this end, the grounding theorist may take the one sense to be derivative of the other and argue that it is only the non-derivative sense of full grounding that has any explanatory connotations. However, the more plausible way to achieve this is again at odds with explanatory deference. The most natural way to distinguish the notions relevant to (3) and (5) is in terms of *immediate* and *mediate* grounding. Using this distinction, (3) and (5) can be replaced by:

(3*) The *B*-propositions are *immediately* fully grounded by the *C*-propositions.

(5*) The *B*-propositions are *mediately* fully grounded by the *P*-propositions.

And if a choice must be made, it is more plausible to take mediate grounding to be derivative of immediate grounding. For instance, Kit Fine (2012, p.51) writes:

> Mediate ground can be defined in terms of immediate grounds. For all relationships of mediate ground can be obtained by appropriately chaining relationships of immediate ground. Thus each relationship of immediate ground can be taken to be a degenerate case of a mediate ground… […] The notion of immediate ground

53

would appear to be [giving] us something genuinely new; and I find it remarkable how strong our intuitions are about when it does and does not hold. […] It is the notion of immediate ground that provides us with our sense of a ground-theoretic hierarchy.[48]

And this gives us the wrong results. Instead of having the explanatory value of the *C*-propositions be deferred to the *P*-propositions, it affirms the explanatory value of the *C*-propositions at the expense of that of the *P*-propositions. And this is the opposite of what we were looking for.

2.3.3 Non-Fundamental Entities

The above criticism holds for any grounding-based conception of reduction. But there are also some worries that arise for the specific framework Rosen adopts. The irreflexivity of grounding naturally leads the grounding theorist to draw an ontological distinction between

---

[48] In the omitted section of the passage, Fine provides an argument for why the definition cannot go the other way around. However, his discussion of the matter seems misleading. There is certainly something intuitive about the idea that mediate grounding is derivative of immediate grounding. But it does not seem that Fine gives any further support for this intuition. His argument goes like this. Immediately before the quoted passage, Fine argues that an immediate ground should be understood as a ground that "need not be seen to be mediated" rather than "is not mediated." This is said to be because it is possible for a fact to be both an immediate and mediate ground for another. Consider the fact that *A* on the one hand, and that *A* v (*A* v *A*) on the other. By virtue of being the left disjunct of the whole disjunction, *A* is plausibly an immediate ground of it. But because the whole disjunction is also immediately grounded by (*A* v *A*), and since (*A* v *A*) is immediately grounded by *A*, it also follows that the whole disjunction is mediately grounded by *A*. The problem for defining immediate grounding in terms of mediate grounding is that, according to Fine, it is hard to see how such a definition can capture the idea that *A need not be seen* as a mediate ground of the whole disjunction. However, this argument seems circular. By taking this to show that an immediate ground "need not be seen to be mediated" (whatever he means by this), Fine seems to presuppose that, when a fact *X* is both an immediate and a mediate ground for some other fact *Y*, the fact that *X* immediately grounds *Y* is independent or prior to the fact that it mediately grounds *Y*. But this is to already presuppose that mediate grounding is derivative of (or at the least, less fundamental than) immediate grounding, which is the very intuition that he sets out to support.

fundamental and non-fundamental entities (or *less* fundamental relative to their grounds). There are several ways this distinction can be cashed out. (The different ways need not be incompatible with each other.) One is to draw the distinction at the subpropositional level. For instance, suppose $p = <a \text{ is } F>$, $q = <b \text{ is } G>$, and that p grounds q. It is common for the grounding theorist to maintain that the fact that p grounds q entails that (i) $p \neq q$, and (ii) that p is more fundamental than q. In this case, one may take (i) and (ii) to consist in the facts that $a \neq b$ and that a is more fundamental than b.[49] Alternatively, one may instead take (i) and (ii) to consist in the facts that $F \neq G$ and that F is more fundamental than G.[50] A third option is to draw the distinction at the level of worldly facts, e.g., something like Armstrongian states of affairs. Given the above supposition, one may maintain that $a = b$, but nonetheless maintain that $[a \text{ is } F] \neq [b \text{ is } G]$, and that $[a \text{ is } F]$ is more fundamental than $[b \text{ is } G]$, e.g., by virtue of 'F' and 'G' corresponding to distinct universals one of which is more fundamental than the other.[51,52] Let us call a view that takes any of these routes the *worldly view of grounding*.[53] In any of these cases, the view entails a substantive distinction within ontology.

A general worry for the worldly view of grounding is that it leads to a significant increase in ontological complexity. But let us suppose that Schaffer's ontological permissivism discussed in the previous chapter successfully dispels this worry. Nevertheless, there are concerns that remain.

---

[49] This is close to Schaffer's (2009) view except that he takes the relata of grounding themselves to be individual entities rather than Russellian propositions.

[50] This best fits how Rosen describes the examples he uses to motivate the grounding-reduction link.

[51] To remain more faithful to Armstrong's original sparse view of universals, one may reject that the universals themselves come with different degrees of fundamentality and instead maintain that predicates express properties that are fundamental to different degrees by virtue of how they are defined in terms of sparse universals. (See Armstrong 1978, ch6.)

[52] Again, similar views have been expressed by Ned Markosian (1998), Paul Audi (2012), and Kelly Trogdon (2013) whose cases also do not go through Russellian propositions.

[53] The individual-based version and the states-of-affairs-based version of the worldly view are not necessarily mutually exclusive, as one may take the non-identity of [a is F] and [b is G] to consist in either a ≠ b or F ≠ G, or both.

To better appreciate the point, let us take a closer look at Schaffer's ontological permissivism than before.

Schaffer's case for ontological permissivism consists of two main pillars. One is his proposal to replace the traditional Ockham's Razor, which prescribes not to multiply entities without necessity, with what he calls The Laser, which prescribes not to multiply *fundamental* entities without necessity (Schaffer 2015). As mentioned above, let us suppose that The Laser successfully dispels the worry of ontological simplicity and thus that non-fundamental entities, either understood as individual entities or states of affairs, are irrelevant to a theory's ontological economy. Even so, this is not the only reason to be wary about positing non-fundamental entities as distinct existents from their grounds. This is because the postulation of non-fundamental entities still incurs a dialectical burden, if not an ontological cost. In essence, The Laser is a methodological principle that helps us choose between theories when all else is equal. And one would be deeply confused to think that a free pass on economy allows one to posit non-fundamental entities on a whim. Positing these entities is to claim that these entities exist. And like any other claim, some sort of justification must be given for them to be legitimately made.[54]

This leads us to the second pillar of ontological permissivism: Schaffer's appeal to Mooreanism. Schaffer's permissivism is part of a larger project that aims to replace the traditional Quinean method of metaphysics. According to Schaffer, the traditional Quinean is fixated on existence questions and asks whether there are numbers, properties, fictional characters, etc. However, Schaffer argues that this is a mistake and takes there to be simple arguments that easily

---

[54] In chapter 4, I call this obligation to justify one's claims 'The Defense Obligation' and discuss a number of issues that concern how it should be understood. The Defense Obligation also plays a central role in the main argument of chapter 4.

answer these questions in the affirmative (Schaffer 2009, pp.356-359). For instance, from the "mathematical truism" that there are prime numbers, it follows that numbers exist; from the "everyday truism" that there are properties you and I share, it follows that properties exist, and so forth. For each of these arguments, he claims that the premises command "Moorean certainty" and that they are far more credible than any philosophical argument against them. From this, he concludes that the Quinean tradition has "confused itself with trivialities" (Schaffer 2009, p.361) and has led us to ask *whether* these entities exist when we should have been asking *how*.

A centerpiece of Schaffer's overall approach to metaphysics is the idea that there is a whole host of existence claims that can be justified on Moorean grounds. That is:

MOOREANISM: If there is a Moorean basis for the claim '*X* exists.', then one is justified in claiming that *X* exists.[55]

And one might think the appeal to Mooreanism is able to cover any non-fundamental entity the grounding view has us posit. However, this is not the case. An immediately obstacle is that the grounding view requires us to posit a *type* of entity that hardly enjoys any Moorean support. As previously mentioned, Rosen takes grounding to hold between facts, which he takes to be Russellian propositions. And Russellian propositions are not the kind of entity that enjoy any substantive Moorean support. The same goes for any other conception of facts, e.g., Armstrongian

---

[55] As Schaffer leaves room for there being existence questions that do license the Quinean approach, Mooreanism should be understood as only providing a sufficient condition for justification.

states of affairs. Armstrongian states of affairs are also theoretical posits that should be esoteric by the lights of commonsense, either by virtue of them being partially constituted by universals or of their connection to the metaphysical tie of instantiation. Either way, they are not the kinds of entity that enjoy any Moorean support.

To get around this problem, the grounding theorist could opt for a principle that does not involve any particular conception of facts. One way to achieve this is to refine the principle so that it prescribes a permissive stance toward entities that are *picked out* by uncontroversial concepts while remaining silent on the type of entity we are positing. For instance, it is plausible that Moorean certainty commands us to affirm the very general claim that certain things are possible, that certain things cause other things, and that there are things that were, is, and will be the case. Following this thought, we may recast the thesis so that it can work in conjunction with types of entity that are acceptable on independent grounds. Here is one way the refined principle might go:

TOKEN MOOREANISM: If there is a Moorean basis for concept *C*, then, for any type of entity *K* that a theory justifiably posits, one is justified in claiming that the tokens of *K* that are picked out by *C* exist.

The modified principle appears to sit well with the three reductive theories we've been using as our guiding examples. However, it is still unfit for a *general* account of reduction. It just so happens that our primary concern lies in theories that target concepts that enjoy Moorean support. But we must not forget that reduction, in its most general form, is a philosophical tool that

can be used for various purposes, not all of which involve uncontroversial notions. For instance, consider Alvin Plantinga's reductive analysis of (non-Lewisian) possible worlds in terms of maximally consistent sets of true propositions.[56] The notion of a possible world is part of the philosopher's lexicon and is obviously beyond the scope of Moorean support. So, the modified approach fails to justify positing facts about (non-Lewisian) possible worlds as distinct existents, or any other theoretical notions that lack a Moorean basis. In this light, the appeal to Mooreanism fails to guarantee the general dialectical cogency of the reductive procedure and is unable to provide sufficient justifications for the many of existence claims the grounding theorist is required to make.[57]

As a final rejoinder, the ontological permissivist may pursue an alternative source of justification. For instance, she may argue that we are justified to posit non-fundamental entities to serve as these more esoteric facts by virtue of the theoretical work that they perform.[58] However, the appeal to theoretical work also faces problems. First, we can easily make sense of cases of reduction that involve things that do not have any meaningful theoretical applications. For instance, it is perfectly sensible to take the notion of *being grue* to be reducible to the notions of *being green* and *being blue*. On the grounding view, this requires us to posit non-fundamental entities to serve as the grue-facts. But because there is no meaningful theoretical purpose that the grue-facts serve, they fall outside of the scope of the suggested approach. Second, the appeal to theoretical work

---

[56] Plantinga (1974)

[57] In the next two chapters, I discuss how dialectical cogency relates to the truth of theories.

[58] This line of response was suggested to me by Phil Bricker. It is also noteworthy that there are different ways to think about the appeal to theoretical work relates to Schaffer's Mooreanism. On the one hand, because Moorean support is only suggested as a sufficient condition for justification, his account has room for there being an alternative source of justification for more esoteric non-fundamental entities. In this light, one available option is to take the thesis below to be supplementary to his account. On the other hand, if it can be argued that accounting for Moorean truths is itself a significant theoretical achievement, then it may not be too implausible to subsume Mooreanism under the approach under consideration. This issue is irrelevant to my discussion below, so I shall leave it unresolved.

itself does not provide strong enough reasons to posit non-fundamental entities as *distinct* existents. At best, the relevant considerations provide reasons for us to accept the general claim that there is some entity or other than plays a certain theoretical role. In order to draw the stronger conclusion, we need reasons to believe that the occupant of a given role is not identical to anything else in our ontology. This means that we are justified to posit non-fundamental entities (as distinct existents) only if the identity view fails: i.e., the appeal to theoretical work fails to justify the postulation of non-fundamental entities in case the work in question can be accounted for without doing so.

2.4 Reduction and Explanatory Roles

2.4.1 Revisiting the Puzzle: Reductive Explanation

For the previous reasons, I think we should not give up on reductive identity so quickly. A solution that presents itself is to invoke *modes of presentation* in some form or other.[59] The exact details may vary, but in any case, the key move is to separate the relata of the identity relation from the relata of the explanatory relation. [60] Consider again the general form of Frege's puzzle. What we need is an account of why claims like (8) below are informative whereas claims like (9) are not:

(8) Mark Twain is Samuel Clemens.

---

[59] This is also the general strategy Jenkins (2011) takes to reconcile grounding with identity claims.

[60] A major juncture lies in whether modes of presentation are posited at the semantic or extra-semantic level. For instance, on Frege's own solution, modes of presentation are abstract objects (i.e., Fregean senses) that account for the semantic link between terms and their referents. But alternatively, those who reject Fregean senses may take modes of presentation to provide extra-semantic or meta-linguistic information, e.g., the fact that 'Samuel Clemens' and 'Mark Twain' are co-referring names. See Bealer (2004) for a critical assessment of this general strategy.

(9) Mark Twain is Mark Twain.

On any account that appeals to modes of presentation, the solution involves identifying the *referents* of the terms 'Mark Twain' and 'Samuel Clemens' while distinguishing the modes in which the same individual is presented by each name. And by taking the informativeness of (8) to consist in the non-identity between these modes of presentation, the puzzle is resolved.

Similarly, reductive identity can be preserved by separating that which is claimed to be identical and that which is claimed to stand in an explanatory relation. Consider how Sider treats the issue. Soon after the passage quoted in Section 2.2, he writes (Sider 2003, p.7, his emphasis):

> [T]o be non-circular, or genuinely reductive, the terms in which the reducing propositions are expressed must be "non-modal". But there is a bit of awkwardness here: in a sense the reducing terms must indeed be modal if the reduction is successful, since if the reduction is successful then the reducing propositions *are* modal propositions, given that analysis is identity. The awkwardness should be resolved as follows. Any reductionist program takes certain notions as being "acceptable". What acceptability amounts to depends on what is driving the reduction—it may be epistemic acceptability, or categoricity, or extensionality, or something else. "Non-modal", then, means "acceptable"—a reduction is non-circular or genuinely reductive if the notions it employs are acceptable according to its standards, whatever those may be.

At face value, these remarks should strike us as odd. The "awkwardness" Sider recognizes is a slight variant of our initial puzzle in that it concerns a more general violation of Leibniz's Law. Suppose that p = <It is possible that X> and that q = <There is a Lewisian world w such that X is

true>. In this case, the supposition that $p = q$, together with Leibniz's Law, licenses the following inference:

(10) p is a modal proposition.

(11) Therefore, q is a modal proposition.

Thus, by generalization, any reduction would seem to be doomed to be circular. But how does moving to acceptability solve this problem? On the face of it, the same reasoning seems to apply to the following as well:

(12) p is an unacceptable proposition.

(13) Therefore, q is an unacceptable proposition.

In this light, his remarks about acceptability are a red herring: there is nothing about "being modal" or "being unacceptable" that allows us to apply these predicates to p but not to q.[61] For us to do this, we need to reject that Leibniz's Law applies in any of these cases. And the only way to do this, it seems, is to treat the above designations of propositions (i.e., 'p', 'q', and our notations in pointed brackets) as opaque; while maintaining that $p = q$, we should also maintain that we cannot substitute 'p' and 'q' *salve veritate*. There are various ways in which the details of this general strategy can be worked out. Here, I do not provide an extensive survey of the matter and instead only propose a view that I take to be promising.

---

[61] There are ways to read Sider that is compatible with my remarks here. For starters, we may take him as suggesting that the sense in which the reducing propositions are modal propositions require a special context in which our designations of propositions are transparent. Furthermore, I conjecture that Sider is implicitly taking the "notions" employed in a reduction are more fine-grained than propositions, and so that the same proposition can be expressed by different notions. And by taking these "notions" to yield different modes of presentation, he may also block the application of Leibniz's Law.

2.4.2 Reductive Identity and Explanatory Roles

The view I propose draws inspiration from David Lewis's recipe for causal-role analyses from his "Reduction of Mind" (Lewis 1994). Here's how his recipe goes for the reduction of mental states. The recipe roughly has three steps. First, by consulting "folk-psychology," we determine the causal role that is associated with the target mental state, call it M. For instance, we look for certain patterns of behavior that are caused by M and the perceptual stimuli that cause M or affect it in various ways. Second, we define M as the occupant of the causal role so defined. Third, through scientific investigation, we look for the physical state P that occupies the causal role in question. From these three steps, we arrive at the conclusion that $M = P$.

