University of California, Santa Barbara

On the metaphysical content of scientific theories

A dissertation submitted in partial satisfaction of the requirements for the degree of
Doctor of Philosophy in Philosophy

Alex S. LeBrun

Committee in charge:

Professor Thomas William Barrett, Co-Chair

Professor Daniel Z Korman, Co-Chair

Professor Kevin Falvey

Professor Jill North, Rutgers University

September 2023

The dissertation of Alex S. LeBrun is approved.

Professor Kevin Falvey

Professor Jill North

Professor Thomas William Barrett, Committee Co-Chair

Professor Daniel Z Korman, Committee Co-Chair

September 2023

On the metaphysical content of scientific theories

Acknowledgements

Chapter 3 overlaps with the following article:

- "What are Empirical Consequences? On Dispensability and Composite Objects"
  Synthese 199: 13201 - 13223 (2021).

I am grateful to *Synthese* for permission to include this material here.

Many people have been indispensable to the completion of this work (and thus deserve some sense of commitment). This work has been made better by the contributions of many colleagues at UCSB in writing groups, the Logic and Philosophy of Science Reading Group, and graduate colloquia. Special thanks to help from Steven Canet, Jon Charry, Morgan Davies, Dylan Goldman, Katherina Gontaryuk, Rick Lamb, and Natalie McCosker. Thanks especially to Jeff Bagwell and Keith Dyck for helping me think clearly through all these issues that we have worked on together. There are many other colleagues whose advice and assistance greatly improved this work: Otávio Bueno, Lenny Clapp, Mark Colyvan, Kevin Falvey, Hartry Field, Dan Hoek, Tom Holden, Trenton Merricks, Jill North, Erica Shumener, Teresa Robertson, and Tuomas Tahko.

My two advisors, Thomas Barrett and Dan Korman, have taught me more about how to be a good thinker, writer, and colleague than everyone else put together. Thomas's support and advice has been truly invaluable, and I am lucky to call him a friend. Without Dan's single question—"Are you saying composite objects are indispensable?"—this dissertation would not exist. I am lucky to have had such outstanding advisors, and everything good in this work is due to their care for my intellectual output.

Every good philosopher needs philosopher friends who know how to stop talking about philosophy. To this need, I thank Tom Costigan, Seán Pierce, and Kyle Dickey. They have both contributed to the philosophical work in this dissertation and to the mental breaks required to complete such a feat.

And, finally, I thank my wife, Clarissa Strayer, whose unbreakable belief in me made the process easy.

VITA OF ALEX S. LeBRUN

September 2023

Education

Doctor of Philosophy in Philosophy, University of California, Santa Barbara, 2023 (expected)

Master of Arts in Philosophy, Northern Illinois University, 2017

Bachelor of Arts in Philosophy, California Polytechnic University, San Luis Obispo, 2014

Professional Employment

2023 - present: Lecturer, California Polytechnic State University, San Luis Obispo

2017 - 2023: Teaching Assistant/Associate, University of California, Santa Barbara

Publications

"What are Empirical Consequences? On Dispensability and Composite Objects" (2021). *Synthese* 199, 13201–13223.

Fields of study

Areas of Specialization: Metaphysics; Philosophy of science

Areas of Competence: Philosophy of Language, Ethics, Biomedical Ethics, History of Analytic Philosophy

Abstract

On the Metaphysical Content of Scientific Theories

by

Alex S. LeBrun

A central concern of metaphysicians of science is uncovering the metaphysical content of our best scientific theories. One method for uncovering this content is with the use of *indispensability* arguments. These arguments infer the existence of some entity or structure on the basis of its presence in the formulation of some theory. While promised as a general argumentative strategy, the literature on indispensability arguments is generally concerned with the question of whether numbers are indispensable to our best scientific theories. In this dissertation, I cash in on the promise that indispensability arguments are generalizable by examining whether composite objects are indispensable to our best scientific theories. The first half of the dissertation (Chapters 2, 3, and 4) examines this question head on. I argue that the extant arguments for the indispensability of composite objects are not convincing, but once we sharpen our definition of indispensability, we see that there are good reasons to think that composites are, indeed, indispensable to our best scientific theories. In the second half, I examine the idea that indispensability is a guide to ontology from the perspective of philosophy of language. In Chapter 5, I present a novel interpretation of Putnam's original indispensability argument, where I argue that he is using the logic of *linguistic presupposition*. In Chapter 6, I try to show that sometimes it is permissible to linguistically subtract one's ontological commitment to some indispensable entity.

Contents

Chapter 1

Introduction

Our best scientific theories often appeal to a variety of entities and structures in the course of their formulations. They will reference theoretical entities like electrons and visible entities like bars of iron. Further, we might think that there is no way to formulate our best theories without referencing, for example, electrons. According to an influential way of thinking, if we were to accept atomic theory but reject the existence of electrons,

> It is like trying to maintain that God does not exist and angels do not exist while maintaining at the very same time that it is an objective fact that God has put an angel in charge of each star and the angels in charge of each of a pair of binary stars were always created at the same time! (Putnam, 1975, p. 74)

Putnam is articulating the intuition behind the indispensability principle:

Indispensability Principle. We ought to ontologically commit to all the entities that are indispensable to our best scientific theories.

According to this principle, it is "intellectual doublethink" (Field, 2016, p. 2) to fail to commit to the entities that are indispensable to the theories one thinks are true. Indispensability here means that we *cannot do without* some entity—that there is no adequate formulation of the theory without appealing to the indispensable thing. The most famous application of this principle has been to try to show that we ought to ontologically commit to numbers, since it seems that they are indispensable to our best theories (Quine, 1956; Putnam, 1975; Colyvan, 2001). Much of the literature has developed responses to the indispensability argument for the existence of numbers. On the one hand, Field (2016) questions whether numbers are indeed indispensable by trying to provide a nominalized formulation of Newtonian Gravitation theory. On the other hand, many have argued that we should reject the indispensability principle as it applies to numbers.[1]

The indispensability principle is meant to justify commitment to all sorts of entities, not just numbers. As Colyvan says, the indispensability argument claims that numbers are indispensable to our best theories in "whatever sense it is in which electrons, neutron stars, and viruses are indispensable to their respective theories" (Colyvan, 2001, p. 12). Accordingly, the indispensability principle can be used to determine ontological commitment for a whole range of traditional metaphysical questions. One of my primary goals in this dissertation is to determine whether composite objects count among the indispensable entities. Chapters 3 and 4 examine this in earnest. If so, then we ought to ontologically commit to them. A secondary goal is to examine the role that philosophy of language plays in determining when something is indispensable. Chapters 5 and 6 take a historical and contemporary look at the connection between language and indispensability. In Chapter 2, I probe the standard definition of 'dispensability', point out some problems with it, and offer an alternative.

[1] See Maddy (1992, 1997); Sober (1993); Yablo (1998, 2005, 2014); Azzouni (2004, 2012); Saatsi (2011, 2016, 2017, 2020).

1.1 The Indispensability Argument for the Existence of Numbers

The indispensability argument for the existence of mathematical objects—of which numbers are one sort—began with works from Quine and Putnam.[2] The argument as it is understood today is as follows:

P1. We ought to ontologically commit to all and only the indispensable entities of our best scientific theories.

P2. Mathematical objects are indispensable to some of our best scientific theories.

P3. So, we ought to ontologically commit to mathematical objects.

The historical accuracy of this argument has been debated, by e.g., Liggins (2008), and in Chapter 5 of this dissertation. Nevertheless, this is the version presented by Colyvan (2001, 2008) and is most commonly discussed.

P1 claims that we ought to ontologically commit to all *and only* the indispensable entities. P1 is justified partially by the indispensability principle, giving us one direction of the premise. But the claim that we ought to ontologically commit to only the indispensable entities of our best theories is both controversial and unnecessary for the validity of the argument. However, as we see in Chapters 2, 3, and 4, it is important to assessing whether composite objects are indispensable to our best scientific theories.

The standard for what counts as indispensable in the course of the argument is generally not given a precise articulation. In common parlance, the term means

[2]See Quine (1948, 1960, 1981) and Putnam (1967a, 1971, 1975).

that which cannot be done without. Accordingly, in the context of scientific theories, an entity's appearance in the formulation of a theory is indispensable when we cannot adequately formulate the theory without that entity. Field is helpful here:

> The utility of theoretical entities lies in two facts:
>
> (a) they play a role in powerful theories from which we can deduce a wide range of phenomena; and
>
> (b) no alternative theories are known or seem at all likely which explain these phenomena without similar entities...
>
> ... [There are dispensable entities for which] we can give attractive reformulations of such theories in which [the dispensable] entities play no role. (Field, 2016, p. 7 - 8)

The idea here is that an entity is dispensable when we can provide an alternative formulation of the theory that does not appeal to the entity, explains the same phenomena, and is suitably attractive. I trace the history of this definition and object to it in Chapter 2.

P2 is on its face mysterious to justify. It is not clear what it would take to show that it is impossible to give an attractive reformulation of some theory that does not appeal to some entity. Putnam (1971, Ch. 5), to his credit, does attempt to give an impossibility argument for a nominalistic (mathematical object-free) reformulation of Newtonian Gravitation theory. But as Barrett (2020b) has argued, is clear that no serious nominalist should be convinced by Putnam's argument. The typical argument for P2 is given just by looking. As Putnam says,

> Newton's law, as everyone knows, asserts that there is a force f_{ab} exerted by any body a on any other body b. The direction of the force f_{ab}

> is towards a, and its magnitude F is given by
>
> $$F = \frac{gM_aM_b}{d^2} \tag{N}$$
>
> where g is a universal constant, M_a is the mass of a, M_b is the mass of b, and d is the distance which separates a and b. (Putnam, 1971, p. 36)

The idea is that the reference to mathematical objects is so central to the content of (N) that we cannot even conceive of a reformulation that does not appeal to mathematical objects. More generally, P2 is justified by claiming that there are (or would be) deficiencies to any nominalistic reformulation of our best scientific theories.

If sound, the argument shows that we ought to ontologically commit to mathematical objects like numbers or sets. The types of responses to the indispensability argument fall in two camps. *Hard roaders* attempt to reject premise 2 by offering reformulations of our best scientific theories that do not appeal to mathematical objects. The *locus classicus* of this is Field (2016), who attempts to reformulate a portion of Newtonian Gravitation theory without quantifying over numbers. *Easy roaders* reject premise 1 by finding some reason to think that the indispensability principle is not true in general or is not true for the case of mathematical objects in particular. See Maddy (1992, 1997); Sober (1993); Yablo (1998, 2005, 2014); Azzouni (2004, 2012); Dorr (2010); Saatsi (2011, 2016, 2017, 2020).

One project of this dissertation is to examine the prospects of an indispensability argument for the existence of composite objects. As this is an indispensability argument, much of the strategies for responding will mirror those of the original indispensability argument.

1.2 The *Enhanced* Indispensability Argument

At the turn of the century, Melia (2000, 2002) and Colyvan (2002) gave arguments that restricted the scope of the indispensability principle. Instead of demanding that we ontologically commit to *all* indispensable entities, most philosophers in the literature now claim that we are only required to commit to all entities that are *explanatorily indispensable*. As Melia says,

> Were there clear examples where the postulation of mathematical objects results in an increase in the same kind of utility as that provided by the postulation of theoretical entities, then it would seem that the same kind of considerations that support the existence of atoms, electrons and space-time equally supports the existence of numbers, functions and sets. (Melia, 2002, pp. 75 - 76)

Here Melia is claiming that it is not enough for the indispensability argument that mathematical objects are indispensable—in the sense of not being able to formulate an alternative without appealing to them. Instead, it must be that mathematical objects are indispensable *in the same way* that atoms, electrons, and space-time are. And this sense, as the literature has developed, has been *explanatory indispensability*. We need, according to this route, an instance where a mathematical fact explains a purely physical fact.

The most discussed example comes from Baker (2005). There is a certain subspecies of periodical cicada in North America that spends 13 years underground in a larval stage before hatching. That this year-length is irregular calls out for explanation. Biologists suggest that a 13-year length life cycle will minimize overlap with predators and other subspecies. The following calls out for explanation:

(*d*) The length (in years) of the life cycle of periodical cicadas is 13.

Why does a 13-year length life cycle minimize overlap with predators and competing subspecies? The proposed answer is that these are prime-numbered periods, and it is a mathematical law that prime-numbered periods minimize overlap. Accordingly, the fact that the year length of this life cycle is prime explains why it is evolutionarily advantageous (given the ecological constraints of the cicada—its predators, competitors, and facts about its environment). Primeness here plays a genuine explanatory role, so claims the enhanced indispensability argument. This is supposedly an instance of a genuine mathematical explanation of a purely physical phenomenon. There have developed other examples since Baker's original.[3]

The responses to the enhanced indispensability argument mirror the responses to the original. There are *hard roaders*, like Tallant (2013), who attempt to give reformulations of the explanation without appealing to primeness. And there are *easy roaders*, like Knowles and Saatsi (2019), who argue that the way that primeness enters into the explanation is not ontologically committing.