For our purposes, we can expand Lewis's talk of causal roles in terms of metaphysical explanations more generally. For the sake of the illustration, I will continue to talk in terms of propositions (whatever they may be) and later discuss how the method generalizes. Let us say 'the A-role' denotes the explanatory role that is associated with concept A. For instance, there are distinctive explanatory contributions that are made by modal, causal, and tensed propositions: propositions of each kind play distinctive but interconnected roles in inquiries about morality, history, in our scientific reasoning, etc.[62] (These facts also bear explanatory connections to one another: e.g., modal facts and tensed facts are crucial for causal reasoning, etc.) In the first step, we follow Lewis and define propositions in terms of the explanatory roles they occupy: e.g., the modal propositions are those that occupy the modality-role, the causal propositions are those that occupy the causality-role, etc. The second step is to find propositions of another kind that can also

---

[62] E.g., for the case of modality, we may roughly say that propositions of the form <It is impossible that X> may explain the truth of moral propositions of the form <It is not obligatory that one does X> (granted that ought implies can), propositions of the form <It is nomologically necessary that, if X occurs, then Y occurs> may explain the truth of causal propositions of the form <X causes Y>, etc.

occupy these roles. Crudely speaking, we may take Lewisian modal realism as claiming that propositions that quantify over Lewisian possible worlds can occupy the modality-role; counterfactual analyses of causation as claiming that counterfactual propositions of varying sorts can occupy the causality-role; the Block Universe Theory as claiming that propositions that involve time-like distance relations (supplemented by a semantic theory of indexicals) can occupy the temporality-role, etc. As the third and final step, we identify propositions that occupy the same explanatory role. Let us call this *the explanatory-role account* or *ERA* for short. This gives us the following formulation of reductive identity:

REDUCTIVE IDENTITY (ERA):   Concept *A* reduces to concept *B* if and only

if every occupant of the *A*-role is identical to the occupant of some *B*-role, but not

vice versa. [63,64]

---

[63] By 'concept', what I really mean is 'bit of conceptual ideology'. Using the case of modality as an example, we may roughly say that the concept of something being *possible* reduces to the concept of there being a certain kind of concrete possible world iff any proposition that expresses something's being possible is identical to some proposition that describes a Lewisian world but not *vice versa*.

[64] The exclusion of the *vice versa* clause is what secures the asymmetry of reduction. For instance, the identification of modal facts with quantificational facts gives us a reduction of modality because, while there are quantificational facts that (supposedly) occupy the same explanatory role as any modal fact, the modal facts do not occupy the same explanatory role of any quantificational fact—they cannot occupy the explanatory role of propositions that do not contain any modal notion. Previously, I had thought that a worthy alternative is to secure the asymmetry by appealing to fundamentality. E.g., one may think that we have a reduction of modality because quantificational facts are (supposedly) more fundamental than modal facts. However, this approach requires us to make judgments about fundamentality in a way that is independent of the reductive procedure. And this seems wrong. Instead, it is more plausible that matters of reducibility inform us about matters of fundamentality. E.g., that modality is irreducible seems to be a powerful reason to think that modality is non-fundamental. But the fundamentality-approach has things backwards: it requires us to think modality is reducible *because* it is non-fundamental. I thank Justis Koon for discussing this matter with me on several occasions and Phil Bricker for pushing me on this issue.

The key to the puzzle of reductive identity lies in taking explanatory roles to serve as modes of presentation. This is partially anticipated by Lewis's causal role analysis. A central aspect of his method is to use a given causal role as a means to pick out a given state as *the mental state M* which is later identified with a physical state *P*. Assuming that the reduction is successful, it turns out that our reference to *M* was already a reference to *P*; it's just that our reference was made in association with the *M*-role. A similar story applies to our case. On the one hand, we are using the modality-role as a means to pick out a collection of propositions as *the modal propositions* which are later identified with a collection of propositions that are picked out by some other explanatory role, namely the *Lewisian-possible-world-role*. Granted that the reduction is successful, it is again revealed that our reference to the modal propositions was already a reference to the Lewisian-possible-world propositions; it's just that it was made in association with the modality-role.

It will be useful to have a shorthand for propositions under different modes of presentation. Let us call a proposition under a given mode of presentation a *conception* of that proposition. For instance, suppose again that p = <It is possible that X> and that q = <There is a Lewisian world w such that X is true>. Furthermore, let us suppose:

(14) r = the proposition designated by 'p' and 'q'

In this case, we can say 'p' involves the *modal conception* of r whereas 'q' is the *Lewisian-possible-world* conception of *r*. Given this terminology, we can reformulate REDUCTIVE EXPLANATION as:

REDUCTIVE EXPLANATION (ERA): If concept A is reduced to concept B, then, for any X that is both the occupant of the A-role and the occupant of the B-role, the A-conception of X is explained by the B-conception of X.

As none of these reformulations contain reference to propositions, the general approach here is compatible with a myriad of ontological frameworks. For instance, if one prefers to work with Armstrongian states of affairs rather than propositions, one can take the occupants of explanatory roles to be states of affairs. In this case, reductive explanation holds between different conceptions of states of affairs rather than propositions. Alternatively, one may wish to combine the two views. That is, one may take propositions to be much more fine-grained than states of affairs and treat the propositions as providing different concepts of states of affairs. In this case, reductive identity is cashed out in terms of states of affairs whereas reductive explanation is cashed out in terms of propositions. While there may be advantages and disadvantages for choosing between these variations, these concern the merits and shortcomings of the relevant ontological frameworks, not the strategy I propose.

2.4.3 Explanatory Roles and Structure

A point of interest concerns how ERA relates to the worldly view of ideology.[65] And this relationship is trickier to spell out than it may initially seem. A natural approach is to take non-

---

[65] It is worth noting that the worldly view of ideology is also compatible with the grounding view. On the identity view, there is only a single layer of structure that underwrites both fundamental and non-fundamental ideology. But the grounding view may posit multiple layers of structure that correspond to the multiple layers of ideology. On this

ontic structure to serve as the occupants of explanatory roles—that ideological reduction bridges two distinct conceptions of the same non-ontic structure. The explanatory value of this bridging consists in the fact that the reducing concept carves nature better at its joints and provides a more conspicuous representation of the world's structure. In the ideal case, the reducing concepts carve nature at its *exact* joints. But even in non-ideal cases, the result of a reduction is nonetheless a redescription of a given phenomenon that sheds better light on the worldly structure underneath it. The tricky part, however, lies in cashing out the idea that the reduced and reducing concepts provide distinct conceptions of the *same* non-ontic structure without reifying structure-talk. If we are to maintain the Quinean connection between quantificational idioms and ontology, we cannot straightforwardly say "*There is* a structure *S* such that concepts *A* and *B* provide distinct conceptions of *S*." without *ontologically* committing to reified structures and thus treating '*S*' as referring to an entity within our ontology. [66]

After close inspection, there seems to be no easy way around this. As far as I can tell, the closest we can get is saying that (i) there are *B*-facts that are structural, and that (ii) the *B*-facts are conceptualized as *A*-facts under certain conditions. And strictly speaking, this has the occupants of explanatory roles being *facts* (again, either understood as propositions or states of affairs) rather

---

picture, we may take non-fundamental structure to be grounded in fundamental structure. (See Sider 2017). Side note: this leads to some awkward terminology. While it is natural to say that the identity view involves a "reductive" view of non-fundamental structure, by the lights of the grounding view, the fact that higher-level structure is grounded by lower-level structure also warrants the label "reductive."

[66] It is easy to see that his structural operator '$\mathscr{S}$' is unhelpful for this task. First, the operator is meant to express absolute structure—the exact joints in nature. So, it cannot be used to express concept A's imperfect relationship to structure, for the claim '$\mathscr{S}(A)$' (i.e., 'A is (absolutely) structural.') is precisely what is meant to be rejected. Second, more importantly, the operation it is meant for yields an expression of the wrong grammatical form. By attaching '$\mathscr{S}$' to an expression, we obtain a sentence of the form 'so-and-so is structural'. This result suits Sider's purpose of making claims about what kinds of structure the world contains, as it allows him to say things like "Quantification is structural." ("$\mathscr{S}(\exists)$"), etc. However, it is unsuitable for making claims about structure identity, e.g., that quantification and logical connectives are distinct structural aspects of the world, or in our case, that the reduced concept is underwritten by the same structure as its reducing concept. And as mentioned the previous chapter (Section 3.3), it is not obvious how these sorts of claims can be expressed with '$\mathscr{S}$'.

than worldly structure. To see this, consider how Sider could deal with the situation. For Sider, any truth that involves a non-fundamental term or concept is a non-fundamental truth. And since concept *A* is stipulated to be non-fundamental (by virtue of the reduction), it follows that the claim we are dealing with is non-fundamental as well. And for him, anything non-fundamental is dealt with by semantic ascent: i.e., by providing a metaphysical semantics of the natural language sentences they are expressed in. Sider makes several suggestions of what such an account may look like, but the one that seems most relevant to us involves producing a biconditional of the form 'Sentence *S* of some natural language *L* is true in *L* iff ϕ' where 'ϕ' is phrased in purely joint-carving, absolutely fundamental terms.[67] Applied to our case, the task boils down to providing a metaphysical semantic account of the previous English sentence "There is a structure *S* such that that concepts *A* and *B* provide distinct conceptions of *S*." that does not quantify over structure.

There are several different ways to pursue such an account, depending on one's view about what is and what is not fundamental. But here is one, very crude way in which Sider may proceed. Let us suppose that *B* is a rock-bottom, fundamental concept. In this case, we mainly need to do two things. First, we need to find a proxy for our talk about structure. For this, we may instead refer to facts (understood either as propositions, states of affairs, or as some sort of construction that is acceptable within his framework) that are relevant to the kind of structure involved.[68] Second, we need a way to express the reduced concept *A*'s imperfect relationship to structure. There are again many different ways the story may go. For the sake of illustration, let us consider

---

[67] See Sider (2011, ch.7) for other suggested forms of such an account.
[68] Consider how he suggests the mereological nihilist should account for truths about hydrogen atoms (Sider 2011, pp.141-144). Since the nihilist denies the existence of composite objects like hydrogen atoms, along with any claim of the form '*x* is a hydrogen atom.', she cannot simply give a translation of 'is a hydrogen atom'. Instead, she needs to give a semantic account of entire sentences by only referring to mereological simples. E.g., he suggests the nihilist to take the metaphysical truth-conditions for hydrogen-sentences in English to be provided by statements about electrons orbiting protons. In this case, we may say electrons, protons, and the property of *orbiting* serve as a proxy for talk of hydrogen atoms.

a sort of psychologistic approach. (The details of what we say here are not central to my point below.) To this end, suppose we take the *B*-facts to *cause* English speakers to be in a certain mental state such that the *B*-facts are conceptualized as *A*-facts. (E.g., we may say that the fact that *p* is true in a Lewisian world *w* to cause English speakers to conceptualize this fact as the fact that *possibly-p*.) From here, swap out anything you think is non-fundamental in this story with something you think is fundamental. For instance, given Sider's affinity for physicalism and mereological nihilism,[69] he may opt to replace reference to English speakers with reference to a collection of particles in a highly complex arrangement, e.g., arranged *the-totality-of-English-speakers-wise*, the reference to conceptualization by whatever brain activity (which can also be replaced by talk of the interaction of particles) that this psychological process may reduce to, etc. (While the same goes for causation, let this one slide for the sake of brevity.) Following this line of thought, we get something like:

(15) 'There is a structure *S* such that concepts *A* and *B* provide distinct conceptions of *S*.' is true in English iff there are particles arranged *the-totality-of-English-speakers-wise* such that the *B*-facts cause them to interact *A-conception-ly* and to interact *B-conception-ly*.

I do not suggest that this even remotely captures what Sider would actually say. Nor do I suggest that (15) is the best version of the account under consideration.[70] The purpose of (15) is simply to

---

[69] Sympathy for physicalism is expressed in various places throughout Sider (2011). But see Sider (2017) for more explicit remarks. For mereological nihilism, see Sider (2013).

[70] For instance, expressions in which the terms '*A*' and '*B*', 'English speakers' are embedded should be substituted for more "exact" descriptions of what is involved. For instance, the arrangement and interaction of particles can perhaps be given by describing their positions in spacetime and their physical interactions therein. The more substantive issue concerns the reference to the *B*-facts. Part of the issue is that the term involves the concept name '*B*'. Even if we grant that the content of *B* is suitably fundamental, insofar as one doesn't think concepts themselves are fundamental entities (or for Sider, that 'Concepts exist.' is a fundamental truth), one may wish to describe the relevant facts without referring to the concepts through which they are picked out. But there is also room for more substantial revisions. I

have a more concrete example to think about the issue. And the main take-away is this. *Ex hypothesi*, the terms involved in the right-hand side of the biconditional are absolutely fundamental; they are already fully cashed out in the most fundamental term there are. Thus, whatever it is that we choose to use as the proxy for structure-talk, it is those *proxies* that end up being involved in explanatory roles and thus being subject to different conceptions. More generally, if we maintain that worldly structure is non-ontic, and that the resources we use for ontology are not applicable to worldly structure, then very few claims about structure become expressible in purely fundamental terms. As a result, the metaphysical semantic account reveals that much of what we say about structure in English is not "really" about structure at all; it is instead about things that we can appropriately quantify over.

This seems to be the situation that Sider's overall framework leaves us with. But perhaps there are ways to avoid this implication by expanding the framework in certain ways. For reasons of space, I will not consider the relevant options in much detail.[71] Instead, I rest with the conditional conclusion that, on roughly *Sider's* framework for the worldly view of ideology, there

---

have assumed that Sider's fundamental language has the resources to refer to facts in general, but this may turn out not to be the case. Also, it may turn out that causation (nor whatever one takes it to reduce to) is unsuitable for capturing concept *A*'s imperfect relationship to structure. In this case, the account would need to be revised much more extensively.

[71] One possible strategy is to make use of predicate functors in the place of quantifiers. The rough idea is that, when we would say "There is a structure *S* such that it is *F*." ("∃*x* (*Sx* & *Fx*)"), we instead say "It is *F-structuring*." ("σ(ρ(κ*SF*)).", which Burgess suggests we read in English as "Doth just self-respectively structures and *F*'s."). (Predicate functors are often discussed in the context of ontological nihilism—the claim that, fundamentally speaking, there are no individuals. See Quine (1960) and Burgess (2005) for an introduction of predicate functors. For discussion in relation to ontological nihilism, see Dasgupta (2009) and Diehl (2018); see Turner (2011) for criticism of its use in defense of nihilism. Also see Donaldson (2015) for its implications for (Sider's version of) the worldly view of ideology.) Since a language of predicate functors (or 'functorese') has been shown to be just as expressive as standard predicate logic, it is likely to allow us to make non-ontologically loaded claims about structure to our heart's content. But a full assessment of the strategy requires much further investigation. A pressing question is whether one can make use of predicate functors exclusively for matters of structure: i.e., whether adopting them in this way pushes one toward ontological nihilism. An obvious source of pressure is ideological simplicity, as a hybrid view that accepts both quantifiers and predicate functors as primitive is ideologically more complex than a view that accepts only one set of them. But also, for what it's worth, it renders Turner's objection to nihilism unavailable, as it relies on the claim that statements that involve predicate functors are no less ontologically committing than their counterparts in standard predicate logic.

is no serious sense in which we can say that there can be two distinct conceptions of the *same* structure. We may of course make loose and metaphorical claims of this sort, as the mereological nihilist can for tables and the nominalist can for universals. But we should not take these claims to provide any serious insight into the matter. This further suggests that explanatory roles are best understood as taking in facts or propositions as their occupants, not worldly structure. Thus, structure does not play a significant role in our understanding of the reductive procedure. The notion of structure helps us determine which of our conceptions of facts are the privileged, elite ones that carve at the joints. While this helps us determine whether a reduction has taken us closer to the bedrock of reality, it tells us little about the route that brought us there.

2.5 Summary

In this chapter, I argued against a solution to the puzzle of reductive identity that characterizes (ideological) reduction in terms of metaphysical grounding. The problems I presented were twofold. First, I argued that an account in terms of grounding fails to capture the fact that reductive relations involve explanatory deference. Second, I argued that, for those who accept a worldly view of grounding, the grounding view of reduction requires us to expand our ontology in a way that is not dialectically cogent. Based on these considerations, I conclude that the puzzle of reductive identity does not motivate the grounding view. Next, I explored a way to solve the puzzle that does not reject reductive identity. To this end, I proposed that we expand Lewis's recipe for a causal analysis and take ideological reduction to consist in identity relations that hold between occupants (i.e., propositions or facts) of explanatory roles.

CHAPTER 3: DIALECTICAL VICES

3.1 Introduction

Dialectical virtues concern features of a theory (either intrinsic or extrinsic) that place its advocates at an advantage in debates. Similarly, dialectical vices concern features that place a theory's advocates at a disadvantage. I suspect that most philosophers are well acquainted with the notion of dialectical virtues and vices, though not under these names. For a quick example, consider arguments that appeal to the burden of proof. Suppose it is shown that, while the burden of proof for a given issue falls on those who advocate theory *T*, it is very unlikely that the relevant proof exists. In this case, many would agree that this provides powerful considerations against *T*. I argue in this chapter that the force of these considerations comes from the dialectic context of the debate. That is, I argue that arguments like the above point out that the target theory has the dialectical vice of bestowing the obligation upon its advocates to present a proof that might not exist.

My main goal is to show that there are at least some dialectical virtues and vices that are epistemically relevant in the sense that we are epistemically better off by accepting theories that are dialectically virtuous and by rejecting those that are not. Let us call this the *Epistemic View of Dialectical Virtues and Vices*, or 'EDV' for short. While EDV seems *prima facie* plausible, there are some serious obstacles that need to be addressed. The main obstacle is the idea that there is a significant gap between our dialectical efforts and the truth of theories. In various contexts, several authors have maintained that matters of the dialectic (e.g., the defensibility of theories) constitute a separate matter from truth. If so, it would seem that dialectical virtues and vices are not

epistemically relevant after all and that they have limited implications for the truth of our theories. Below, I call this gap between our dialectical interactions and truth *the pragmatic-epistemic gap*. In what follows, I make efforts to partially close the pragmatic-epistemic gap and argue that matters of the dialectic are more closely related to truth than it may otherwise seem.

I approach this issue by way of a case study. The case study centers around a theory that is often associated with what I take to be a dialectical vice. In Section 3.2, I lay down the groundwork for the case study by examining the nature of this dialectical vice and how it relates to the theory in question. In Section 3.3, I explain how this dialectical vice relates to the evaluation of this theory, and then expand this into an account of the epistemic relevance of dialectical virtues and vices in general.

3.2 Case study: The Error Theory of Normativity

3.2.1 Overview

The case study involves the global error theory of normativity that takes absolutely *all* of our normative judgments are false, or "The Error Theory" hereafter. More precisely, the case study concerns a particular line of criticism that can be raised against The Error Theory. As a *global* error theory of normativity, the theory also rejects epistemic normativity and thus entails a strong form of epistemic nihilism that says, roughly speaking, nothing is a *reason for belief*. According to various critics, epistemic nihilism is self-defeating: if nothing is a reason for belief, then *a fortiori*,

nothing is a reason to believe that epistemic nihilism is true. And since The Error Theory entails epistemic nihilism, we may also say that The Error Theory is self-defeating in the same way.[72]

What is interesting about this line of criticism concerns the *kind* of self-defeat that is relevant. In various philosophical contexts, we are interested in criticisms that allow us to conclude that a given theory is false. But it has been noted by many that the kind of self-defeat that we are currently concerned with fails to warrant such a conclusion. Accordingly, apart from the question of whether epistemic nihilism (and thus, The Error Theory) is indeed self-defeating, there is a further question of what the criticism properly implies.

In this section, I first clarify what the relevant kind of self-defeat is, explain how it applies to The Error Theory, and discuss some shortcomings of how the criticism from self-defeat is addressed in the literature. In the following section, I discuss how the notion of dialectical virtues and vices help overcome these shortcomings and what this implies for the general relationship between _EDV_ and the pragmatic-epistemic gap.

3.2.2 The Error Theory and Self-Defeat

Let us start by laying out the theory more clearly. The tagline for The Error Theory is that *all* of our normative judgments are false. For ease of exposition, I follow how this idea is cashed out by Bart Streumer (2013), who is one of the very few who explicitly defends The Error Theory and embraces its consequence for epistemic nihilism. Streumer characterizes the theory in terms

---

[72] The reason the case study involves The Error Theory rather than epistemic nihilism is because the way Bart Streumer (2013) treats the issue of self-defeat on behalf of The Error Theory is deeply relatedly to our discussion later.

of property-talk and takes it to consist of the claims that (i) normative judgments are beliefs that ascribe normative properties, that (ii) normative properties do not exist, and therefore, that (iii) all of our normative judgments are false.[73] There are two features of The Error Theory that are most relevant to our purposes. The first is that The Error Theory involves a *global* rejection of normative properties. That is, as opposed to a *local* error theory of normativity that only rejects normativity of a given domain, say morality, The Error Theory rejects normativity across *all* domains.[74] Thus, not only does the theory reject moral properties, or instrumental / practical normative properties (i.e., normative properties that are dependent on our goals and desires), but it also rejects epistemic normative properties that concern what we ought to believe. The second is that The Error Theory does not merely reject some thick conception of normativity, nor does it come with a reductive account of what our normative discourse is "really" about (cf. expressivism of normativity). Rather, it is the claim that normative judgments are simply *false* because they are beliefs that presuppose the existence of properties that (allegedly) do not exist.[75]

The two features just mentioned provide the basis for the criticism from self-defeat. The reasoning can be summarized as follows:[76]

---

[73] Those who typically understand negative existential claims in terms of uninstantiated properties should take claim (ii) with a grain of salt. E.g., if one typically analyzes negative existential claims of the form "There are no *K*'s." in terms of the property *K*'s being uninstantiated, one may be tempted to take (ii) as claiming that the property of *being a normative property* is an uninstantiated second-order property. But this does not seem to be what Streumer means by (ii). Ultimately, Streumer concedes that The Error Theory entails that nothing is a *reason for belief*. But on the reading under consideration, The Error Theory would entail that the first-order property of *being a reason for belief* is not a normative property. To circumvent this issue, those who wish to understand (ii) in terms of uninstantiated properties should take it to mean that, roughly speaking, the properties that are *commonly associated with normativity* are uninstantiated. This way, we get the intended entailment that there are no instances of the property of *being a reason for belief*. To make this reading more salient, one may take (ii) as claiming that *the normative properties* are uninstantiated. I thank Ned Markosian and Seungsoo Park for discussing this matter with me.

[74] The terms *local* vs. *global* error theories are taken from Lutz (2014).

[75] Some context: very few (unsurprisingly) defend the Error Theory explicitly and the view is mostly raised as an undesirable consequence of moral error theory. See, for instance, Cuneo (2007), Stratton-Lake (2002), Gibbard (2003), Parfit (2011). These references are taken from Streumer (2013).

[76] The argument more or less follows the reasoning expressed in Streumer (2013, p.197).

1. The property of *being a reason for belief* is a normative property, if such a property exists.[77] (*Premise*)

2. The Error Theory entails that there are no normative properties. (*Premise*)

3. The Error Theory entails that nothing is a reason for anything. (*By 1, 2*)

4. The Error Theory entails that there is no reason to believe The Error Theory. (*By 3*)

5. The Error Theory is self-defeating. (*By 4*)

Understood as a line of *criticism*, it is tempting to draw the further conclusion that:

6. The Error Theory is false. (*By 5*)

Intuitively, the reasoning outlined above provides a powerful reason to reject The Error Theory. However, several authors note a difficulty for moving from (5) to (6). In order to make the relevant inference, The Error Theory needs to be self-defeating in a way that indicates that the theory is false. But it is not entirely clear that this is the case. Let us take a step back and inspect the inference from (4) to (5). To make this inference, the fact that there is no reason to believe The Error Theory must be something that conflicts or "defeats" the theory. But notice that there is no such conflict with respect to the tenets of the theory. If anything, it conflicts with one's *claim* or *argument* that the Error Theory is true. This should become apparent the more we think about it. After all, it is thinkers and not theories that need *reasons for beliefs*. In this light, it is not entirely obvious that (6) follows from (5).

---

[77] Alternatively, one may run the argument in terms of the property of *having a reason for belief*, which can be ascribed to epistemic subjects, instead of the property of *being a reason for belief*, which can be ascribed to propositions. On this alternative approach, the key step of the argument is that The Error Theory entails that no epistemic agent ever *has a reason* to believe The Error Theory. The discussion below is compatible with either way of putting the point.

The point here can be clarified by sharply distinguishing between two kinds of self-defeat. The general issue is well explored by Jeremy Morris (2008). According to Morris, there are two different ways in which an expression can be self-defeating.[78] For illustration, compare two paradigmatic cases of self-defeat, namely the Liar Paradox and Moore's Paradox. Consider these two sentences:

(L) This sentence is false.

(M) *p*, but I don't believe *p*.

Intuitively, both (L) and (M) are self-defeating sentences. But there is a key difference in the way in which they are self-defeating. For (L), we may say the sentence is *logically self-defeating* in the sense that the truth-conditions of (L) are inconsistent with what the sentence says: the truths-conditions of (L) are such that it is true only if it is false. But the same cannot be said about (M). Unlike (L), there seems to be nothing *logically* peculiar about (M). It is certainly possible for *p* to be true while I do not believe this to be the case; thus, unlike (L), it is not the case that (M) is true only if it is false. Thus, (M) is not logically self-defeating. To account for the self-defeating nature of (M), Morris argues that we need to consider features of the act of *asserting* (M). His diagnosis goes like this. Plausibly, it is constitutive of a successful assertion that it conveys what the speaker

---

[78] Terminological point: instead of talking in terms of there being two *kinds* of self-defeat like I do, Morris takes there to be a single kind of self-defeat that arises for different reasons. That is, instead of distinguishing between *logical* and *pragmatic* self-defeat as I do below, he distinguishes between cases of self-defeat that arise due to *logical reflexive inconsistencies* and those that arise from *pragmatic reflexive inconsistencies*. His general account of self-defeat, which he calls 'P4' states that: if the relevant use of [an expression] 's' is successful, then that use of 's' is a failure on account of the inconsistency of that use of 's' with the conditions of success for that same use of 's'. On this account, a given expression comes out as logically self-defeating in case the relevant sort of inconsistency arises due to the fact that 's' refers back to its own logical features; when the inconsistency arises due to the fact that 's' refers to the conditions for a successful use of 's', it comes out as pragmatically self-defeating. I take the differences between his and my ways of speaking to be inconsequential.

believes.[79] So, for a speaker to successfully assert (M), the assertion must convey that the speaker believes *p*. But since (M) also says that the speaker does not believe *p*, a successful assertion of it entails that the aforementioned success conditions are not met. In this light, we may say (M) is *pragmatically self-defeating* in the sense that the act of asserting (M) is successful only if it fails.[80]

Applying Morris's distinction to our case, it is evident that The Error Theory is *not* logically self-defeating. There is nothing about The Error Theory that is logically inconsistent with the consequence that there is no reason to believe The Error Theory. Thus, if The Error Theory is self-defeating at all, it must be pragmatically self-defeating. Below, I say more about how to understand the self-defeating nature of The Error Theory. But at a general level, we can already appreciate the difficulty in moving from (5) to (6). As evidenced by (M), pragmatic self-defeat is not generally indicative of falsity. Thus, even if it is granted that The Error Theory is pragmatically self-defeating, this does not allow us to directly infer that the theory is false.

The main goal of the case study is to explore ways to overcome the above difficulty. In the remainder of this section, I give a more detailed discussion of how the notion of pragmatic self-defeat applies to The Error Theory. Next, I make preliminary remarks about how the charge of pragmatic self-defeat can be developed into a proper criticism of The Error Theory. In the following section, I conclude the case study by providing my proposed account of the matter and discuss how the lessons from the case study can be generalized.

---

[79] This claim about the constitutive nature of assertion is not entirely uncontroversial. For example, Rescorla (2009) argues for an unrestricted model of assertion that does not generally require speakers to assert what they believe. The issue Rescorla pursues bears a complicated relationship to our topic the disentangling of which goes beyond the purview of this dissertation. To sidestep this issue, we may at least say that it is constitutive of assertion that it *presents* the speaker as believing what is being asserted.

[80] This diagnosis is bolstered by the fact that several authors aim to solve Moore's Paradox by appealing to the relation between the sentence and the speaker. E.g., see Lee (2001).

(Clarificatory note: While Morris's discussion involves a notion of pragmatic self-defeat that applies to linguistic expressions like sentences, it is also sensible to talk about the pragmatically self-defeating nature of the propositions that a given use of them express. By further extension, it is also sensible to talk about the pragmatically self-defeating nature of expressive acts. In what follows, I leave open how these three notions of pragmatic self-defeat relate to each other.[81] Furthermore, while I primarily talk in terms of pragmatically self-defeating expressions, I will at times talk in terms of self-defeating propositions and expressive acts depending on the context.)

3.2.3 Is The Error Theory Pragmatically Self-Defeating?

Here, I do not give a general survey of the different ways in which the charge of pragmatic self-defeat can be developed. Instead, I only discuss two noteworthy approaches. Before diving in, it will be helpful to consider how much leeway we have. Let us start by refining the notion of pragmatic self-defeat. As hinted earlier, the general idea is that an expression e is pragmatically self-defeating in case the use of that e is successful only if it fails. More precisely, it concerns cases where the relevant failure obtains by virtue of an inconsistency between the success conditions of the use of e and what e expresses.[82] When an expression is self-defeating in this fashion, it appears to be so with respect to at least two parameters. First, an expression is pragmatically self-defeating

---

[81] That being said, it strikes me as plausible that the pragmatic self-defeating nature of sentences consists in that of propositions, and that the self-defeating nature of propositions consists in that of expressive acts. For instance, it seems plausible to say the *sentence* (M) is pragmatically self-defeating by virtue of the fact that each speaker's use of (M) expresses a *proposition* whose pragmatically self-defeating nature holds by virtue of the fact that a speaker's *act* of asserting the proposition expressed by her use of (M) is pragmatically self-defeating—i.e., the fact that the speaker's act of asserting the relevant proposition is successful only if it fails.

[82] The by-virtue-of claim is added in order to exclude cases where the use of e is a failure due to irrelevant factors, e.g., being grammatically nonsensical. See Morris (2008, p.209) for relevant discussion.

with respect to speakers. In the case of (M), the sentence is self-defeating with respect to *all* speakers. But this need not always be the case. For example, the sentence '*p* but Victor does not believe *p*.' is pragmatically self-defeating when asserted by Victor but not when it is asserted by Sam. Second, an expression is pragmatically self-defeating with respect to kinds of expressive acts. Returning again to (M), the self-defeating nature of (M) was explained in terms of the success conditions for the act of *assertion*. But there are plenty of cases of pragmatic self-defeat that involve other kinds of expressive acts: e.g., one may *boast* that one is always humble, *yell* that one never yells, etc.[83] Furthermore, the same expression can be self-defeating with respect to one kind of expressive act but fail to be so with respect to another: e.g., the sentence 'I am always humble' is pragmatically self-defeating with respect to the act of *boasting* but not with respect to the act of *yelling*. In light of these considerations, I define 'pragmatic self-defeat' as follows:

> PRAGMATIC SELF-DEFEAT: An expression *e* is *pragmatically self-defeating* with respect to speaker *S* and expressive act $A =_{def} S$'s use of *e* is a successful instance of *A* only if it fails to be so by virtue of the fact that *S*'s use of *e* expresses a proposition that is inconsistent with the success conditions for *A* regarding that same use of *e*.[84]

The self-defeating nature of The Error Theory supposedly holds irrespective to any particular speaker. So speaker-relativity will not play an important role in my discussion below. Instead, the

---

[83] See Morris (2008, pp.210-211, footnote 9).

[84] This definition is meant to capture the general spirit of Morris's account mentioned in footnote 7. While Morris puts forth his account (P4) only as a sufficient condition, for the sake of simplicity, I treat this as a stipulative definition of pragmatic self-defeat. Also, the keen reader may notice that the definition involves an ambiguity regarding the terms 'successful' and 'fails'. I address this issue below, especially in footnote 15.

main question I am asking is: with respect to what expressive act is The Error Theory pragmatically self-defeating? Alternatively: what expressive act (if any) has success conditions that are inconsistent with The Error Theory? Below, I consider two possible answers.

The first approach is to argue that The Error Theory is pragmatically self-defeating with respect to the act of assertion, or more simply, that it is *assertorically self-defeating*. To this end, one may hope to develop the previous diagnosis for (M) into a more general account that is applicable to The Error Theory. In light of the above definition, the self-defeating nature of (M) can be said to consist in the fact that, for any speaker S, S's use of (M) expresses a proposition (i.e., <S does not believe p.>) that is inconsistent with the success conditions for that same (assertoric) use of (M), namely, that S's use of (M) is successful only if S believes p. That is, it is by virtue of this fact that S's use of (M) is successful only if it fails. By generalization, we can derive the following sufficient condition for assertoric self-defeat:

BELIEF-INCONSISTENCY TEST: Sentence 's' is *assertorically* self-defeating with respect to speaker S if there is a proposition p such that (i) S's use of 's' is successful only if S believes p, and (ii) S's use of 's' expresses a proposition that is inconsistent with the proposition <S believes p.>.

A serious obstacle for this approach, however, is that it is not entirely obvious that The Error Theory satisfies this condition. Consider the sentence:

(ET) The Error Theory is true.

According to the BELIEF-INCONSISTENCY TEST, (ET) comes out as assertorically self-defeating (with respect to speaker *S*) in case The Error Theory is inconsistent with:

(BET) <*S* believes The Error Theory.>

However, it is not entirely obvious that it is. As we previously saw, The Error Theory at most entails:

(NR) <There is no reason to believe The Error Theory.>

The problem is that there is no apparent reason to think (NR) is inconsistent with (BET).

Despite this difficulty, the approach is still noteworthy because several of Struemer's remarks forces him to concede that The Error Theory is assertorically self-defeating, at least when it is asserted by *him*. As I discuss later, Streumer's defense of The Error Theory heavily leans on the claim that the theory is literally *unbelievable*, i.e., that it is impossible for anyone to believe that The Error Theory is true. In support of this rather alarming claim, he argues for the following (equally alarming) theses about the nature of belief:

BELIEF OF CONSEQUENCES (BOC): One cannot believe *p* without believing the consequences of *p*.

BELIEF OF REASONS (BOR) One cannot believe *p* while believing that there is no reason to believe *p*.

For reasons of space, I do not recite or assess Streumer's arguments for the above theses.[85] For the sake of exposition, I simply explain how Streumer's acceptance of these theses commits him to the claim that The Error Theory is assertorically self-defeating.