I am unsure of the relation between the original and the enhanced indispensability arguments. In chapter 2, I attempt to trace some of the arguments for the original indispensability principle. And it seems to me that similar arguments cannot be given for the explanatory indispensability principle. Given this, much of my discussion in this dissertation is focused on the original indispensability principle. I do, however, abut the questions of scientific explanations in Chapters 2 and 4. The idea I am probing, though, is that important explanations are guides to the indispensable content of our best scientific theories. This, in oppo-

[3]See Bangu (2008); Lange (2016); Baron (2014, 2016, 2021).

sition to the idea that only the entities that are indispensable to explanations are ontologically committing.

1.3 Overview of Chapters

In Chapter 2, I examine the definition of dispensability and its relation to the indispensability argument. I argue that the reigning definition entails that too many things are dispensable to our best scientific theories, and this entailment is at odds with the purpose for which we seek a conception of dispensability. In light of my arguments, I present a positive proposal that radically shifts our understanding of how dispensability and indispensability arguments work. I will use this new definition of dispensability in my indispensability argument for the existence of composite objects in Chapter 4.

In Chapter 3, I examine a recent rejection of a composite object *dispensability* argument. The rejection claims that our empirical evidence distinguishes between ordinary and composite-free theories, and it empirically favors the ordinary ones (Hofweber, 2016, 2018). I claim that this response to the dispensability argument is not tenable. This is because it presupposes an indefensible thesis about when two empirical consequences are distinct or the same. My argument provides some insight into what our empirical consequences are, and I conclude that empirical evidence is radically metaphysically neutral. This gives us some insight into the significant content of our scientific theories—the content that a scientific realist is committed to—and I show how this insight relates to questions about theoretical equivalence more broadly.

In Chapter 4, I attempt to show that there is no science without composites. I do this by providing an argument that some composite objects are indispensable to some of our best scientific theories. My argument is novel in that it leverages compositional explanations of properties to show that composites are indispensable. Those who are sympathetic to a science without composites have a bevy of tools to paraphrase composites away, but I show that these will not work. This chapter concludes my examination of the prospects of an indispensability argument for the existence of composite objects.

Chapters 5 and 6 concern the relation between the indispensability argument, philosophy of language, and ontological commitment. In Chapter 5, I reject the connection between Putnam and the indispensability argument as presented in P1 - P3. Hilary Putnam (2012) believed that mathematical claims are objectively true but that there are no mathematical objects. There are some initial worries with Putnam's position. First, it seems inconsistent with the conclusion of the so-called Quine-*Putnam* indispensability argument which concludes that there are mathematical objects. Second, it seems inconsistent to affirm that 2+2=4 is objectively true but deny that there are numbers. In Chapter 5, I resolve both of these seeming inconsistencies. To the first, I present a novel interpretation of Putnam's indispensability argument that departs radically from the Quine-Putnam version. To the second, I extract a theory of ontological commitment from close examination of Putnam's comments. I connect this theory of commitment to some recent trends in philosophy of physics.

Finally, in Chapter 6, I present a defense of Melia's easy road strategy of *weaseling*. In metaphysics of science, one question is what ontological commitments are incurred by our best scientific theories. While it is natural to think that one is ontologically committed to anything that is appealed to in the theory, the weasel instead thinks that they can affirm some theory that appeals to Xs and then "prune away" commitment to *X*s. The weasel's strategy crucially relies on the practice of linguistic subtraction—when in ordinary speech we subtract away some of what we say. In Chapter 6, I provide a purely pragmatic theory for determining when an instance of linguistic subtraction is permissible or felicitous. I then apply my theory to three cases of interest in metaphysics of science.

Chapter 2

On dispensability and indispensability

Most philosophers of science and metaphysicians agree that electrons are indispensable and that absolute rest is dispensable to our best scientific theories. What's more, they admit that these beliefs have metaphysical consequences: we should ontologically commit to electrons and reject the structure of absolute rest. But these are easy cases. What of difficult ones? Of numbers? Composite objects? Causation?

Indispensability and dispensability arguments infer from the formulation of our best scientific theories to some claim that we ought or ought not commit to some entity or structure. It is not immediately clear, though, what parts of theories are dispensable. For these arguments to do any work, we must have a clear conception of what it takes for an entity to be dispensable. The historical and contemporary literature has coalesced around a definition, best articulated by Colyvan (2001, 71):

Colyvan's definition. An entity (or structure) X is *dispensable* to a theory T if and only if there exists a theory T^- in which:

(i) T^- doesn't appeal to Xs,

(ii) T^- is empirically equivalent to T, and

(iii) T^- is suitably attractive.

Condition i says that a dispensing theory must no longer appeal to some relevant entity or structure. (Plausibly, replacing appeal to electrons with schmelectrons, which have all the same properties as electrons, is not a way of avoiding appeal to electrons.[1]) Condition *ii* says that a dispensing theory must be empirically equivalent to the original. Two theories are empirically equivalent just in case they have the same empirical consequences. This roughly means that the theories make the same predictions and are confirmed by the same observations. (See Chapter 3 for a detailed examination of empirical equivalence.) Condition *iii* says that a dispensing theory must be suitably attractive. It is important that we do not, for example, move to a theory that is so unattractive that it's not a legitimate candidate for belief.

There is a desideratum on any definition of dispensability: it ought to get the right result in easy cases. The purpose of this definition is to help provide a reasonable metaphysics of science. If the definition does not fulfil its purpose, it fails. I argue that Colyvan's definition fails this desideratum. My hinge case is causation. Colyvan's definition entails that causation is trivially dispensable to our best scientific theories. And causation is not *trivially* dispensable if it is dispensable at all.

I will then parlay my criticisms of Colyvan's definition into a positive proposal. Colyvan's definition presumes that a dispensing theory must always preserve only empirical content. But sometimes, as I argue, we demand that a dispensing theory preserve more than empirical content. If I am right, this reveals an unconsidered first step in any dispensability or indispensability argument. We must *first* identify what content must be preserved, and only then can we ask whether some entity or structure is dispensable or not. Indispensability arguments, to borrow a phrase, do not tell us what exists, they tell us what *else*

[1] Also plausibly, a theory can appeal to some entity or structure just by presupposing, rather than stating, its existence.

exists.[2] They aim to tell us what we must accept beyond that which we already do.

2.1 Preliminaries

We began with two inferences. First, that the indispensability of electrons to our best scientific theories entails that we ought to ontologically commit to them.[3] Second, that the dispensability of absolute rest to our best scientific theories entails that we ought to reject the structure of absolute rest.[4] In the literature on indispensability, these inferences are respectively justified by appeal to the following principles:

Indispensability. If some entity or structure is indispensable to any of our best scientific theories, then we ought to metaphysically commit to that entity or structure.

Dispensability. If some entity or structure is dispensable to all of our best scientific theories, then we ought not metaphysically commit to that entity or structure.

These principles serve as a thruway between the formulations of our best scientific theories and some consequence for our metaphysical picture of the world.