The reasoning goes as follows. Because (NR) is a consequence of The Error Theory, according to (BOC), it follows that one cannot believe The Error Theory without believing (NR). According to (BOR), this implies that it is impossible to believe The Error Theory, which in turn implies that no actual speaker believes The Error Theory. According to (BOC) again, this implies that, if a speaker accepts (BOC) and (BOR), then she cannot believe The Error Theory without also believing that she herself does not believe The Error Theory. We thus arrive at the home stretch. If one asserts (ET), for that assertion to be successful, it must convey that the speaker believes The Error Theory. But as we saw above, a speaker who subscribes to (BOC) and (BOR) believes The Error Theory only if she believes that she does not believe it. Thus, if such a speaker successfully asserts (ET), it follows that the success conditions for the assertoric use of (ET) are not met. To put the point more conspicuously, according to the above reasoning, when a speaker who subscribes to Streumer's view of belief asserts (ET), she is committed to affirming:

(*) The Error Theory is true, but I do not believe The Error Theory.

And (*) is directly analogous to (M).

The discussion so far shows that, by *Streumer's own lights*, The Error Theory is assertorically self-defeating. And because his reasoning constitutes one of the few explicit defenses of The Error Theory in the literature, it gives us enough material to move on with the case study.

---

[85] See Streumer (2013, pp.195-199) for the full argument; for criticisms and responses, see Streumer (2017b).

Nevertheless, it will be helpful for us to have a way to talk about the self-defeating nature of The Error Theory that does not rely on such a controversial view about belief. To this end, I argue that The Error Theory is pragmatically self-defeating with respect to the act of advocation, or more simply, that it is *advocationally self-defeating*. The general idea goes like this. Plausibly, it is constitutive of the act of advocation that it conveys what the speaker takes to be reasons to believe what is being advocated. Thus, to advocate The Error Theory, one must provide reasons to believe The Error Theory. But because The Error Theory entails (NR), if one successfully advocates The Error Theory, one thereby ends up providing a refutation of it.

Like before, the leading idea here is that a successful act of advocating The Error Theory entails its own failure. However, unlike assertoric self-defeat, this thought involves two different senses of 'success' / 'failure'. On the one hand, an expressive act can be said to be "successful" in the sense that it successfully *qualifies* as an instance of a certain kind of expressive act. Let us call this sense 'constitutive success'. On the other hand, an expressive act can be said to be "successful" in the sense that it *achieves* what the relevant kind of expressive act is meant to achieve. Let us call this sense 'goal-related success'.[86] The previous discussion of assertoric self-defeat solely

---

[86] The distinction between constitutive and goal-related success is related to the distinction between constitutive norms and constitutive goals, which I discuss in chapter 4. In brief, constitutive norms are the norms one must follow in order to qualify as correctly participating in a given activity; constitutive goals are the norms that concern what is to be achieved within the activity in question. Constitutive success can be understood as an adherence to the constitutive norms of expressive acts; goal-related success can be understood as an adherence to their constitutive goals. The distinction also helps us further refine the previous general definition of pragmatic self-defeat. By distinguishing between constitutive and goal-related success, we may notice that the previous definition encompasses four different general notions of pragmatic self-defeat. The difference between these notions can be expressed as follows (let 'C' stand for 'constitutive success' and 'G' stand for 'goal-related success'):
    C-C pragmatic self-defeat: the use of an expression is constitutively successful only if it constitutively fails.
    C-G pragmatic self-defeat: the use of an expression is constitutively successful only if it goal-relatedly fails.
    G-C pragmatic self-defeat: the use of an expression is goal-relatedly successful only if it constitutively fails.
    G-G pragmatic self-defeat: the use of an expression is goal-relatedly successful only if it goal-relatedly fails.
Although it is technically possible to also distinguish between the different varieties of assertoric and advocational self-defeat, the only notions that are relevant to my discussion are C-C assertoric self-defeat and C-G advocational self-defeat.

concerned constitutive success: assertoric self-defeat was said to occur when an expressive act successfully qualifies as an instance of assertion only if it fails to do so. But this is not exactly the way I am arguing that The Error theory is advocationally self-defeating. When a speaker performs an act that constitutively succeeds at advocating The Error Theory, it does not seem to follow that the act in question fails to *qualify* as an advocational act.[87] Instead, it is more plausible that it entails the goal-related failure of that act. That is, it seems more plausible that The Error Theory is advocationally self-defeating in the following sense:

ADVOCATIONAL SELF-DEFEAT: An expression *e* is *advocationally self-defeating* $=_{\text{def}}$ for proposition *p* that is expressed by the use of *e*, an act of advocating *p* is constitutively successful only if it goal-relatedly fails by virtue of the fact that *p* is inconsistent with the goal-related success conditions for advocating *p*.

It seems plausible that an advocation of *p* is goal-relatedly successful if it provides reasons to believe *p*. By this, I simply mean that the goal is to put forth considerations that support the truth of *p*. But because The Error Theory entails (NR), the theory is inconsistent with there being any such reasons that can be provided. If any such reasons *are* provided, by modus tollens, it follows that The Error Theory is false. Thus, it is impossible for any constitutively successful act of advocating The Error Theory to achieve the goal of advocation.

---

[87] For this to follow, the fact that a speaker's expressive act successfully qualifies as an act of advocating The Error Theory must entail that the speaker does not believe that there are reasons to believe The Error Theory. But this requires that we assume (i) (BOC) and (ii) that it is constitutive of advocation that one believes what one advocates. And neither of these can be plausibly assumed.

3.2.4 The Pragmatic-Epistemic Gap

Let us take stock. Previously, we considered two ways to flesh out the charge that The Error Theory is pragmatically self-defeating. (Going forward, unless noted otherwise, my use of 'pragmatic self-defeat' is restricted to assertoric and advocational self-defeat.) Our discussion up to now illustrates the two lessons I aim to draw from this case study. First, it helps illustrate that pragmatic self-defeat (or more specifically, assertoric and advocational self-defeat) is a dialectical vice. Understood as a piece of criticism, it is a *vice* in the sense that it picks out something "bad" about a theory that discourages us from endorsing it. Second, it is a *dialectical* vice in the sense that the bad-making feature in question concerns the fact that any advocate of such theories will find themselves occupying an awkward dialectical position whether they must assert what is not assertable or advocate what cannot be advocated.

Second, it sheds further light on the difficulty in moving from (5) to (6). The difficulty lies in the fact that the (5) and (6) concern The Error Theory's relationship to two distinct things. Given our previous discussion, (5) concerns *our* (supposed) inability to either assert or advocate the truth of The Error Theory. That is, it concerns the theory's relationship to us *theorists*, or more precisely, the dialectical efforts we make in the course of our inquiry. In contrast, (6) is about its truth value. And this concerns the theory's relationship to *the world*. Thus, we cannot directly move from (5) to (6): for all we know, the true theory of normativity could be a theory that we are unable to assert or advocate.

The apparent gap between pragmatic self-defeat and falsity is an instance of something more general. There is nothing special about the domain of normativity and it appears to be equally arguable that, for all we know, the true theory of *any* domain could be a theory that is inconsistent with our dialectical efforts. Let us call this apparent gap between our dialectical efforts and truth the "pragmatic-epistemic gap." In Section 3.3, I make efforts to partially close the pragmatic-epistemic gap. To *completely* close the gap is to show that the above is not a genuine possibility: i.e., that the true theory of any given domain is discoverable (at least in principle) through our dialectical efforts. Instead of making this stronger claim, I argue for a weaker, partial solution: I argue that no theory can be coherently defended on the basis of the pragmatic-epistemic gap. As the final part of the case study, in the first half of Section 3.3, I explain why The Error Theory cannot be defended by appeal to the pragmatic-epistemic gap. In the second half, I expand this into an account of the epistemic relevance of dialectical virtues and vices in general.

Here are some preliminary considerations that will help us better appreciate what my proposed account aims to achieve. The first issue to note concerns what the account *does not* aim to achieve. As I discuss later, my argument in Section 3.3 rests on a substantive assumption about the relationship between our dialectical efforts and truth. I say the argument rests on an "assumption" because I am unaware of any non-question-begging arguments in support of it. This reflects a general difficulty when dealing with radical skeptics like The Error Theorist and the epistemic nihilist. For instance, consider how Terence Cuneo prefaces his criticism of epistemic nihilism, which equally carry over to The Error Theory (Cuneo 2007 pp.120-121):[88]

---

[88] The passage opens with the *assumption* that the epistemic nihilist accepts that the claim in question because his overall challenge is posed as a series of dilemmas. In our terminology, this is tantamount to proceeding with the charge that epistemic nihilism is pragmatically self-defeating.

Assuming that the epistemic nihilist accepts the claim that there are no epistemic reasons [to accept epistemic nihilism]—might we now have the material for an argument to the effect that epistemic nihilism is false? […] I believe not; there is no plausible non-question-begging argument to establish that epistemic nihilism is false. […] Were a person to press such an argument against epistemic nihilism, she would endeavor to establish that there are reasons by tacitly assuming that some statements provide evidential support to believe others. But if epistemic nihilism were true, it would be impossible that there were premises of an argument that provide evidential support for its conclusion.

The difficulty here is that the denier of epistemic facts is unlikely to be moved by epistemic considerations. Thus, any such considerations seem to inevitably beg the question against the nihilist.[89] Similarly, because my argument relies on the existence of epistemic facts, it shares the same fate as an argument against The Error Theory.

That being said, there are significant achievements that my argument aims to make. An argument that completely closes the pragmatic-epistemic gap and thereby refutes The Error Theory would also be an argument that solves *the* problem of radical skepticism. But this is not what I aim to do here. Instead, I aim to reveal what must be given up in order to uphold the pragmatic-epistemic gap and show that The Error Theory is even less tenable than its critics suggest. To sharpen the benefits of my argument, I briefly consider how theorists have previously responded to the pragmatic-epistemic gap and discuss some shortcomings that my argument aims to address.

---

[89] The issue here is reminiscent of Hume's "Skepticism with regard to reason" (*A Treaties of Human Nature* 1.4.1.).

3.2.4.1 Parfit's and Cuneo's Compromise

As previously mentioned, the issue of pragmatic self-defeat is mostly discussed in the context of epistemic nihilism. And in light of the pragmatic-epistemic gap, critics of epistemic nihilism often give up on claiming that the view is *false*. Instead, they often suggest that pragmatic self-defeat should be understood as providing a non-epistemic, pragmatic reason to reject epistemic nihilism. For instance, after making note of the pragmatic-epistemic gap, Derek Parfit (2011, p.293) goes on to point out that the criticism reveals "how deep such skepticism goes, and how blank this skeptical state of mind would be." And given his concession of the gap, it is clear that the depth and vacuity of this form of skepticism act as non-epistemic, pragmatic reasons to reject it. Another example is Cuneo's argument (2007, pp.120-121) that the previously quoted passage is meant to preface. After expressing the aforementioned doubts, he points out that the view implies that "there could be no argument for anything," which he criticizes as being "sufficiently unattractive that any minimally adequate philosophical theory will be at pains to avoid being committed to it" (2007, pp.122). Again, given his preface, it is clear that Cuneo also takes this criticism as appealing to certain non-epistemic, pragmatic standards of the "attractiveness" and "adequacy" of theories.[90]

---

[90] A similar dialectical situation can be found outside of the context of epistemic nihilism. E.g., during his exploration of personal ontology (the pursuit of the question 'What are we?'), Eric Olsen (2007, pp.208-210) identifies a number of philosophical claims that he takes to be "pathological," i.e., philosophical claims that are impossible to live by (e.g., solipsism, the denial of free will, etc.). One such pathological view he considers is the rejection of practical normativity according to which there are no reasons for our actions. Although he concedes that the pathological nature of this view does not provide a reason to think that it is *false*, he does suggest that it provides a kind of criticism that one might find "satisfying" to some degree or other. In relation to my discussion so far, it seems natural to read him as rejecting pathological views on the basis of non-epistemic, pragmatic reasons. I thank Simon Langford for bringing my attention to this example.

Applied to The Error Theory, Parfit's and Cuneo's arguments translate to abstaining from claiming (6) and instead reaching for the weaker conclusion that The Error Theory "ought to be rejected" in a non-epistemic, pragmatic sense of 'ought'. From a very general perspective, I think Parfit and Cuneo are correct to appeal to practical consequences. In fact, in light of the previous remarks about radical skepticism, it may as well be inevitable. But I think they are considering the consequences of the wrong thesis. By focusing on the practical consequences of epistemic nihilism (and by extension, The Error Theory), their arguments amount to a cautionary tale that vividly describes just how practically or pragmatically unattractive epistemic nihilism is. But instead of being engrossed in the practical implications of epistemic nihilism, we can make progress on the epistemic front as well by critically assessing the practical implications of upholding the pragmatic-epistemic gap.

3.2.4.2 The Mere Defender's Rejoinder

An additional drawback of Cuneo's and Parfit's compromise is that it falls to Streumer's (2013) rejoinder to various kinds of practical arguments. As mentioned earlier, Streumer argues that it is literally impossible to believe The Error Theory, even for its own defenders. (In this light, the defenders of the Error Theory are not, strictly speaking, Error Theorists. Instead, they are *mere defenders*.) But by appealing to the pragmatic-epistemic gap, he argues that this has no bearing on the truth of the theory. Furthermore, he wields the unbelievability thesis to block a variety of objections that appeal to the practical consequences of accepting The Error Theory. For the sake of illustration, here are two examples (Streumer 2013, pp.206-208):

THE NORMATIVE OBJECTION: The Error Theory is a malignant view that threatens all of our normative judgments, including our deepest and most important moral convictions (e.g., that it's morally wrong to torture innocent children for fun). Therefore, we should reject the Error Theory.

THE OBJECTION FROM BAD FAITH: Normally, when we believe that one of our own beliefs is untrue, we give up this belief. But self-proclaimed Error Theorists rarely give up all of their normative judgments. Therefore, they are guilty of a form of bad faith.

Streumer responds to these objections by pointing out that the unbelievability of the Error Theory prevents any *actual theorist* from facing these kinds of consequences. For the NORMATIVE OBJECTION, he argues that The Error Theory is not malignant in the aforementioned way because it is impossible for an actual theorist's normative judgments to be (rationally) undermined in light of the theory; thus, the theory fails to pose a threat to any of our important normative judgments. For the OBJECTION FROM BAD FAITH, he argues that defenders of the theory are not intentionally ignoring the truth but are instead unable to believe it; thus, the defenders do not exhibit bad faith.[91]

The lessons here are twofold. The first is about the point of focus. Intuitively, the objection from self-defeat has a different *flavor* from the kinds of objections represented above. But on

---

[91] Complication: Streumer's full response involves his claim that we can "come close" to believing the Error Theory by believing different parts of it at different times (Streumer 2013, pp.202-203). I omit this for issue for the sake of brevity.

Parfit's and Cuneo's approach, the difference disappears, and the issue of self-defeat turns out to be one among the many undesirable consequences of The Error Theory. And to make matters worse, Streumer's strategy appears to be effective against these kinds of objections.[92] In this light, it is desirable to develop the criticism from self-defeat in a way that is immune to this kind of response, both for the sake of garnering a better understanding of the issue of pragmatic self-defeat and for the sake of developing it into a stronger line of criticism against The Error Theory. The second is about the intended audience. Like many, neither Parfit nor Cuneo take themselves to be engaging with an actual theorist. Rather, they engage with an imaginary antagonist from whom we are suggested to distance ourselves. But since Streumer is an actual person who explicitly defends The Error Theory, we are in need of an argument that applies not only to the hypothetical believer, but to the actual, mere defender as well.

3.3 The Epistemic Relevance of Dialectical Vices

3.3.1 Concluding the Case Study

To recap, the goal of the case study is to explore a way to partially close the pragmatic-epistemic gap by providing reasons to believe that the true theory of normativity cannot be outside

---

[92] In response, Spencer Case (2019) attempts to break the apparent impasse by delegating the decision to a "neutral arbiter" (by which he likely means a hypothetical ideally rational agent) who is neither committed to accepting or rejecting The Error Theory. Interestingly, he also argues that a demonstration that a philosophical position cannot be held by anyone is the strongest possible evidence we could practically expect from the critic of the theory. However, Case's argument appears to presuppose that the pragmatic-epistemic gap can be closed. Otherwise, there is no reason to think that the "evidence" obtained by the neutral arbiter's rulings has any bearing on truth. In this light, my argument in the following section can be seen as supplementing Case's argument.

of what we can assert or advocate. To this end, I point to presuppositions of the act of debating and exchanging arguments.

I start with the following claim: one's presence in a debate presupposes that one takes exchanging arguments to be, in principle, the kind of activity that guides us to the truth of the matter, given that one believes there is such a truth to be guided towards. By this, I do not mean that there must be a fact of the matter for every debate. There are many debates that are not over facts of the matter (e.g., debates about whether to get a dog or a cat) and it may turn out that some of the debates we engage in are not really about anything (e.g., debates about the "true principles" of alchemy). What I mean is that, in case the participants believe that there is a truth to be discovered, they must also believe that they can arrive at it through argumentation. I also do not mean that the act of debating is guaranteed to lead us to the truth sometime or later. Some debates may fail to be settled in a given moment in history and there may be debates that are never settled during our limited time in the universe. The claim is about what we can achieve *in principle.* Lastly, I do not claim that argumentation is *in fact* an in-principle method to discover truths. Rather, I claim that the participants must accept that it is in order for their presence in the debate to be meaningful. That is, in case one believes that there is a truth to be discovered but also believes that the truth cannot be discovered through argumentation, then one's words in the debate hall would be pointless.

Below, I further refine this claim and its implications when I give the general account for dialectical vices. But the above gloss suffices for the case study in that it shows the epistemic gap is significantly narrower than it may appear to be. Defending the Error Theory by appealing to the pragmatic-epistemic gap entails that the true theory of normativity cannot be discovered through argumentation. Thus, if one accepts this possibility, her participation in the debate would be

epistemically meaningless in the sense that she would be participating in an activity that, by her own lights, does not bring us any closer to the truth. The point here is not that the defense is self-defeating in the sense that one's defense undermines the position one is defending. Rather, it is that such an attempt is incoherent. By participating in a debate, one thereby commits to rejecting that the pragmatic-epistemic gap holds for whatever the debate is about. Thus, if one defends The Error Theory by appealing to this gap, one's remarks about the matter fail to properly count as being part of the debate. The same point can be made from the opposite direction. If one enters the debate over normativity, then one must presuppose that the truth of the matter is discoverable through argumentation. This holds for *everyone* in the debate, regardless of whether one opposes, believes, or merely defends The Error Theory. And this presupposition rules out the possibility that the true theory is pragmatically self-defeating; because such a theory cannot be represented in the debate at all, it is impossible for the debate to be settled in its favor. Hence, insofar as the advocate of The Error Theory are present in the debate, it is incoherent for them to uphold the pragmatic-epistemic gap.

The above argument suggests that the advocate of The Error Theory cannot concede that the Error Theory is pragmatically self-defeating, nor should the critic accept this as a viable move. Thus, if The Error Theory is to be defended at all, it must be shown *not* to be pragmatically self-defeating. I do not believe that this can be done. But even it is, we may at least say that Streumer's mere defender is no better off than the sincere believer of The Error Theory. A noteworthy limitation of my argument, however, is that it does not completely close the pragmatic-epistemic gap. At most, it shows that *no participant* of a given debate can coherently uphold the possibility that the truth of the matter cannot be discovered through the process of argumentation. But the Error Theorist could accept everything I say and simply refuse to engage in the act of debating. In

this case, my argument fails to apply. I accept this as a limitation and am unaware of a non-question-begging argument that can move the silent Error Theorist. However, this is not a special problem for my argument. In relation to my previous remarks, it is part of the problem of radical skepticism, as it involves doubts about the power of rationality, if not our capacity to grasp the truth. This is a problem we all face, and I do not intend to solve it here.

3.3.2 Generalizing on the Case Study

The backbone of the previous argument is the claim that the participants of a debate must presuppose that argumentation is connected to truth. To expand this into an argument for the epistemic relevance of pragmatic self-defeat, we need a means to evaluate the veracity of this presupposition. If this presupposition turns out to be false and our faith in the process of argumentation is misplaced, then it would be illegitimate to take the results of our dialectical interactions to be indicative of which theory is true. This brings us to more or less the same difficulty mentioned above. In order to give a non-question begging argument to the effect that argumentation is indeed connected to truth, the argument must somehow demonstrate that the connection holds without presupposing its own capacity to reveal the truth of the matter. And it is hard to see how this could be done. But again, this is the general problem of radical skepticism, which is not what I aim to solve here. Thus, I do not attempt to *argue* for the connection between argumentation and truth. Instead, I simply *assume* that the connection holds and focus on the task of describing its nature.

While the above remarks are made for pragmatic self-defeat, it can easily be expanded into an account for dialectical virtues and vices in general. Pragmatic self-defeat is likely to be the most severe dialectical vice there is, and perhaps is the only vice that defeats any comparison of theoretical virtues and vices. But if there are any other dialectical virtues that are epistemically relevant, we may reasonably expect them to be more loosely connected to truth, namely by affecting the likelihood of the relevant theories being true.[93] The general idea goes like this. Dialectical virtues and vices are exhibited by features of theories that give their advocates a leg up or a leg down in debates. One way to understand this is in terms of a probability boost with respect to whether a given theory "wins out" in a debate.[94] Similarly, dialectical vices can be understood as features that decrease this probability. (I take dialectical virtues and vices to be comparative notions on a continuum, rather than taking each of them to be substantive. For the sake of brevity, I proceed by only referring to dialectical virtues.) Here, I do not give an exact definition of what it is for a theory to "win out" in a debate. But a rough account can be given as follows. Intuitively, the fate of a theory depends on its advocate's ability to discharge the dialectical obligations she takes up in the context of the debate. For example, one such obligation is the *defense obligation*, which I discuss in Chapter 4. In a nutshell, the defense obligation concerns one's obligation to defend the claims one makes in the context of the debate. More specifically, one is obliged to defend one's claims against the legitimate challenges that are raised by her interlocutor. And in the event that one fails to defend a claim that is central to her position, we may say that she (and her theory) is thereby "defeated" in the debate. In what follows, I leave open what other kinds of

---

[93] One may think it is odd to evaluate the probabilities of philosophical theories. For instance, if one thinks philosophy (especially metaphysics) mostly concerns with necessary truths, the probability of philosophical theories should be either 1 or 0. We can get around this by taking these probabilities as subjective probabilities such as *credence of beliefs*. See Brenner (2017) for further discussion.

[94] Formally, feature F is a dialectical virtue $=_{def}$ P(Theory T is victorious | Theory T has feature F) > P(Theory T is victorious). *Mutatis mutandis* for features that are dialectic vicious.

dialectical obligations there are. But at a general level, we may define the relevant sense of victory as: theory *T wins out* in debate *D* iff the advocate of *T* is the only participant of *D* that successfully discharges all of her dialectical obligations regarding *T* in the context of *D*.

Given the connection between argumentation and truth, we may expect that a theory's being victorious positively affects its likelihood of being true. The way in which this latter probability is affected depends on how tight the former connection is. One way to think about the matter is:

TRUTH PREVAILS: In principle, theory *T* wins out in every debate if and only if *T* is true.

The idea here is to maintain a connection tight enough to *guarantee* that, in principle, the act of exchanging arguments leads us to *the* truth of the matter. There are two purposes for taking this to hold "in principle." First, it is meant to deal with cases of rational error. In certain cases, it seems possible (and probably mundane) for a debate not to be settled in the true theory's favor because the participants collectively misinterpret the evidence. To rule these kinds of cases out, we may restrict our attention to idealized debates that only involve fully rational agents. Second, it is meant to deal with misleading evidence. In localized cases, even fully rational agents may be led astray when the currently available evidence prescribes the wrong theory. The solution, then, is to think in terms of what the evidence prescribes when we have access to all the evidence there is to gather. I use the term 'evidence' broadly and refer to anything that can get a philosophical argument off the ground. Although I remain neutral on the matter, this can be understood as encompassing

sentences that are particularly difficult for the Quinean to paraphrase or empirical facts that conflict with a philosophical theory.[95] (Think of this to be analogous to the physicalist's appeal to "final physics.")

With TRUTH PREVAILS, we have a straightforward account of the epistemic relevance of dialectical virtues. Say "the final victor" is the theory that wins out in every debate in which it is involved. On TRUTH PREVAILS, the probability of a theory's being true is equal to the probability of its being the final victor. Hence, anything that probabilistically supports a theory's being the final victor should also probabilistically support its being true. And since a theory's being dialectically virtuous increases its probability of being the final victor, it also increases its probability of being true.[96,97]

One drawback of a connection this strong is that it rules out the possibility of ties. And even without being pessimistic about the act of debating, it may be desirable to admit some degree of humility. To this end, we may relax the connection by instead adopting:

---

[95] The debate between Huemer (2009) and Bradley (2017) briefly touches on issues of adopting this broad sense of evidence.

[96] The formal argument is as follows ('V' = 'Theory T is the final victor, 'T' = 'Theory T is true', 'D' = 'Theory T has a particular dialectical virtue D'):
1. $P(V) = P(T)$
2. $P(V \mid D) > P(V)$
3. Therefore, $P(T \mid D) > P(T)$

[97] One might question the probabilistic support from being dialectically virtuous to being the final victor, since, for all I've said, dialectical virtues probabilistically support a theory's winning in *a* debate, not all of them. While this worry can be addressed by pointing out that the one probabilistically supports the other, it invites the problem of transitivity I discuss below. Fortunately, the solution I discuss below is applicable here as well, as being the final victor screens off a theory's being dialectically virtuous with respect to its being true.

TRUTH COMPELS: In principle, theories that are undefeated in every debate in which they are involved are substantially more likely to be true.

Unlike the previous thesis, TRUTH COMPELS maintains a probabilistic connection between being undefeated and being true. This way, when more than one theory comes out undefeated and we have several "finalists" rather than a single final victor, we may humbly admit that we are unable to pick out which theory is true while maintaining that any of the finalists are substantially more likely to be true than those that failed to make it to the final round. What counts as "substantially more likely" depends on one's preferred view of epistemic humility. On a minimalist view, one may take the probability of a non-finalist being true is 0 and the probability of being true to be equally distributed among the finalists. If one thinks it is appropriate to maintain a greater level of humility, one may take it to be more likely than not that the true theory is among the finalists, in which case the sum of their probabilities would be some value greater than .5. On a view that admits an even higher level of humility, one may perhaps take the sum of their probabilities to be lower than .5. But again, on the pain of rendering argumentation obsolete, there must be a limit to how much humility is tolerable and how low this sum can be.

This relaxed connection invites a complication, however. The account that results from TRUTH COMPELS requires that the probabilistic support in question is transitive. That is, it requires that the probabilistic support from dialectical virtues to truth is mediated by the probabilistic support from being one of the finalists to truth.[98] The problem is that probability

---

[98] The formal argument is as follows (the same translation scheme applies as before except 'F' = 'Theory *T* is one of the finalists'):
1. $P(T \mid F) > P(T)$

boosts often fail to be mediated in this way. Consider Tomoji Shogenji's counterexample to transitivity (Shogenji 2003, p.1). Arguably, the following are true:

(1) One's being an academic philosopher makes it more likely that one has a doctoral degree.

(2) One's having a doctoral degree makes it more likely that one is well paid.

However, since one's being an academic philosopher may (and perhaps does) negatively affect the probability of one's being well paid, we cannot infer from these two statements that:

(3) One's being an academic philosopher makes it more likely that one is well paid.

Fortunately, Shohenji offers a solution that carries over to our case. The desired inference has the following form (let the variables range over propositions).

(4) X raises the probability of Y.

(5) Y raises the probability of Z.

(6) Therefore, X raises the probability of Z.

In order for the inference to go through, we need to rule out the possibility that X negatively affects the probability of Z. In the same article, Shogenji proves that this can be done in case a screening-

---

2. $P(F \mid D) > P(F)$
3. Therefore, $P(T \mid D) > P(T)$

Informally: A theory's being among the finalists makes it more likely that the theory is true; A theory's being dialectically virtuous makes it more likely to be among the finalists; Therefore, a theory's being dialectically virtuous makes it more likely that the theory is true.

off condition is met. Formally, Y screens off X with respect to Z in case the following condition is met (the following is a sufficient condition for transitivity):[99]

SCREENING-OFF (SO): $P(Z \mid X \,\&\, Y) = P(Z \mid Y) \,\&\, P(Z \mid X \,\&\, \sim Y) = P(Z \mid \sim Y)$

Informally: SO requires that (i) the probability of Z given X and Y is equal to that of Z given Y alone and that (ii) the probability of Z given X and not-Y is equal to that of Z given not-Y alone. That is, it requires that, once the truth value of Y is determined, adding X to the evidence pool does not affect the probability we assign to Z.[100] If this condition is met, Y prevents X from influencing the probability of Z in any way, so, *a fortiori*, it blocks off any potential negative influence from X to Z. Hence, transitivity is preserved.

Returning to dialectical virtues, transitivity holds in case a theory's being among the finalists screens off its being dialectically virtuous with respect to its being true. And it is easy to see that it does. Whether or not a theory is among the finalists is part of the culmination of the dialectic—the result of the dialectic fully running its course. And after the dialectic has been completed, any advantage or disadvantage in the dialectic arguably has no bearing on the truth of theories. That is, once we know whether or not a given theory is among the finalists, adding the fact that the theory is dialectically virtuous should not affect the probability of the theory's being true in any way. To maintain otherwise is to say that these advantages and disadvantages in the

---

[99] William Roche (2012) shows that transitivity is preserved under weaker conditions as well. For instance, when $P(Z \mid X \,\&\, Y) \geq P(Z \mid Y)$ and $P(Z \mid X \,\&\, \sim Y) \geq P(Z \mid \sim Y)$. Roche also explains that this weaker condition is not necessary for transitivity, either.

[100] A concrete example may be helpful. <Killer stabbed Victim> screens off <There is a testimony that Killer stabbed Victim> with respect to <Killer committed murder>. This is because, once the truth value of <Killer stabbed Victim> is determined, the fact that someone testified against Killer makes no difference to the probability of whether Killer committed murder.

dialectic are related to truth outside of the dialectic context, which is absurd. Thus, the transitivity problem is solved.

I concede that both TRUTH PREVAILS and the variants of TRUTH COMPELS are very strong assumptions to make. However, I take these assumptions to be widely accepted (though only implicitly) on the pain of radical skepticism. If my claim about the act of debating is correct, anyone who thinks philosophy is in the business of discovering truths is committed to some suitable version of TRUTH COMPELS. Thus, on my account, any such person already believes that (at least some) dialectical virtues and vices provide guidance to truth.

CHAPTER 4: THE DIALECTICAL VICE OF PRIMITIVISM

4.1 The Argument

In the previous chapter, I explored the general relationship between the act of debating and the discovery of truths. In this chapter, I examine this relationship in the more specific context of reductionism and primitivism. To recap, the main goal of this dissertation is to shed light on the general appeal and attractiveness of (ideological) reductionism. As a means to sharpen the discussion, I focus on two commonly held attitudes regarding the relationship between reductionism and primitivism, namely:

UPSPOKEN RULE: If something can be reduced, it should be reduced.

LAST RESORT: One ought to accept primitivism only as a last resort when all other accounts fail.

I argue that these two attitudes are driven not by a positive aspect about reductionism, but instead by a dialectical disadvantage that can be associated with primitivist positions in general. In a nutshell, while the participants of a debate are required to defend the claims they make in a debate, I argue that it is often particularly difficult for the primitivist to properly meet this requirement. And in light of our discussion in the previous chapter, I argue that this disadvantage is not a mere

quirk of our practice of exchanging arguments but is something that has serious implications when evaluating a primitivist theory of various domains.

My overall argument goes like this. I start by deriving a sufficient condition for being defeated in a debate. I derive this condition from a basic dialectical norm that concerns one's obligation to defend one's claims. (Note: I use 'claim' and 'asserted proposition' interchangeably.) As I discuss later, there is some disagreement over how this norm should be precisely formulated. Let 'The Defense Obligation' be a neutral way to refer to this norm. According to some (Rescorla 2009b; 2009c), The Defense Obligation is to be identified with:

> THE DEFENSE NORM: When challenged to defend an asserted proposition, one must either defend it or else retract it.

The thesis captures the plausible idea that, for one to properly respond to a challenge, one must either provide a defense (i.e., a cogent, non-circular argument) or retract the claim that is being disputed. However, some theorists deny that one is required to respond to *all* challenges in this fashion. Instead, they argue that The Defense Obligation should be qualified in the following way. Say a challenge to an asserted proposition *p* is *legitimate* iff the norms of debating are such that one cannot dismiss the challenge without either defending *p* or else retracting it.[101] According to these theorists, The Defense Obligation should be understood as:[102]

---

[101] This definition is meant to be more or less equivalent to what appears in Rescorla (2009b, pp.45-46).
[102] This is the thesis Rescorla (2009b; 2009c) attributes to his opponents.

THE DEFAULT-CHALLENGE NORM: When *legitimately* challenged to defend an asserted proposition, one must either defend it or else retract it.

Later in Section 4.3, I give a more detailed presentation of this debate. For now, let us set this issue aside. To this end, let us restrict our discussion to legitimate challenges while leaving open whether there are any challenges that are in fact illegitimate in the relevant sense. What matters for my current purposes is that, on either conception of The Defense Obligation, one is required to retract the claims that one is unable to properly defend. That is, on either conception, The Defense Obligation plausibly entails:

(*) When legitimately challenged to defend an asserted proposition p, if one is unable to defend p, then one must retract it.

At face value, (*) only seems to express the requirement to respond to challenges that are *actually* raised. But given our interests in philosophical debates, it is desirable to extend our discussion to *possible* challenges as well. Suppose you assert p in the course of a debate. At some point, you but not your interlocutor become aware that p cannot be defended against a particular challenge. In this case, the fact that your interlocutor is oblivious of this challenge should not put you in the clear. Rather, it seems that you ought to retract your assertion of p regardless of whether

it is actually challenged by your interlocutor.[103] In this light, it is reasonable to expand (*) as follows, which is also the first official premise of my argument:

(1) For any proposition *p* one asserts, if one is unable to defend *p* against a possible legitimate challenge, then one must retract *p*. (Premise)

Using (1), we can derive a sufficient condition for being defeated in a debate. In some cases, retracting a given claim may not seriously affect the overall case one is making in a debate. E.g., one may be able to make the same point by asserting a different proposition in its place. But intuitively, there are claims the retraction of which is much more damaging to one's dialectical situation. Most notably, if one retracts the very proposition that expresses one's position in a debate, it is plausible to say that one thereby forfeits the debate in question. But we may also think of other propositions that one's position is dependent on. Let us cast a wide net and say that, among the propositions one asserts, a proposition is *central* to one's position in a debate iff one cannot retract *p* without forfeiting the debate.[104] Together with (1), this notion of centrality entails the following sufficient condition for being defeated in a debate:

(2) For any proposition *p* that is central to one's position in a debate, if one is unable to defend *p* against a possible legitimate challenge, then one is defeated in that debate.