A straightforward argument for the indispensability principle appeals to inference to the best explanation (IBE).[5] Suppose you see stains in the wallpaper

[2]The phrase is Baker's, which I don't fully endorse: "It is not that science tells us what exists; science tells us what *else* exists" (Baker, 2007, 18).

[3]See, e.g., Melia (2000, 474 - 475), Field (2016, 43), Colyvan (2001, Ch. 4.3), and Dorr (2010, §4).

[4]See Norton (2003) and Friedman (1983, p.112).

[5]This is not the only argument for the principle. Another historically famous argument stems from scientific realism. See Putnam (1971), Colyvan (2001), and Field (2016, 1989).

and warped floorboards, and the best explanation for this is that the pipe behind the wall burst. According to a standard form of IBE, in accepting the burst pipe as the best explanation, we commit to the entities and structure that are required in order to put forward that explanation (viz., the burst pipe and its causal relationship to the empirical phenomena) (Field, 1989, 15). *Mutatis mutandis* for our best scientific theories: if a scientific theory is the best explanation for some phenomena, then upon accepting as much, we are committing to the entities and structure that are required in order to state the theory. A definition of dispensability is meant to pick out exactly those entities and structure that are required in order to state the theory.

The most common justification for the dispensability principle relies on naturalism.[6] If one believes that the only reliable guide to metaphysics is science, then if some entity or structure is dispensable to our best scientific theories, we should abandon commitment to that entity or structure. But some reject this variety of naturalism. If so, they might endorse a principle that weakens the consequent of the dispensability principle, e.g., that an entity or structure's dispensability provides some defeasible reason to not commit to it.

We cannot even interpret the dispensability and indispensability principles unless we understand what it means for an entity to be dispensable. We must have a definition of dispensability in order to make these principles precise. Colyvan's definition is orthodoxy within philosophy of science.[7]

---

[6] See Colyvan (2001, Ch. 2.2).

[7] Here are three representative samples. In the 1950s and 60s, philosophers were concerned with the ontological status of *all* theoretical entities. These philosophers often cited the fact that we can construct relatively attractive, empirically equivalent, theoretical-entity free theories. See Craig (1953, 1956), Carnap (1956), Goodman (1957), Scheffler (1957), Hempel (1958), Nagel (1961), Nagel (1965), Maxwell (1962), Putnam (1965), and Hooker (1968a,b). Second, the indispensability argument for the existence of numbers claims that numbers (or some other mathematical objects like sets) are necessary parts of our best scientific theories. Field (2016) is

Colyvan's conditions are relatively straightforward, though I wish to note something about condition ii, my target in the current essay. When one offers a dispensing theory, one is showing that we can retain all of the relevant content of the original without some entity or structure. We shall call this content that must be preserved the *privileged* or *scientifically important* content of the theory. To dispense with some entity or structure X, we provide a suitably attractive theory that preserves the privileged content of the original theory and doesn't appeal to Xs. According to Colyvan's definition, the privileged content of a theory is the theory's *empirical* content, captured in condition ii. Part of the appeal of Colyvan's definition is that it is maximally empirically conservative: if we accept it and the indispensability principle, then we are only required to commit to the empirical phenomena and exactly as much structure and as many entities as are needed to explain the empirical phenomena. In this way, Colyvan's definition presupposes that a theory's privileged content is exactly its empirical content.

As we saw above, the desideratum on a definition of dispensability is that it gets the right result in easy cases. More precisely, a definition of dispensability ought to be materially adequate when conjoined with the dispensability principles: it should entail the dispensability of entities or structure we obviously ought to reject and it should not entail the dispensability of entities or structure we obviously ought not reject.

---

the locus classicus of attempting to provide empirically equivalent, attractive, number-free alternatives to scientific theories, which he did to a portion of Newtonian Gravitation Theory. Third, some are concerned with the *dis*pensability argument in object metaphysics that claims composite objects are dispensable to our best scientific theories (Dorr, 2002; Brenner, 2018; LeBrun, 2021). There, philosophers presuppose that what it takes to show composites to be dispensable is that we provide alternative theories (or a schema for constructing alternatives) that do not appeal to composites, are suitably attractive, and are empirically equivalent to our ordinary theories.

2.2 Against Colyvan's Definition

My objection to Colyvan's definition is that empirical equivalence isn't exactly the relation that a successful dispensing theory bears to the original theory, and that this contributes to his definition failing the desideratum.

Here I provide two examples. The first motivates the thought that *ii* doesn't do enough to guarantee that a dispensing theory preserves the privileged content of the original theory. I don't take this first one to be a counterexample to Colyvan's definition. There are responses that he can give to it, but my alternative diagnosis is more plausible. The second example is a more traditional counterexample. Colyvan's definition entails that some entities which aren't obviously dispensable are trivially dispensable.

2.2.1 Geometry

We consider the history of axiomatizations of geometry. The traditional way of formulating geometry is analytic geometry, which appeals to points and lines on a coordinate system together with unit of distance. Analytic geometry appeals to a primitive distance function which maps pairs of points to real numbers: the distance between a and b is n. Because this geometric system uses a coordinate system with a unit of distance, it requires the apparatus of the real numbers.

Synthetic geometry, axiomatized by Hilbert (1930) and Tarski (1959) attempts to do away with a coordinate system and a distance predicate, and thus numbers. (Don't accord philosophical weight to the names 'analytic' and 'synthetic'.) Synthetic geometry will not entail that the distance between any two points is equal to some real number n. In fact, a distance predicate (as a polyadic relation between a pair of points and a real number) is incomprehensible in synthetic geometry. Instead, it gets by with relative notions like congruence—*the distance between two points a and b is the same as the distance between b and c*. Accordingly, synthetic geometry does not require numbers, a coordinate system, or a metric.

All the same, it is well-known that these two formulations of geometry capture all of the same relevant theorems and axioms. Synthetic geometry can accommodate all of the theorems of analytic geometry without the use of numbers. Thus, it seems that synthetic geometry explains everything that analytic geometry does, but without the use of numbers. If so, then numbers are *dispensable* to theories of geometry. And this consequence has generally been the lesson from the move to synthetic geometry.[8]

Consider whether synthetic geometry meets conditions *i* - *iii*. Regarding *i*, it seems clear that synthetic geometry does not appeal to numbers. Likewise, regarding *iii*, synthetic geometry is at least as attractive as analytic geometry. Now consider *ii*, the demand that a dispensing theory be empirically equivalent to the original theory. It almost seems like a category mistake to ask whether synthetic and analytic geometry are empirically equivalent. Neither theory has empirical consequences. So, *prima facie*, it seems unanswerable whether condition *ii* is met, even though it seems that synthetic geometry dispenses with numbers.

Certainly, Colyvan's defender has replies. They may say that there is a sense in which the two theories have empirical consequences—in particular, when we assume them to be theories of space. Analytic and synthetic geometry *as theories of space* are empirically equivalent. If so, we can count synthetic geometry as a case of dispensing with numbers. The problem with this reply is that it seems that

[8]See Burgess (1984), Burgess and Rosen (1997, IIA), and Field (2016).

synthetic geometry *as a theory of geometry* also dispenses with numbers. Or they may say that trivially these geometric formulations are empirically equivalent. They have the same empirical consequences: none at all. The problem with this reply is that it would entail that $\forall x(x = x)$ dispenses with numbers as well. It has the same empirical consequences as both analytic and synthetic geometry, but does not appeal to numbers (or points or lines, for that matter).[9] Or they may deny the relevant dispensability principle which says that dispensability is relevant for pure mathematical theories. Instead, they insist that dispensability only matters for physical theories. Strictly speaking, this response neutralizes the counterexample, as the example would no longer entail anything about what we ought to commit to. However, I am not especially moved by this response. The example is meant to bring out something important about *dispensability* as it applies to all theories. Denying a variety of the dispensability principle seems to change the subject.

So, this example puts *some* pressure on Colyvan's definition, but there are ways to defend it. My primary aim here is to motivate the following framing of this example. We agreed that some core claims of analytic geometry must be preserved in any adequate axiomatization of geometry. This is the privileged content of analytic geometry. The privileged content includes Playfair's axiom, that there is at most one line that can be drawn parallel to another given one through an external point. But the privileged content does not include a measurement, which assigns a numerical value to each line segment. Synthetic geometry shows that we can preserve the privileged content of analytic geometry without appealing

[9] One could reply to this example by claiming that $\forall x(x = x)$ is not sufficiently attractive on the grounds that it does not preserve conservativeness over analytic geometry. I consider this in detail below in §3.3.

to numbers. More generally, we might offer the following two-step procedure of dispensing: identify the privileged content of a theory, and then any successful dispensing theory will be one that preserves that content while doing away with the dispensable part. And while the privileged content usually includes empirical consequences, it might have nothing to do with the empirical realm, as with the dispensing of numbers in geometric axiomatizations.

2.2.2 Causation

My second example targets Colyvan's definition at its core. His definition entails that some not obviously dispensable parts of our theories are trivially dispensable. It thus fails to satisfy the desideratum. The basic idea, in line with the lesson from geometry, is that Colyvan's definition wrongly identifies a theory's privileged content.

Suppose a ball is thrown at a window and the window shatters. Further suppose our best scientific theories explain that the throwing of the ball caused the window to shatter. Call the theory that explains this T^1. Though this is just a toy example, there are purported instances of genuine causation in our best physical theories.[10] For simplicity, let's assume that if one accepts a theory that contains a causal explanation, they are committing to the *structure* of causation (rather than, e.g., the existence of causal forces). And let's assume that the relata of causal relations are events.

It is a live debate whether causation is dispensable to our best scientific theories.[11] Philosophers in this debate carefully examine these theories and see whether the role played by causation is dispensable or indispensable to them.

[10] See the examples given by Mumford and Anjum (2010, 2011b,a).

[11] Cf. Woodward (2015) and Weaver (2019).

I hereby take it that, in T^1, causation is neither obviously dispensable nor obviously indispensable. Good metaphysics of science is needed to judge. I will show, though, that Colvyan's definition entails that causation is trivially dispensable to T^1.

Let's consider T^1 in detail. T^1 is the theory that the throwing of the ball caused the window to shatter. Some trivial consequences follow from T^1, like that the throwing of the ball occurred, and that the shattering of the window occurred, and they occurred in sequential order. Some non-trivial consequences also follow from T^1. First, that the two events are not merely sequentially ordered. There is a difference between mere temporal sequencing and causation, and T^1 entails that the throwing of the ball and shattering of the window are not mere temporal sequences. Second, that events which are causally related are nomologically entangled. There's a sense in which if the first event occurred, the second *had* to occur. It was no accident that the window shattered following the throwing of the ball.

Here's how we can trivially dispense with causation from T^1 if we adopt Colyvan's definition. We construct an alternative theory, T^{1-}, which is comprised of only the trivial consequences identified above. It will entail that the throwing of the ball occurred, that the shattering of the window occurred, and that these two events occurred in sequential order. Crucially, it will not entail that there is a difference between causation and mere sequential ordering, and it will not entail that the two events occurred with nomological necessity. T^{1-} will be comprised of exactly those consequences of T^1 that are non-causal.

At first glance at least, T^{1-} meets Colyvan's conditions for dispensing with causation. First, it does not appeal to causation, satisfying i. We have genuinely eliminated the structure of causation in T^{1-}. Second, it is empirically equivalent to T^1, satisfying *ii*. Every empirical consequence entailed by the original theory will be entailed by T^{1-}. In both theories, the observations and predictions are identical: *if* the ball is thrown at the window, then the window shatters; and these events will occur sequentially. There's good reason for their empirical equivalence. A necessary condition on causation is sequential ordering of events. And the only empirical consequences of causal explanations are the sequential ordering and occurrence of the events. So, as long as a theory entails the same sequential ordering and occurrence consequences as some theory with causal explanations (and there are no other differences between the two), the two are empirically equivalent. Accordingly, T^{1-} satisfies Colyvan's condition *ii*. And this simple causation dispensing theory is not egregiously unattractive in terms of unification, fruitfulness, etc. (We will examine this in detail shortly.) It preliminarily satisfies *iii*.

This simple dispensing procedure is generalizable. Every scientific theory that appeals to causation has a variant that is empirically equivalent, does not appeal to causation, and is sufficiently attractive. So, Colyvan's definition of dispensability permits the trivial dispensing of causation, and the dispensability principle entails that we ought not commit to the structure of causation. Something has gone wrong. It seems like, regardless of whether causation is actually dispensable to our best scientific theories, we cannot show this via the simple dispensing method. Thus, we should reject Colyvan's definition because it fails this desideratum.

2.2.3 Colyvan's Reply

There's a conspicuous response on behalf of Colyvan: the simple causation dispensing theory just isn't attractive and so T^{1-} does not dispense. There are two versions of this objection, and we shall treat each separately.

The first version of the attractiveness objection goes like this: A condition on a successful dispensing theory is that it is not objectionably unattractive, and T^{1-} is objectionably unattractive, so it does not dispense with causation. For this objection to have any force, we must identify features of T^{1-} that explain why it is unattractive. It cannot be that T^{1-} fails to make the appropriate predictions or observations, since we crafted the theory to have exactly the same empirical content. So we cannot complain that the simple causation dispensing theory fails on any grounds that impinge on the empirical. Nor is T^{1-} inconsistent or incoherent. T^{1-} also does not fail on aesthetic virtues like simplicity, beauty, or unification; it is more simple than T^1 and explains more phenomena using fewer theoretical posits.