(By 1, definition of centrality)

---

[103] To clarify, I do not claim that this holds for *any* conversational context. E.g., in a court setting, it is reasonable to think that an attorney is only obliged to respond to the challenges that are actually raised in court. But the point I am making seems plausible for most debates in an academic context.

[104] The purpose of this expansion is to iron out some technicalities. For instance, if you are a modalist who posits primitive modal operators, then the claim that these modal notions are primitive is baked into the very position you are defending. In contrast, if you are a universal theorist who posits a primitive notion of *instantiation*, it is not entirely obvious that this is part of *the position* that you are defending in the debate. But by casting a wide net like this, we need not worry about such details: insofar as your theory cannot be defended without introducing the notion of instantiation, my argument below applies just the same.

Now consider what this implies for the primitivist's dialectic. Let us abbreviate 'primitivism about notion *I*' as 'I-primitivism'. Also, let us call a claim that a given notion is primitive a 'P-claim'. For the reasons expressed above, the following is trivially true:

(3) The P-claim for *I* is central to primitivism about *I*. (Premise)

Let a challenge to an assertion of *p* is *brute* iff it is a mere request for the justifications for *p* and is not accompanied by any countervailing considerations. I submit that it is legitimate to brutely challenge any given P-claim. Let us call those who endorse I-primitivism an 'I-primitivist'. My claim is that:

(4) It is legitimate to brutely challenge the I-primitivist's P-claim about *I*. (Premise)

Because a brute challenge is not accompanied by any supporting considerations, there is no particular issue it requires one to address. Instead, it baldly requires one to provide justifications for the relevant claim. Thus, from (2)-(4), it follows that:

(5) If the I-primitivist cannot provide justifications for her P-claim for *I*, then she is defeated in the debate. (By 2-4)

In simpler terms, the argument so far is meant to show that (i) the primitivist has the burden of proof regarding her central P-claim, and that (ii) her fate in the dialectic is directly tied to her ability to discharge this burden. And in many cases, I argue that it is genuinely difficult for the primitivist to discharge this burden of proof. Generally speaking, there appears to be no positive, non-circular argument one can give in defense of *any* P-claim. Often times, the only way forward is through the process of elimination: by showing that no reductive procedure for the relevant

notion can be successful. And in such cases, I argue that the primitivist is at a general disadvantage. In sum, I argue that:

(6) There is virtually no positive, non-circular argument one can give in defense of a P-claim. (Premise)

And from (5)-(6), I draw the general conclusion that:

(7) The primitivist faces a comparative disadvantage in the dialectic. (By 5, 6)

In the rest of the chapter, I provide more background for the above argument and spell out the key ideas expressed above. In Section 4.2, I provide more background on how to understand the normative talk I am engaging in. Section 4.3, I give a fuller presentation of The Defense Obligation and discuss how it relates to my argument. Specifically, I address premise (4) and argue that the premise is acceptable on either of the previously mentioned conceptions of the norm. In Section 4.4, I focus on steps (6)-(7), discuss how it explains UNSPOKEN RULE and LAST RESORT, and address some possible objections.

4.2 Constitutive Norms and Constitutive Goals

Before diving in, it will be helpful to clarify a few points about the nature of the requirements and obligations we are concerned with.[105] There are two senses in which one is "required" to defend one's claims (given that one is able to do so), both of which are relevant to our discussion. One sense of this requirement concerns the *constitutive norms of correctness* of

---

[105] Much of the exposition on constitutive norms in this subsection are taken from Michael Rescorla (2009b).

debating, or its *constitutive norms* more simply. The constitutive norms of an activity are the norms one must follow in order to qualify as correctly participating in the activity in question. And like any other activity, there are norms that are constitutive of debating. For a more mundane example, consider the act of playing soccer. The most well-known rule of soccer is probably the handling offense; while the ball is in play, players other than the goalie are prohibited from deliberately touching the ball with their hands or arms. Supposing that the handling offense is a constitutive norm of soccer, soccer players are *required* not to intentionally touch the ball with their hands in the sense that they cannot correctly play soccer without doing so. In a similar vein, the aforementioned norms express that the norms of debating are such that one cannot correctly participate in a debate without responding to challenges in the previously mentioned ways.

The other sense of 'requires' concerns the *constitutive goal* of debating. Constitutive goals are norms that concern what is to be aimed or achieved within the activity. And these are different from constitutive norms in that one may disregard the constitutive goal of an activity while not violating any of its constitutive norms. Consider again the case of soccer. Arguably, the game of soccer is structured around the constitutive goal of scoring more goals against the opposing team than they can against yours—that is what soccer players are supposed to be attempting to achieve. But one can disregard this constitutive goal while following every constitutive norm of the game. For instance, suppose a player is involved in a match-fixing scheme and deliberately obstructs her team's efforts to win. In this case, while the player violates the constitutive goal of soccer, she still has to "play by the rules" and play the game correctly to make sure that the opposing team wins (and to avoid getting caught). Applied to the act of debating, the constitutive goal of a debate is arguably to "win," or, at the least, to avoid being defeated. That is, each participant of a debate is primarily meant to strive at demonstrating one's own position as the epistemically superior option

on the table. In this sense, we may say one is "required" to defend one's claims in the sense that it is conducive to the constitutive goal of debating. That is, although one may correctly participate in a debate while foregoing its constitutive goal, it lays out what is required for one to participate sincerely.

One final clarification: some may think that my remarks above wrongfully imply that the purpose or value of debating entirely lies in convincing others of your position.[106] But no such conclusion follows. In activities like soccer, we may easily distinguish the constitutive goal that the players pursue from the *reason* a person decides to play soccer at a given moment. By saying that the constitutive goal of soccer is to score goals, all I am saying is that the game of soccer is structured around the attempt to score goals: that's how you play the game. And this by no means implies that a person becomes a player on the pitch for the sole purpose of scoring goals. Instead, people play the sport for a wide variety of reasons, e.g., one may play it as a profession, to stay in shape, to spend time with their friends, etc. Similarly, we may distinguish the constitutive goal of debating from the reason one enters a debate. In some cases, the constitutive goal of debating *may* coincide with one's reason for engagement. For instance, we often challenge claims made by others primarily because we think they are false. But one may engage in a debate for a different set of reasons: e.g., to cultivate a better understanding of one's intellectual peers, or perhaps as a pass-time game of intellectual chess. In these cases, the reason one decides to engage in a debate is not necessarily to "win" and to convince others of one's position. Nevertheless, it remains true that the activity itself is structured around the demonstration of positions. The way we learn more about our peers' perspective is by asking how they would respond to certain challenges; the way we let

---

[106] I thank Erin Wiebe for bringing this issue to my attention.

the pass-time continue is by showing that the matter (however frivolous) is yet to be settled, and so forth.

4.3 The Defense Obligation

4.3.1 The Two Conceptions of The Defense Obligation

In this section, I give a fuller presentation of The Defense Obligation by discussing the debate over how it should be formulated. My discussion here has two main purposes. First, it is intended to provide background information about The Defense Obligation and how it should be understood. Second, I aim to show that my overall argument in this chapter is compatible with the two conceptions of The Defense Obligation that are in competition. In what follows, I provide a brief overview of the debate and touch on some issues that arise for each conception.

The two conceptions are respectively backed by the *dialectical foundationalist*, who endorses THE DEFAULT-CHALLENGE NORM, and the *dialectical egalitarian*, who endorses THE DEFENSE NORM.[107] The two camps disagree about the extent to which we are supposedly required to defend our claims. More specifically, they disagree over whether we are ever entitled to dismiss a given challenge without providing a defense nor retracting our claim (i.e., whether there are any challenges that are not legitimate). According to the dialectical foundationalist, the answer is that we sometimes are. The exact details vary between different versions of the view, but

---

[107] My presentation of the debate heavily draws from Michael Rescorla's (2009b; 2009c) work on the topic, although I later criticize some of the conclusions he arrives at. It is also worth noting that, while his characterization of dialectical foundationalism is technically neutral over the many variations of the view, his overall language does seem to apply to Robert Brandom's (1994) view more naturally. My presentation here inherits this aspect of Rescorla's exposition.

the core idea is that there are certain propositions that express a privileged "default" position in the dialectic in the sense that the burden of proof supposedly falls on those who oppose them rather than those who assert them. [108] Accordingly, it is said to be illegitimate to challenge these propositions without providing supporting considerations. To summarize, say a proposition *p* is *dialectically basic* iff it is not legitimate to brutely challenge *p*. Dialectical foundationalism, then, is the view that dialectically basic propositions exist. In opposition, dialectical egalitarianism is the view that no propositions are dialectically basic. Accordingly, the egalitarian argues that we are never entitled to dismiss a challenge without a defense: for her, *all* challenges to any proposition are legitimate. As a result, they take the restriction to legitimate challenges to be redundant, in which case, THE DEFAULT-CHALLENGE NORM reduces to THE DEFENSE NORM.

In relation to my own argument, the only premise that this debate potentially bears on is (4). Accordingly, the main task of this section comes down to showing that (4) is compatible with both dialectical foundationalism and egalitarianism. At the surface level, it is easy to see that dialectical egalitarianism raises no tension with (4): if *all* challenges are legitimate, then it is *a fortiori* legitimate to brutely challenge any given P-claim. However, there are some deeper issues to be wary of. Specifically, I discuss below that there are aspects of Rescorla's defense of the view that clashes with my overall stance in this chapter. For dialectical foundationalism, it is also easy that the view lays the grounds for a possible objection to (4): one may reject (4) by endorsing foundationalism and arguing that there are at least some P-claims that are dialectically basic. In

---

[108] See Rescorla (2009c, pp.88-89) for a brief survey of the variations. For commonly provided examples, Rescorla (2009c, p.89) summarizes them as: (a) reports about one's mental states; (b) perceptual reports; (c) basic autobiographical reports; (d) "hinge" propositions (e.g., that the world has existed for more than five minutes, etc.); (e) elementary logical and mathematical truths; and (h) elementary analytic truths.

response, I argue that, even by the foundationalist's lights, P-claims should not be accepted as dialectically basic. Let us look into these two issues in turn.

4.3.2 Rescorla on Dialectical Egalitarianism

Overall, Rescorla gives a compelling defense of dialectical egalitarianism. But there are certain aspects of the way he delivers it that conflict with my purposes in this chapter. Let us first take a look at how he defends the view. His defense starts by attacking one of the main arguments for dialectical foundationalism. Typically, dialectical foundationalists motivate their view by arguing that dialectical egalitarianism entails a global form of skepticism. The argument rests on a regress problem. Say a *persistent interlocutor* is someone who brutely challenges all of a speaker's assertions.[109] Now suppose a persistent interlocutor challenges you to defend a given claim. To meet this challenge, you arguably need to provide a cogent, non-circular argument. But to do so, you need to make more claims, which are again brutely challenged by your interlocutor. To meet this further challenge, you would need to make even more claims, which are also brutely challenged, so on and so forth. Following Rescorla, call this 'the dialectical regress'. If dialectical egalitarianism is true, then all of the persistent interlocutor's brute challenges are legitimate: at every step, you would be obliged to make further claims that are vulnerable to your interlocutor's brute challenges. Thus, dialectical egalitarianism seemingly entails that:

(8) The dialectical regress is an infinite regress.

---

[109] Rescorla credits Adam Leite for this term.

From here, the foundationalist draws two key inferences. First, she takes (8) to entail that:

(9) It is impossible for one to vindicate any assertion against a persistent interlocutor.

Second, she takes (8) to entail the skeptical conclusion that:

(10) One ought to withhold judgement from all propositions.

On the pain of this skeptical conclusion, the dialectical foundationalist argues that we must stop the regress by taking certain propositions to be dialectically basic. The introduction of dialectically basic propositions serves this purpose by allowing there to be a point in the dialectical regress where the burden of proof is supposedly shifted from the speaker to the interlocutor (Rescorla 2009b, p.45; 2009c, p.88). That is, once your conversation with the persistent interlocutor reaches a point where you assert a dialectically basic proposition *p*, because your assertion of *p* is immune to brute challenges, the onus is on your interlocutor to provide considerations that undermine *p*. And in the event that your interlocutor is unable to do so, the dialectic is suggested to come to an end.

In response, Rescorla argues that the dialectical egalitarian can block the skeptical conclusion by rejecting the inference from (9) to (10) (Rescorla 2009b, pp.48-50). He points out that, in making this inference, the foundationalist presupposes a very tight connection between one's *state* of being justified in believing a proposition and one's *ability* to express one's justifications for it. Following Rescorla, say one *vindicates* an asserted proposition *p* iff one meets all of her interlocutor's legitimate challenges against *p*, all legitimate challenges raised against the propositions one asserts to meet the challenges against *p*, all legitimate challenges raised against

the propositions one asserts to meet these further challenges, so on and so forth. He argues that the inference goes through only if one accepts something along the lines of:

VINDICATION THESIS: One is justified in believing *p* only if one can, at least in principle, vindicate *p*.[110]

The egalitarian's way out, then, is to deny this. Without something like the VINDICATION THESIS, the impossibility of vindicating our claims does not entail that none of our beliefs are justified. And Rescorla provides good reasons to reject the thesis. Here is a quick summary of the reasons he provides (2009b, pp.49-50): we are seemingly justified to believe a proposition we learn through testimony even when we cannot remember exactly when we learned it; we are seemingly justified to hold beliefs based on perception (e.g., <I am raising my arm>) even though it is unclear how the assertions of these beliefs can be defended; very young children and non-linguistic animals seem to be able to hold justified beliefs even though they are unable to perform a defense. On these grounds, he argues that the dialectical egalitarian has a plausible way to block the inference from (9) to (10).

I think Rescorla's criticism successfully shows that dialectical egalitarianism is a viable option on the table. However, there are concerns about the overall view he uses it to defend. An interesting feature of Rescorla's position is that, while he rejects the inference from (9) to (10), he

---

[110] The qualification 'at least in principle' is only meant to rule out cases where one is unable to perform a defense due to dialectically and epistemically irrelevant factors such as exhaustion, inebriation, etc.

accepts (9) as a consequence of his view.[111] Accordingly, he takes the main task of his defense to consist in showing that, at least from the egalitarian perspective, (9) is innocuous. But this doesn't seem right. Rescorla is right to point out that, without the VINDICATION THESIS, (9) does not entail skepticism. However, skepticism is not the only epistemological problem the egalitarian needs to worry about. In Chapter 3, I argued that the act of debating is our main philosophical method for discovering truths. To this end, I argued for:

> TRUTH COMPELS: In principle, theories that are undefeated in every debate in which they are involved are substantially more likely to be true.

And although (9) does not contradict TRUTH COMPELS, it renders it vacuous. In Section 4.1, I argued that one is effectively defeated in a debate in case one fails to properly defend those propositions that are central to one's position. (This was premise (2) of the argument.) To properly defend a proposition, one arguably needs to vindicate it in the sense that we previously introduced. But (9) seems to entail that, in principle, it is impossible for our claims to be properly vindicated: to secure your failure, all your opponent needs to do is brutely challenge every claim that you make. And since this "ingenious" strategy is available to everyone, it would follow that, in

---

[111] In Rescorla (2009b, p.47), he takes (9) to be *entailed* by dialectical egalitarianism. But this can be misleading. His actual motivation for (9) does not come from dialectical egalitarianism *per se*, but instead involves his aversion of the other possible responses to the dialectical regress. The result of one's dialectical interaction with a persistent interlocutor must be one among the following: (a) one vindicates one's initial claim by providing an infinite chain of arguments; (b) one vindicates one's initial claim by providing a circular argument; (c) one vindicates one's initial claim by asserting a dialectically basic proposition; (d) one fails to vindicate one's claim. (c) is obviously incompatible with dialectical egalitarianism; he takes (a) to be impossible due to our mortal limitations; he takes (b) to be simply begging the question. And it is this process of elimination that he leads him to endorse (d), which is tantamount to accepting (9).

principle, it is impossible for anyone to properly win in a debate. This means that the constitutive goal of debating can never be achieved, which makes the act of debating mostly pointless.

The upshot is that (9) threatens the epistemic value of exchanging arguments. And relinquishing this value is no more tolerable than that of skepticism. Rescorla fails to notice this implication of (9) mainly because he takes his criticism of the VINDICATION THESIS to imply more than it actually does. In several places, Rescorla seems to suggest that the rejection of the VINDICATION THESIS encapsulates the view that the state of *holding a justified belief* is completely independent of the act of *defending a claim*; so much so that our dialectical interactions have no epistemological consequences whatsoever. Call this view 'anti-vindicationism'. His reasoning from the negation of the VINDICATION THESIS to anti-vindicationism is most visible in the following passage. After reciting the previously mentioned reasons to reject the VINDICATION THESIS, he writes (2009c, p.94):[112,113]

---

[112] The original passage addresses what he calls 'the justification thesis' which says: "Epistemic justification intimately involves the ability, at least in principle, to defend one's beliefs against legitimate challenges." It is clear that the justification thesis is simply a rough formulation of the VINDICATION THESIS.