The only thing that Colyvan could identify to justify the claim that T^{1-} is objectionably unattractive is that it fails to preserve the non-trivial consequences of T^1. T^{1-} doesn't distinguish between cases of mere temporal sequencing and cases of causation. The theory doesn't even have the linguistic resources to distinguish them. Moreover, T^{1-} doesn't tell us how, when there is a causal relation between two events, we think that their occurrences hold with nomological necessity. *This* is the sense in which T^{1-} is objectionably unattractive.

Colyvan (or one sympathetic to Colyvan's definition), however, is not privy to this objection. It is inconsistent with a core tenet of his view. Recall that part of Colyvan's view is that the privileged content of a theory is the *empirical content* of that theory. Colyvan's definition is suited toward an empirically minded philosopher who wishes to be maximally conservative over the empirical. Condition *ii* of his definition was meant to guarantee that the dispensing theory captured the privileged content, which is exactly only its empirical consequences. Colyvan cannot then object to T^{1-} on the grounds that it does not preserve T^1's privileged content, since by his own standard it does. T^{1-} is empirically equivalent to T^1, and Colyvan's definition presupposes that the privileged content is preserved if two theories are empirically equivalent.[12] Accordingly, Colyvan would impugn his own view if he said that T^{1-} did not capture the privileged content of the original theory.

I endorse the claim that T^{1-} does not preserve the privileged content of T^1, and for this reason it does not dispense with causation. But Colyvan cannot give this response to the simple causation dispensing theory. At the very least, Colyvan's formal notions of attractiveness, having to do with features of a theory like elegance, simplicity, and so on, are inadequate for explaining how T^{1-} is objectionably unattractive. Instead, to show that T^{1-} is objectionably unattractive, we must invoke the content that the theory fails to preserve. And once we do that, we are better served by my alternative conception of dispensability

The second version of the attractiveness objection goes like this. Colyvan can concede that T^{1-} is suitably attractive, but instead strengthen condition *iii*. It is not the case that a dispensing theory must be suitably attractive; rather, it must be at least as attractive as the original theory. The idea behind this objection is intuitive. We ought to accept the best theory available. T^1 is a *more attractive* theory

[12]Strictly speaking, Colyvan's definition does not—as written—*say* that the privileged content is the empirical content, as it is simply a definition. Rather, the spirit of, and the motivation for, the definition presuppose that the priveleged content is the empirical content.

than T^{1-}, so even if we can "get by" without causation, this isn't enough to show that causation is dispensable. Of course, this is a concession to my argument, but it is not ad hoc.

The response is to replace condition *iii* with the following:

iii+ T^{-} is at least as attractive as T.

This would likely respond to the counterexample. It is plausible that T^{1-} is slightly less attractive than T^{1}, and if our definition of dispensability had condition *iii*+, T^{1-} would not dispense with causation.

There are, however, independent reasons to reject *iii*+ as a condition on dispensing. My argument here takes us into considerations about dispensability in general. In particular, if our definition of dispensability requires that a dispensing theory be no less attractive than the original theory, then (in)dispensability arguments collapse into arguments only about theory choice. And I will argue this is a bad result.

Suppose that (in)dispensability arguments collapse into arguments about theory choice. By this I mean that once we determine which theory is the best among a slate of alternatives, all entailments of dispensability and indispensability are settled: the entities that are appealed to in the best theory are indispensable (to that theory), the entities not appealed to in the best theory are dispensable (to that theory). There is nothing more to be said about the (in)dispensable parts of that theory. If so, then (in)dispensability considerations are redundant. Once we determine which theory is the best, no new metaphysical entailments can be gained by asking which parts of the theory are dispensable or indispensable.

However, dispensability and indispensability considerations are *not* redundant. We can accept that some theory is our best—that there are no alternatives that are more attractive according to the theoretical virtues like simplicity, fruitfulness, etc.—and still have questions about whether all the entities and structure that are appealed to within that theory are *required* in order to formulate the theory. The idea here is that the virtues which determine the best theory may not perfectly match the reasons for metaphysical commitment. If some theory is less cognitively cumbersome to humans, or is more beautiful, or is more likely to generate novel predictions, which are all theoretical virtues, this doesn't entail that the metaphysics of that theory is more correct than the alternatives. This isn't to say that theoretical virtues play no part in determining the correct metaphysics of science, just that they are not perfect determiners.

There are examples in the history of science where, plausibly, some theory is deemed our best, but we are hesitant to endorse some entity as indispensable. At the turn of the 20th century, chemists debated the existence of atoms despite their appearance in our best theories. The theories atoms appeared in were incredibly well confirmed, fruitful, unifying, and had all the relevant theoretical virtues we take to be indicative of true scientific theories; they were among our best. Yet many chemists were reluctant to commit to the existence of atoms until Perrin's 1913 experiment showing that atoms were responsible for Brownian movement, at which point the consensus around atoms shifted. It seems plausible that scientists justifiably accepted that the theories in which atoms appeared were the best explanations of the relevant phenomena, but they believed we didn't have enough evidence to show that atoms were *indispensable*.[13] If this story is correct, then (in)dispensability considerations are not redundant. We should not demand that a dispensing theory is at least as attractive as the original theory, only that it

[13]Cf. Maddy (1997), Castro (2013), Brown (2015), and Boyce (2018).

should be attractive enough. As a result, this second version of the attractiveness response should be rejected.

My resulting picture of dispensability looks like this. Determining whether some entity is dispensable or indispensable is not tantamount to looking only to the *most attractive theory* and seeing which entities are appealed to within that theory. Rather, we use the theoretical virtues to identify a collection of candidate best theories in some domain. These theories will all share the privileged content, and otherwise will differ similarly to how T^{1-} and T^1 do—in the theoretical structure and entities involved. These theories must be suitably attractive, meeting some threshhold for candidates for belief.[14] And we need not assume that the theoretical virtues will single out a unique *best* theory. Once we have identified this collection of theories, we can determine the dispensable and indispensable parts. The indispensable parts are the entities and structure that are shared among all candidate best theories. Some entity or structure is *dispensable* if there is at least one candidate theory that does not appeal to that entity or structure.

2.3 Some lessons

We ought to reject Colyvan's definition. *ii* is the wrong condition for guaranteeing that a dispensing theory preserves all a theory's priveleged content. Sometimes, a dispensing theory must preserve more than just the original's empirical consequences, e.g., a candidate dispensing theory for causation must preserve the nomological necessity between events linked by causation (or we must explain why we don't need to preserve this). Our rejection of Colyvan's definition has profound impacts on the way we understand dispensability and indispensability

[14]Craigean theories, e.g., will plausibly not meet this threshhold.

arguments.

The first impact is, in a sense, dialectical. The traditional picture of dispensability or indispensability is this:

> We aim to determine the metaphysical import of our scientific theories. A successful indispensability argument will show that some entity or structure's existence is "given by", or follows from, our best scientific theories. A successful dispensability argument will show that some entity or structure's existence is not given by, does not follows from, our best scientific theories. In this way, sound dispensability and indispensability arguments tell us what *science says exists*.

If my arguments against Colyvan's definition are sound, though, this traditional picture is undermined. For recall: T^{1-} fails to dispense because it does not preserve all of T^{1}'s privileged content. There are, then, two steps to any dispensability or indispensability argument. The first step, absent in the traditional picture and smuggled into Colyvan's condition *ii*, is to determine a theory's privileged content. The privileged content of theories of space and time might be different than the privileged content of a theory with causal explanations. For the case of T^{1}, the privileged content included the non-trivial consequences about causation. Before we can even adjudicate whether some dispensability or indispensability argument succeeds, we must have a univocal answer on the theory's privileged content. The second step is to determine what else we must commit to. We are committed to whatever is required to explain a theory's privileged content.

It is understandable why the traditional picture included condition *ii*. Colyvan, and many others who were concerned with dispensability, is an empiricist who traces his roots to Quine. Naturally for him, we are only committed to whatever else is required to explain the empirical phenomena. But for those of us who do not share these proclivities, we must first have an answer to the question of what the privileged content of a given theory is.

The second impact of our rejection of Colyvan's definition is that it clarifies the three ways one may respond to a given dispensability or indispensability argument. The first way is to reject that the argument succeeds in establishing that some entity or structure is dispensable or indispensable, in the sense that the conditions for dispensing haven't been met. The second way is to reject the relevant dispensability or indispensability principle. If one is not an austere naturalist, they may reject some dispensability argument on the grounds that they don't accept the relevant dispensability principle. The third way to reject a dispensability or indispensability argument is illustrated by my arguments here. We may reject a putative dispensability or indispensability argument on the grounds that the argument presupposes the wrong privileged content for dispensing. We might, e.g., agree that causation is dispensable to capturing some theory's empirical consequences, while simultaneously claiming that a successful dispensing theory must preserve more than just the empirical consequences. This response constitutes a rejection of the dispensability argument.

Some big picture worries remain. Whatever problems Colyvan's picture had, at least it provided a complete picture of dispensability. It provides an algorithm for determining what the significant content of a theory is. Everyone agrees that the empirical content is significant and metaphysically committing. But what else beyond that? Colyvan's picture says that the other significant content is whatever is needed to explain the empirical stuff. But I am proposing a rejection of Colyvan's view in favor of one which says that, sometimes, in some theories, the significant content is the empirical stuff, *plus* some other "privileged" content, and additionally whatever is needed to explain all of that. How do we know what this privileged content is? How do we know, e.g., that a theory of causation must preserve some extra-empirical content?

These are deep and difficult questions about the project of the metaphysics of science. The tools that we have at our disposal for determining a theory's significant content seem to be, on the one hand, Colyvan's empiricism, and on the other, *a priori* metaphysics. Metaphysics of science must forge a middle ground, providing rational reconstruction of scientific theories that is neither pure empiricism nor pure *a priori* metaphysics. What I have done here is provide an argument for this middle ground.

Chapter 3

What are Empirical Consequences? On Dispensability and Composite Objects

In Lorentz's ether theory, ether was postulated as a substance that acted as the medium for the transmission of light through space. Philosophers of science tell the following story for why physicists no longer accept the existence of ether (see, e.g., Norton (2003)).

> Lorentz's ether theory had many predictions and observations, including length contraction, which is the phenomenon that a moving object's measured length will be shorter than its proper length. Einstein proposed an alternative theory, special relativity, that has the exact same predictions and observations (Bradley, 2021, 9), including length contraction, without needing to posit the existence of ether. Special relativity thus showed physicists that ether is dispensable.[1]

Einstein showed that ether is dispensable to theories of light by providing an attractive alternative theory which (i) does not appeal to ether and (ii) is empirically equivalent to Lorentz's. Physicists took the dispensability of ether as good reason to abandon ontological commitment to it. Today, philosophers of science and metaphysicians operate on the same understanding of dispensability. Some claim that numbers are *in*dispensable to scientific theories by arguing that one cannot provide nominalistic alternatives that meet these two conditions

[1]It is common for philosophers of science to tell a similar story for why physicists no longer believe in absolute space. See Friedman (1983, p. 112). Additionally, absolute space and ether are taken to play roughly the same role in Lorentz's ether theory—as providing a privileged inertial frame.

(cf. Colyvan (2001) and Field (2016)). Philosophers generally take the dispensability of an entity to have ontological consequences; if an entity is dispensable, we should abandon ontological commitment to it.

One argument within the metaphysics of ordinary objects is best understood as a dispensability argument. This Composite Object Dispensability Argument (CODA) concludes that ordinary composite objects like metal bars are dispensable to our best scientific theories (cf. Rosen and Dorr (2002), Sider (2013), and Brenner (2018, 660)). The idea is that the only things needed to explain phenomena like conduction are the microphysical particles that "make up" the metal bar. The CODA offers a strategy for constructing a variant of any scientific theory, and these variants supposedly meet the criteria for dispensing with composites. They are meant to be attractive theories that (i) do not appeal to composites and (ii) are empirically equivalent to the ordinary theories that do appeal to composite objects. Proponents of this dispensability argument take themselves to have shown that composites are dispensable to any scientific theory which they appear, and this is meant to be evidence that there are no composite objects like metal bars.

Here I am concerned with empirical equivalence and its relation to the CODA. Two theories are empirically equivalent in virtue of sharing the same empirical evidence or content, which is understood as having the same empirical consequences.[2] According to the CODA, the *composite-free* theories have the same

[2]Here we are presupposing some rough distinction between the empirical and non-empirical. Such a presupposition raises questions and concerns about the theory-ladenness of observation (see Fodor (1984)). There are difficult questions about any particular distinction between the empirical or non-empirical, and whether consequences like *There is an electron in the bubble chamber* are empirical. Here, we work with an intuitive distinction between theory and observation, and we rest easy knowing there are difficult boundary cases. Unlike the logical positivists, we are not drawing the boundary between meaningfulness and nonsense, and so the question of whether some particular consequence is empirical or not is less pressing than it was for them. See Lewis (1988, 4) for a similar motivation.

empirical consequences as the *ordinary* theories that appeal to composites; our empirical evidence is neutral between them. For example, an ordinary theory might have the empirical consequence that there is a metal bar in the lab. A composite-free alternative would have the empirical consequence, roughly, that some microphysical particles "arranged metal bar-wise" are in some particular place. These two empirical consequences are taken to be the same.