[113] Here is some more textual evidence. In the course of pointing out that the foundationalist's argument from skepticism presupposes the VINDICATION THESIS, he comments that "[t]he impossibility of answering iterated brute challenges *shows nothing about epistemic justification* unless we assume the Vindication Thesis" (Rescorla 2009b, p.48, my emphasis). Also, after reciting his reasons for rejecting the thesis, he writes (Rescorla 2009b, p.50):

> All too often, the debate between dialectical foundationalism and egalitarianism is conflated with some epistemological debate. But dialectical foundationalism and egalitarianism are rival views about the structure of reasoned discourse. They entail epistemological claims only when conjoined with additional doctrines relating reasoned discourse to epistemic status. One can reject most such doctrines, so one can coherently combine either dialectical foundationalism or egalitarianism with virtually any conception of epistemic justification.

Here, he is primarily complaining about how egalitarianism is often associated with skepticism on the basis on the VINDICATION THESIS. By expanding on this point, he is also complaining how dialectical foundationalism is often associated with epistemic foundationalism on similar grounds.

The dispute between dialectical foundationalism and egalitarianism is not epistemological. It is not about whether we are justified in believing certain propositions. […] Unfortunately, this issue has been obscured by a persistent tendency to combine or conflate egalitarianism with various controversial epistemological perspectives. […] One need not accept any such epistemological perspective just because one accepts egalitarianism. Epistemological consequences follow from egalitarianism only if we assume one or another dubious doctrine, such as [the VINDICATION THESIS], linking epistemic status and status within reasoned discourse.

And because he thinks he establishes anti-vindicationism, he does not explore whether there are any epistemological worries other than skepticism that stem from (9).

However, anti-vindication does not follow from the negation of the VINDICATION THESIS. It is one thing to sharply distinguish epistemic status from status within reasoned discourse; it is another thing entirely to deny that there are any interesting connections between the two. And one may reject the VINDICATION THESIS while upholding a weaker connection between one's status in the two domains. Rejecting the VINDICATION THESIS amounts to rejecting that the vindication of *p* is not *necessary* for our being justified in believing *p*. But this is compatible with vindication being *sufficient* for justification, or at least our getting closer to it. That is, it seems plausible to maintain that:

Push-Pull: For any two incompatible propositions *p* and *q*, if one can vindicate *p* but not *q*, then it is epistemically better to believe *p* rather than *q*.

The general idea here is that matters of vindicability push us away from certain propositions and pull us toward certain others; hence the name. And this reflects the simple idea that we are capable of improving our epistemic status by believing the propositions we are capable of vindicating. As it was for TRUTH COMPELS, (9) undermines PUSH-PULL not by contradicting it, but by making it vacuous; the existence of such a simple yet effective means to undo our vindicative efforts would make it impossible for us to improve our epistemic status in this fashion. And this gives us a reason to resist (9) regardless of what we say about the VINDICATION THESIS.

Rescorla might object that I am reading too much into (9). For instance, Rescorla emphasizes that (9) only pertains to our vindicative efforts against a *persistent* interlocutor. And since not all interlocutors are persistent like this, he reassures us that it is both "possible and routine" for us to vindicate our assertions against non-persistent interlocutors on the basis of our shared background beliefs (Rescorla 2009b, p.54).[114] Based on this, he may argue that the value of exchanging arguments can be secured by the fact that we are well able to vindicate our claims against non-persistent interlocutors. However, while the solution does sound plausible, it appears to clash with the spirit of egalitarianism. The problem I raise is not so much about whether we are ever able to meet all of an interlocutor's challenges. Rather, it is about whether our vindicative

---

[114] These remarks are originally made in response to the objection that (9) entails that it is impossible for us to vindicate our assertions in the context of any debate. Because Rescorla takes his criticism of the VINDICATION THESIS to have dispelled all epistemological worries against (9), he does not consider the epistemological implications I am pointing to. Instead, he treats the issue as a general absurdity that one may associate with (9).

efforts allow us to retain our assertions. Say an assertion is *globally vindicable* iff, in principle, it can be vindicated against all possible interlocutors. In contrast, say that an assertion is *reasonably vindicable* iff, in principle, it can be vindicated against all non-persistent interlocutors. In its core, the proposed solution is that, although no assertion is globally vindicable (since no assertion can be vindicated against a persistent interlocutor), we are entitled to retain our assertions that are reasonably vindicable. However, it is hard to see how the dialectical egalitarian can pursue this line of argument. Regardless of whether you are able to vindicate an assertion of p against one group of interlocutors, if there is a different set of interlocutors who are armed with legitimate challenges that you cannot meet, then you are arguably obliged to retract p. In Section 4.1, I argued for an expanded retraction norm, namely:

(1) For any proposition p one asserts, if one is unable to defend p against a possible legitimate challenge, then one must retract p.

Since the inability to defend p entails the inability to vindicate p, (1) entails:

(1*) For any proposition p one asserts, if one is unable to vindicate p against a possible interlocutor, then one must retract p.

According to the egalitarian, the persistent interlocutor's brute challenges are indeed legitimate. Thus, (1*) and (9) jointly entail that we are obliged to retract all of our assertions. By rejecting the VINDICATION THESIS, there is room to maintain that we are entitled to retain our *beliefs*. But unless we reject (9), it seems that we are never entitled to keep our assertions of them. Here is a different way to put the point. For the dialectical egalitarian to pursue this line of argument, she needs to have a principled reason to ignore conversational contexts that involve a persistent interlocutor. The *dialectical foundationalist* has such a reason. For her, the persistent interlocutor

violates the norms of debating by brutely challenging dialectically basic claims. Thus, our inability to convince her does not oblige us to retract our claims. But there seems to be nothing that can do this for the dialectical egalitarian. Since she takes all challenges to be legitimate, she has no non-arbitrary means to distinguish the challenges that are raised by each type of interlocutor. That is, from the egalitarian perspective, there should be no significant difference between a claim that is reasonably but not globally vindicated and a claim that fails at both. From her perspective, both involve the failure to meet a legitimate challenge that obliges us to retract the relevant assertions.

In conclusion, it is best for the dialectical egalitarian to block the inference from (8) to (9). To sharpen the point, let us distinguish two ways in which a regress can be said to be vicious. On the hand, say a regress R of task T is *procedurally vicious* iff the fact that R is an infinite regress entails the impossibility of T. On the other hand, say a regress is *philosophically vicious* iff the fact that R is procedurally vicious has a serious consequence that we have philosophical reasons to resist. Rescorla's approach was to concede that the dialectical regress is procedurally vicious while denying that it is philosophically so. What I hope to have shown is that this approach is untenable. Thus, I argue that the dialectical egalitarian's only way forward is to somehow show that the dialectical regress is procedurally benign.[115] For reasons of space, I do not explore whether or how this can be achieved. For my own purposes, it suffices to say that this is the only remaining option for the dialectical egalitarian and that this version of the view is compatible with my overall argument of this chapter.

---

[115] The resulting view would be the dialectical analogue of Peter Klein's (1999) infinitism about epistemic justification.

4.3.3 The Foundationalist's Challenge

Let us turn to whether premise (4) is objectionable on the basis of dialectical foundationalism. My claim is that even the dialectical foundationalist should not accept P-claims as primitive, which strikes me as the more commonly held view. For the most part, even when we do not have an analysis we can readily deploy, it is more typical to think that the majority of our notions (or at least the ones that are often relevant in metaphysics) are analyzable in some way or other.[116] Below, I consider two ways in which this attitude can be resisted and respond to each of them.

Approach 1: P-claims are dialectically basic because primitive concepts are fundamental bits of ideology.

Problem: The approach conflates the need for a *defense* with the need for an *explanation*. A defense of *p* is ultimately a matter of providing *reasons to believe* that *p* is true; an explanation of *p* is a matter of providing *reasons why p* is true. And these are two different things. It is easy to conflate the two because the explanation of *p* can often be used to defend *p*. Let *p* be a proposition that is brutely true and let *q* be the proposition <*p* is brutely true>. Arguably, *q* provides a reason to believe *p*: if one accepts that <*p* is brutely true>, then one thereby has a reason to believe that *p* is true. But *q* doesn't explain *p*; since we've stipulated that *p* is brutely true, it is not explained by anything. Accordingly, just because you are not required to explain something does not mean that you are not required to defend it. To say that a primitive concept is fundamental is to say that there

---

[116] Some may think that the declining popularity of conceptual analyses suggests otherwise. But the view in question is not that the traditional type of conceptual analysis can be given to the majority of our notions, but instead that most notions are excluded from primitive ideology. And this point is compatible with adopting other means of analyses, e.g., use-condition analyses, Sider's metaphysical semantics approach, conceptual engineering, etc.

is nothing that explains our employment of it. That is, if I claim that the notion of *possibility* is primitive, I am thereby claiming that there is nothing that explains why the possible states of affairs are possible. In this case, I am alleging that my employment of this notion represents the world at its most fundamental level. Thus, it becomes unreasonable to demand that I explain why these states of affairs have the modal status that they have. But there is no reason to think that this also makes it unnecessary to defend this claim. To say otherwise is to maintain that we have some special epistemic access to the world at its fundamental level that entitles us to make claims about it without a defense. And there is no reason to think that this is the case.

Approach 2: Another approach is to restrict the claim to a special set of "basic" concepts and claim that can be presumptively accepted as primitive.[117] The details of this objection depend on what it is for a concept to be "basic." But the general idea is that there is a special philosophical intuition that certain concepts are unanalyzable.

Problem: I am doubtful that any concepts are basic in any meaningful way. To my mind, it is utterly implausible that we have direct intuitions about whether a concept is *primitive*. And neither does any part of our comprehension of a concept seem to shed any light on its analyzability: I expect that nobody at any point in their philosophical career had ever encountered a concept and immediately *felt it in their guts* that the concept is unanalyzable. There may be cases in which it is genuinely difficult to imagine how an analysis might go. But this hardly implies anything significant about dialectical basicness. At best, it makes room for an inductive argument that a given notion is primitive. But as I discuss later, this is compatible with the view I am arguing for.

---

[117] This approach was suggested to me by Justin Dealy.

This leads me to believe that the approach mistakes some other intuition for the alleged intuition for the basicness of concepts. One plausible culprit may involve intuitions about comparative fundamentality. Despite my previous remarks about the absolutely fundamental, intuitions about what explains what do seem to give us some indication of whether a given concept is comparatively fundamental. For instance, the notion of *greenness* intuitively concerns more fundamental phenomena than *grueness*; more drastically, the notion of *property instantiation* concerns matters that are much more fundamental than the notion of a *product warranty*. It is likely that these judgments are rooted in the fact that we expect *grueness* to be explained in terms of an object's being green under certain conditions and not being green under others, and similarly for *product warranty* and *property instantiation*. By reflecting on a wide variety of concepts in this fashion, we may even be able to pick out a handful of concepts that are *reasonable candidates* for being primitive in the sense that we would not be alarmed to see a theory that takes such a notion to be primitive. However, this is as far as we can get with comparative fundamentality. To make an actual choice within this pool of candidates, we need reasons to believe that a given notion concerns absolutely fundamental matters. To have such a reason, we need a further reason to believe that the matters in question cannot be explained by anything else: we cannot simply intuit that we have reached rock-bottom. And this brings us back to my diagnosis of the primitivist's dialectic: apart from demonstrating that all possible explanations fail, there appears to be no reason that we can possibly have.

4.4 The Primitivist's Dialectic

4.4.1 Defending Primitiveness

In this section, I further motivate premise (6) and spell out how this puts the primitivist at a dialectical disadvantage. Premise (6) states that there is virtually no positive argument one can give in defense of a P-claim. By 'positive argument', I mean an argument for a position that does not rest on the failure of its rivals. I say "virtually" mostly for the sake of humility and to make room for outlying cases. But in most cases, there is seldom anything one can provide as direct *evidence* for primitivism. And the only way to defend primitivism seems to be through the process of elimination: by showing that no reductive procedure for the notion in question is successful. This constitutes a general dialectical disadvantage for the primitivist. Since it is virtually impossible to perform such a defense in our lifetime, it would seem that we are never able to properly defend primitivism of a given concept as opposed to remaining agnostic about its reducibility. Furthermore, it makes the case for primitivism entirely reliant on the failure of reductionism. Because of this, the mere prospect of a reduction is enough to put significant pressure on the primitivist.

The argument here is not intended to suggest that we should refrain from taking *any* notion as primitive. This would not be a viable means of theorizing. Instead, I am arguing that the theoretical costs that are often associated with primitivism are at least partially dialectical in nature. And like any other theoretical cost, there may be various ways in which the cost in question can be balanced out or be mitigated in some other way.[118] Also, my argument allows there to be cases

---

[118] That being said, it seems likely that an argument for primitivism that is based on a cost-benefit analysis would collapse into an argument that shoots down reductive proposals. Consider three of the main theoretical virtues that are

where we have inductive reasons to actively doubt the prospects of a reduction. For instance, if there has been a sufficiently large number of failed reductive proposals, it may be justified to believe that no such reduction is possible. In other cases, it may simply be difficult for us to imagine how an analysis could even go. That being said, I do take my argument to put pressure on various forms of primitivism that are found in the literature. I argue that, for most of the notions that are actively debated in metaphysics (and philosophy more broadly), there are not enough inductive grounds that cast doubt on their reducibility. Consider our leading examples of causality, temporality, and modality. Although one may have reasons to doubt the specific reductive theories that have been offered in these areas, the consensus is hardly that we should stop pursuing a reductive theory altogether.[119] In this light, it is fair to say that those who accept primitivism for *these* notions are situated in an uphill battle.[120]

Thus is my diagnosis of the primitivist's dialectic. These considerations give us simple and easy explanations for LAST RESORT and THE UNSPOKEN RULE. Let us start with LAST RESORT. The attitude is that primitivism should be accepted only when all other accounts fail. By (6), the process of elimination is the only way a primitivist can justify her P-claim. By (5), her

---

often relevant to a cost-benefit analysis: (i) simplicity, (ii) unification, and (iii) explanatory power. In most cases, the first two are a losing battle for the primitivist, as it would seem that reductive theories usually lead to greater simplicity and unification than primitivism. Thus, a cost-benefit analysis would be in the primitivist's favor only if primitivism leads to greater explanatory power. But for this to be the case, there needs to be certain phenomena that the reductionist cannot explain very well, which would seem to constitute a reason to believe that the reductive procedure fails to go through.

[119] Ned Hall and L. A. Paul (2013) address such a line of inductive reasoning for the case of causation.

[120] How much does this actually affect debates that are currently held? This depends. As far as actual members of the philosophical community are concerned, I doubt that these considerations will persuade anyone to switch camps. Part of this is because, as I see it, the majority of philosophers who identify as non-reductionists do so not because they actively embrace primitivism, but instead because they are unsatisfied with the reductive theories that have been proposed thus far. If so, then 'primitivism' can be understood as a general banner that unites philosophers who are skeptical of the reductive ideas that are currently and/or have been historically prominent. In this light, a particular philosopher can easily avoid criticism based on my discussion by rebranding oneself as an opponent of specific reductive approaches rather than an advocate of primitivism. However, my discussion still has the merit of clarifying how we should conceptualize philosophical discussions of various sorts and of directing our attention to how dialectical contexts can and should influence the theories we adopt.

inability to go through the process of elimination results in her being defeated in the debate. But obviously, you cannot go through the process of elimination while conceding that the opposing position is not eliminable. Therefore, it follows that the primitivist cannot concede to the reductive procedure without forfeiting the debate. This means that the tenability of primitivism relies on the untenability of reductionism. Thus, it is unwarranted to reach for a primitivist account unless the prospects of a successful reduction are substantially discredited. The reasoning so far also sets the stage for an explanation of UNSPOKEN RULE. The rule is that, if something *can* be reduced, then it *should* be reduced. On the face of it, this appears to make a stronger claim than LAST RESORT. As we see above, LAST RESORT is a methodological attitude that concerns what position should be favored at what stage of the dialectic. In contrast, UNSPOKEN RULE appears to make a stronger claim about when reductionism should be *accepted*. This stronger claim can be justified by combining LAST RESORT with TRUTH COMPELS. If the primitivist concedes that a reduction is possible, then she thereby forfeits the debate. By TRUTH COMPELS, this implies that primitivism is untenable, or to the least, that it has a substantially decreased likelihood of being true. Thus, if one concedes that something *can* be reduced, it is no longer reasonable for one to accept primitivism, in which case, reductionism remains as the only viable option.

These explanations for LAST RESORT and UNSPOKEN RULE have some interesting implications. When thinking about the matter, it is easy to think that these attitudes are rooted in some independent feature of primitivism and reductionism. For instance, when asked in a casual setting, many of my colleagues suggested that these attitudes had something to do with the greater simplicity that can be achieved by reductive theories, or something of that nature. Although I was doubtful of a simplicity-based account, my initial hypothesis was broadly along similar lines; I too thought that there must be something "good" about reduction such that the mere possibility of a

reduction is enough to pull us in. But the explanations I arrived at turn this idea on its head. On my account, LAST RESPORT and UNSPOKEN RULE are less about the internal characteristics of these theories but are instead more about where the natural twists and turns in the dialectical lead us. In this light, the initial hypothesis mistook the explanans for the explananda. Initially, the goal was to explain these attitudes by identifying some independent good-making feature of reductionism. But my analysis reveals that this gets things backwards and that it is the dialectical forces expressed by these attitudes that makes reductionism attractive.

4.4.2 Objections

I close the chapter by considering some objections. One may resist the move from (6) to (7) by arguing that defending a reducibility-claim is no easier than defending a P-claim. For instance, one may think that, insofar as the reductionist faces the same task of refuting the other reductive proposals on the table, she is in the exact same dialectical situation as the primitivist.[121] The case is more pressing when we consider the fact that there could be more than one reductive procedure that goes through. A prime example is the set-theoretical reduction of natural numbers: the fact that the natural numbers seem reducible to both van Neumann ordinals *and* Zermelo ordinals appear to count as a reason to reject that the reduction is successful in either case.[122] In this light, it would seem that the success of a reduction does not solely rely on extensional adequacy,

---

[121] I thank Justin Dealy for suggesting this objection.
[122] Simply put, the problem is that the two "reductions" entail different extra-arithmetical truths. E.g., the statement "2 is a member of 3." is true if we employ von Neumann ordinals but not if we employ Zermelo ordinals. See Benacerraf (1965).

but also requires that the proposed procedure is uniquely successful. Hence, a proper defense of a reducibility-claim would need to demonstrate that all other reductive attempts are unsuccessful.

This objection can be met by observing that the prospects of a given reductive procedure itself act as evidence for its uniqueness. Successful reductive procedures are not a dime a dozen and subject matters that admit multiple reductive procedures are even more scarce. Because of this, when one is close to a successful procedure, this itself gives one a reason to believe that the procedure is uniquely successful. In this fashion, the reductionist is generally fighting against better odds than the primitivist.

Another possible objection is that, contrary to (6), a positive case for primitivism *can* be made by appealing to the theoretical work the notion can do. This objection can be bolstered by the fact that, when a notion is suggested to be primitive, it is often accompanied by an appeal of this sort. For example, David Lewis famously advocated a notion of primitive naturalness by appealing to the "new work" that can be carried out by a theory of sparse, natural properties (Lewis 1983; 1986). More recently, various grounding theorists advocate a primitive notion of grounding by appealing to the work that modal notions like supervenience or necessitation are incapable of performing.[123]

However, it does not take long to notice that the appeal to theoretical work only goes so far. For both of the above cases, the appeal to theoretical work only urges us to *recognize* the notions in question. And the P-claims for each notion rely on the failure of the rival, reductive views. For naturalness, Lewis famously expresses his indecision for *how* to account for naturalness

---

[123] The stock example is that modal notions fail to track more discriminating explanatory relations that hold for necessary truths; while the mathematic truth '2+2=4' does not obtain by virtue of, say, the fact that Stevie Wonder is a great composer, a modal analysis of grounding is unable to exclude such cases.

(e.g., primitivism about naturalness, an analysis through Armstrongian universals, and an analysis through tropes). This clearly suggests that he thinks the appeal to theoretical work offers no special support for primitivism.[124] For grounding, there is an unsettled debate between those who accept primitive grounding and those who offer an analysis in terms of supervenience and absolute fundamentality.[125] And if the analysis is successful, then the theoretical work that is envisioned for grounding gives no support to primitivism about grounding: if the reduction is successful, any work assigned to it can be carried out by the reduction base as well. Thus, the theoretical work would hardly serve as a distinguishing factor between primitivism and reductionism.

---

[124] After quickly going through some unsuccessful analyses, he writes (Lewis 1986, p.63) "Unless we are prepared to forgo some of the uses of the distinction between natural and unnatural properties, we shall have no easy way to define it without circularity. That is no reason to reject the distinction. Rather that is a reason to accept it—as primitive, if need be." But immediately after, he discusses analyses with better prospects (e.g., an analysis in terms of Armstrongian universals and one in terms of tropes) and states to be undecided among the three options.

[125] Primitivism about grounding is the standard view, more or less. But see Phil Bricker (2006) for a suggested partial analysis in terms of supervenience and absolute fundamentality.

BIBLIOGRAPHY

Armstrong, David (1978), *Nominalism and Realism: Universals and Scientific Realism* (Volume 2), Cambridge University Press

Audi, Paul (2012) "Grounding: Toward a Theory of the In-Virtue-Of Relation", *Journal of Philosophy* 109 (12) 685-711

Baker, Alan (2003) "Quantitative Parsimony and Explanatory Power", *British Journal for the Philosophy of Science* 54 (2):245-250

Barnes, Elizabeth (2014) "Going Beyond the Fundamental: Feminism in Contemporary Metaphysics", *Proceedings of the Aristotelian Society* 114(3pt3):335-351

Baron, Sam, & Tallant, Jonathan, (2018) "Do not Revise Ockham's Razor Without Necessity", *Philosophy and Phenomenological Research*, Vol. XCVI No. 3

Bradley, Darren (2017) "Philosophers Should Prefer Simpler Theories", *Philosophical Studies* 175 (12):3049-3067

Brenner, Andrew (2016) "Simplicity as a Criterion for Theory Choice in Metaphysics", *Philosophical Studies* 174 (11):2687-2707

Bealer, George, (2004) "An Inconsistency in Direct Reference Theory", *Journal of Philosophy* 101 (11):574-593

Benacerraf, Paul (1965) "What numbers could not be", *The Philosophical Review*, *74*(1): 47-73

Bliss, Ricki, (2013) "Viciousness and the Structure of Reality", *Philosophical Studies*, 166: 399–418.

Bliss, Ricki, and Trogdon, Kelly, (2021) "Metaphysical Grounding", The Stanford Encyclopedia of Philosophy (Winter 2016 Edition), Edward N. Zalta (ed.), URL = plato.stanford.edu/archives/win2021/entries/grounding

Brandom, Robert, (1994) *Making It Explicit*, Harvard University Press

Bricker, Philip (2006) "The Relation Between the General and the Particular: Entailment vs. Supervenience" in Dean Zimmerman (ed.), *Oxford Studies in Metaphysics Volume 2*, Oxford University Press

--. (2016) "Ontological Commitment" *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = plato.stanford.edu/archives/win2016/entries/ontological-commitment

--. (2018) "A Sketch of Reality" in *Modal Matters: Essays in Metaphysics*, Oxford University Press

Burgess (2005) "Being Explained Away", *The Harvard Review of Philosophy* 13 (2):41-56

Cameron, Ross, (2015) *The Moving Spotlight: An Essay on Time and Ontology*, Oxford University Press

Case, Spencer, (2019) "The Normative Error Theorist Cannot Avoid Self-defeat", *Australasian Journal of Philosophy*, DOI: 10.1080/00048402.2019.1582079

Correia, Fabrice & Schnieder, Benjamin (2012) "Grounding: an opinionated introduction" in Correia, Fabrice. and Schnieder, Benjamin (eds) *Metaphysical Grounding: Understanding the Structure of Reality*, Cambridge University Press

Cowling, Sam (2020) "Ideology and Ontology", in *Routledge Handbook of Metametaphysics*, Routledge

Cuneo, Terance, (2007) "Epistemic Nihilism" in *The Normative Web: An Argument for Moral Realism*, New York: Oxford University Press

Curtis, Benjamin (2014) "The Rumble in the Bundle", *Nous*, 48(2):298-313

Da Vee, Dean (2020), "Why Ockham's Razor Should be Preferred to the Laser," *Philosophical Studies*, 177:3679-3694

Daly, Chris, (2010) *An Introduction to Philosophical Methods*, Broadview Press

Dasgupta, Shamik, (2009) "Individuals: an essay in revisionary metaphysics", *Philosophical Studies*, 145 (1): 35-67

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

Diehl, Catherine, (2018) "A Language for Ontological Nihilism" *Ergo* 5:971-996

Divers, John, (2002) *Possible Worlds*, Routledge

Donaldson, Thomas, (2015) "Reading the book of the world", *Philosophical Studies* 172 (4):1051-1077

Dorr, Cian, (2016) "To be F is to be G" *Philosophical Perspectives* 30 (1):39-134

Fiddman, Mark & Rodriguez-Pereyra, Gonzolo, (2018) "The Razor and the Laser," *Analytic Philosophy*, 59(3): 341-358

Fine, Kit, (2012) "Guide to Ground" in Correia, Fabrice. and Schnieder, Benjamin (eds) *Metaphysical Grounding: Understanding the Structure of Reality*, Cambridge University Press,

Finocchiaro, Peter (2019a) "The Explosion of Being: Ideological Kinds in Theory Choice", *Philosophical Quarterly* 69 (276):486-510

--. (2019b) "Ideology and its Role in Metaphysics", *Synthese* 198 (2):957-983

Gibbard, Allan, (2003) *Thinking How to Live*, Cambridge: Harvard University Press

Griffith, Aaron (2017) "Social Construction and Grounding", *Philosophy and Phenomenological Research* doi:10.1111/phpr.12376

Hall, Ned and Paul, L. A. (2013) "Metaphysically Reductive Causation", *Erkenntnis* 78:9-41

Haslanger, Sally (2005) "What Are We Talking About?: The Semantics and Politics of Social Kinds", *Hypatia* 20 (4):10-26

Hofweber, Thomas (2005) "Supervenience and Object-Dependent Properties" *Journal of Philosophy* 102 (1):5-32

Huemer, Michael (2009) "When is Parsimony a Virtue", *Philosophical Quarterly* 59 (235):216-236

Jenkins, Carrie, (2011), "Is Metaphysical Grounding Irreflexive?", *The Monist*, 94: 267–76.

Jenkins, Katherine (2016) "Amelioration and Inclusion: Gender Identity and the Concept of *Woman*", *Ethics*, 126: 394–421

Klein, Peter (1999) "Human Knowledge and the Infinite Regress of Reasons", *Philosophical Perspectives,* 13:297-325

Kornblith, Hilary, (2002) *Knowledge and Its Place in Nature*, Oxford University Press

--. (2011) "Why Should We Care About the Concept Knowledge?", *Episteme* 8(1):38-52

Krämer, Stephan (2010), "How Not to Defend an Ontological Cheat," *Analysis* 70 (2)"290-296

Lee, Byeongduk. (2001) "Moore's Paradox and Self-ascribed Belief", *Erkenntnis* 55: 359-370

Leite, Adam (2004) "On Justifying and Being Justified", *Philosophical Issues* 14 (1):219-253

--. (2005) "A Localist Solution to the Regress of Epistemic Justification", *Australasian Journal of Philosophy* 83(3):395-421

Lewis, David (1973) "Causation", *Journal of Philosophy* 70 (17):556-567

--. (1983) "New Work for a Theory of Universals", *Australasian Journal of Philosophy* 61 (4):343-377

--. (1986) *On the Plurality of Worlds*, Malden, Mass.: Wiley-Blackwell

--. (1994) "Reduction of Mind" Samuel Guttenplan (ed.), *Companion to the Philosophy of Mind*. Blackwell. pp. 412-431

--. (2000) "Causation as Influence", *Journal of Philosophy* 97 (4):182-197

Lutz, Matt, (2014) "The 'now what' problem for moral error theory", *Philosophical Studies* 171, 351-371

Markosian, Ned, (1998) "Brutal Composition", *Philosophical Studies* 92 (3):211-249

Margolis, Eric & Laurence, Stephan (2007) "The Ontology of Concepts", *Noûs* 41 (4):561-593

Maurin, Anna-Sofia (2012), "Bradley's Regress" *Philosophy Compass* 7 (11):794-807

Mikkola, Mari (2015) "Doing Ontology and Doing Justice: What Feminist Philosophy Can Teach Us About Meta-Metaphysics," *Inquiry*, 58 (7-8): 780-805

Morris, Jeremy, (2008) "Pragmatic Reflexivity in Self-Defeating and Self-Justifying Expressions", *Argumentation* 22, 205-216

Nelson, Michael, (2008) "Frege and the Paradox of Analysis" *Philosophical Studies* 137 (2):159-181

Nolan, Daniel (1997) "Quantitative Parsimony", *British Journal for the Philosophy of Science* 48 (3):329-343

--. (2014) "The Dangers of Pragmatic Virtues", *Inquiry: An Interdisciplinary Journal of Philosophy* 57 (5-6):623-644

Oliver, Alex, (1996) "The Metaphysics of Properties", *Mind* 105 (417):1-80

Olsen, Eric, (2007) "Nihilism" in *What Are We?*, New York: Oxford University Press

Parfit, Derek, (2011) *On What Matters Vol.2*, New York: Oxford University Press

Perovic, Katarina, (2017) "Bradley's Regress", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), plato.stanford.edu/archives/win2017/entries/bradley-regress

Plantinga, Alan, (1974) *The Nature of Necessity*, Oxford: Clarendon

Quine, Willard Van, (1951) "Ontology and Ideology," *Philosophical studies* 1:11-15

--. (1960) "Variables Explained Away", *Journal of Symbolic Logic* 32 (1):112-112

--. (1983) "Ontology and Ideology Revisited", *The Journal of Philosophy* 80 (9): 499-502

Rescorla, Michael, (2009a) "Assertion and Its Constitutive Norms" *Philosophy and Phenomenological Research*, 79 (1): 98-130

--. (2009b) "Epistemic and Dialectical Regress", *Australasian Journal of Philosophy* 87 (1):43-60

--. (2009c) "Shifting the Burden of Proof", *Philosophical Quarterly* 59 (234):86-109

Roche, William, (2012) "A Weaker Condition for Transitivity in Probabilistic Support", *European Journal for the Philosophy of Science* 2:111–118

Rubio, Daniel (2022) "Ideological Innocence", *Synthese* 200 (5):1-22

Schaffer, Jonathan (2009) "On What Grounds What", in Manley, David, Chalmers, David & Wasserman, Ryan (eds), *Metametaphysics: New Essays on the Foundations of Ontology*, Oxford University Press

—. (2012) "Grounding, Transitivity, and Contrastivity" In Correia, Fabrice & Schnieder, Benjamin (eds.), *Metaphysical Grounding: Understanding the Structure of Reality*, Cambridge University Press

--. (2010) "Monism: Priority of the Whole" *Philosophical Review* 119 (1):31-76

--. (2015) "What Not to Multiply Without Necessity" *Australasian Journal of Philosophy* 93 (4):644-664.

—. (2016) "Ground Rules: Lessons from Wilson" Aizawa, Ken & Gillett, Carl (eds.), *Scientific Composition and Metaphysical Ground*, Palgrave-Macmillan

Shalkowski, Scott, (1994) "The Ontological Ground of Modality", *Philosophical Review* 103 (4):669-688

Sider, Ted, (2003) "Reductive Theories of Modality", Loux, Michael & Zimmerman, Dean (eds.), *The Oxford Handbook of Metaphysics*, Oxford University Press

--. (2011) *Writing the Book of the World*, Oxford University Press

--. (2013) "Against Parthood", *Oxford Studies in Metaphysics* 8:237-293

--. (2017) "Substantivity in Feminist Metaphysics" *Philosophical Studies*, 174:2467–2478

Shah, Nishi. (2010) "The Limits of Normative Detachment", *Proceedings of the Aristotelian Society* 110 (3pt3):347-371

Shogenji, Tomo, (2003) "The Transitivity of Probabilistic Support", *British Journal for the Philosophy of Science* 54: 613-616.

Skow, Brad, (2015) *Objective Becoming*, Oxford University Press

Sober, Elliot (2015) *Ockham's Razer: A User's Guide*, Cambridge University Press

Stratton-Lake, Philip, (2002), Introduction of *The Right and the Good* by David Ross, Oxford University Press

Streumer, Bart, (2013) "Can We Believe the Error Theory?", *The Journal of Philosophy* 110: 194-212

—. (2018) "Why We Really Cannot Believe the Error Theory" in Machuca (ed) *Moral Skepticism: New Essays*, Routledge

Tallant, Johnathan (2009) "Ontological cheats just might prosper," *Analysis* 69 (3):422-430

Thompson, Naomi, (2016) "Metaphysical interdependence", in Jago Mark (ed.), *Reality Making*, Oxford University Press

Turner, Jason, (2011) "Ontological Nihilism", *Oxford Studies in Metaphysics* 6:3-54

Trogdon, Kelly, (2013) "An introduction to grounding", in Hoeltje, M., Schnieder, B. & Steinberg, A. (eds), *Varieties of Dependence*, Philosophia Verlag

Torrengo, Guilano (2014) "Ostrich presentism", *Philosophical Studies* 170 (2):255-276

Rosen, Gideon, (2009) "Metaphysical Dependence: Grounding and Reduction", in Hale, Bob & Hoffman, Aviv (eds.), *Modality: Metaphysics, Logic, and Epistemology*, Oxford: Oxford University Press

van Fraassen, Bas, (1987) *The Scientific Image*, Oxford University Press

Willard, M. B., (2014) "Against Simplicity" *Philosophical Studies* 167 (1):165-181

Williams, Michael (2000) "Dretske on Epistemic Entitlement", *Philosophy and Phenomenological Research* 60 (3):607-612

--. (2004) "Is Knowledge a Natural Phenomenon?", in Richard Schantz (ed.), *The Externalist Challenge*, De Gruyter

--. (2015) "What's So Special About Human Knowledge?", *Episteme* 12 (2):249-268

Wilson, Jessica (2014) "No Work for a Theory of Grounding", *Inquiry*, 57(5–6): 1–45.

—. (2016) "Grounding-Based Formulations of Physicalism", *Topoi* 37 (3):495-512

Wright, Crispin, (2003) "Truth in Ethics" in *Saving the Differences*, Cambridge: Harvard University Press