Some, though, disagree with the CODA's claim of empirical equivalence. For example, Hofweber claims that composite-free theories are *trivially* empirically inequivalent to their ordinary counterparts. He says,

> There is lots of evidence that supports the [composite] object theory over the things arranged object-wise theory. The object theory predicts that there is a bar of metal in the lab, the object-wise theory doesn't predict it. That there is such a bar can be confirmed with the observation that there is such a bar of metal in the lab. (Hofweber, 2016, 199)

Hofweber is claiming that the empirical evidence for our scientific theories—the observations and predictions—is "clearly in favour of [the existence of composite] objects" (Hofweber, 2018, 321-322). The claim is not merely that we have more or better *scientific* reasons to prefer ordinary theories, but rather that we have better *empirical* reasons to prefer ordinary theories. He is clear that "[empirical] scientific evidence does in fact distinguish" between composite-free and ordinary theories, and it favors the ordinary ones (*ibid.*). Hofweber here presupposes a thesis about the individuation conditions of empirical consequences. In particular, he is committed to the thesis that differences in the "thick" mereological content between two empirical consequences suffices for a difference between those empirical consequences. As a result, if two theories' empirical consequences differ in their "thick" mereological content, then they are empirically inequivalent. Otherwise it could not be that our empirical evidence supports the ordinary theory over the composite-free one. In this way, Hofweber is committed to the empirical significance of "thick" mereological content; he is presupposing that "thick" content *matters* to the scientific theory. Accordingly, the empirical consequences of a composite-free theory are trivially inequivalent to the empirical consequences of our ordinary theories. Call this the trivial response to the CODA.

At this point, the dialectic is brought to a halt. The CODA claims that composite-free theories are clearly empirically equivalent to their ordinary counterparts. The trivial response claims that composite-free theories are trivially empirically *in*equivalent to their ordinary counterparts. Without some clear understanding of the individuation conditions for empirical consequences, we cannot adjudicate this disagreement.

My topic is the individuation conditions of empirical consequences. My proximate aim is to settle whether the trivial response to the CODA is tenable. As we will see, my ultimate target is anyone who claims that "thick" makes an *empirical* difference. I will argue that there is no good conception of empirical consequences that will permit the trivial response to the CODA. My argument proceeds by considering successful cases of dispensing, like when we rid our physics of ether. My thesis supports the position I call empirical quietism, which entails that we cannot settle any distinctively metaphysical disputes by appealing to empirical evidence. Although quietism may seem obvious to some, there are two additional philosophical payoffs that the following discussion yields.

First, because empirical equivalence plays a prominent role in many ontological arguments, it is imperative that we understand the conditions under which two theories are empirically equivalent. For any pair of empirically equivalent theories, there are at least three pressing ontological questions. First is the underdetermination question, which asks whether we should be committed to either theory's ontology; it seems that the existence of an empirically equivalent alternative should threaten to undermine our confidence in the theory we accept (cf. Bas (1980), Laudan (1990), Stanford (2009), and Worrall (2011)). Second is the theoretical equivalence question, which asks whether the two theories are fully equivalent; it may be that two theories seem to have different ontological commitments, but that this is a merely apparent difference (cf. North (2009), Curiel (2014) Barrett (2015), Barrett (2019), Barrett (2020a), and Weatherall (2019b)). Third is the dispensability question, which asks whether we have reasons to prefer one theory's ontology to another; it may be that we have reasons similar to parsimony to accept one of two empirically equivalent theories. For philosophers of science and metaphysicians, determining exactly when two theories have the same empirical consequences is important for the role that empirical equivalence plays in these arguments, and my conclusion will entail that some conceptions of empirical equivalence—ones antagonistic to quietism—should be rejected.

Second, my discussion will allow us to draw broader lessons on theoretical equivalence. The literature on theoretical equivalence is concerned with the conditions under which theories are fully equivalent, in the sense of saying the same thing about the world. Some people in that literature endorse a position adjacent to Hofweber's. Whereas Hofweber argues that an empirical consequence's "thick" content is relevant for individuation, these folks argue that a *non-empirical* consequence's "thick" content is likewise relevant for individuation. For example, North (2009) argues from ostensible differences in the structure of two formulations of classical mechanics—a difference in "thick" content—to the inequivalence of those formulations. If my arguments against those who ascribe empirical significance to "thick" content are correct, we should tread lightly. If we will have learned anything, it is that two theories having the same or different consequences is a complicated matter, not to be decided by only considering metaphysically rich content.

3.1 Preliminaries

The trivial response claims that composite-free theories are trivially empirically inequivalent to ordinary scientific theories, and it rejects the CODA on those grounds. We begin by investigating exactly what the trivial objection is committed to. The CODA, as an argument in its own right, is rarely discussed in the literature. It is usually implied by the claim that appeal to ordinary objects is "pragmatic" (cf. Healey (2013, 53), Brenner (2018, 660)[3]). I hope the following explanation of the CODA shows that there is philosophical value in pursuing it explicitly.

3.1.1 The CODA

Dispensability arguments are given by those who draw ontological commitments from an entity's dispensability. Consider the following dispensability principle: If some entity is dispensable to our best scientific theories, then we ought not

[3]Healey (2013) does not explicitly endorse the conclusion of the CODA—that we ought to reject the existence of composite objects—but he does accept that what scientists regard as composed is partially determined by the context in which the scientist is operating.

be committed to its existence. Scientific realists may be inclined to accept the dispensability principle if they believe that science is the best guide to answering ontological questions. This dispensability principle seems to be what drove physicists to abandon commitment to ether.

To show that an entity is dispensable to some scientific theory, we must provide an alternative theory that dispenses with that entity. A dispensing theory is one that fits the following account, adapted from Colyvan (2001, 77):

Dispensability. An entity (or structure) X is dispensable to a theory T if and only if T has an attractive variant T^- for which:

(i) T^- does not appeal to Xs, and

(ii) T^- has the same empirical consequences as T.

If there's a theory that appeals to some entity (or structure[4]) we suspect to be dispensable, to show its dispensability we provide an attractive theory that does not appeal to the entity (or structure) and has the same empirical consequences as the original. (i) requires that the variant does not appeal to the dispensable entity; if it does, we haven't shown that the entity is unnecessary. (ii) requires that the variant has the same empirical consequences; if it doesn't, then that suggests that the entity does play an explanatory role.[5] These conditions are each necessary and jointly sufficient for showing an entity to be dispensable.

Empirical consequences, intuitively, are the observations and predictions of a theory. Theories make observations and predictions about the world; they tell

[4]Cf. (North, 2009, 64) and Barrett (2020a, 2 - 3).

[5]What if T^- explained more than T did? Does the definition entail that T' does not dispense with the entity in question? Per the definition provided, it seems that the entity is not dispensable. Some may find this problematic, since it seems like we ought to prefer T^- to T. But note that we have reasons beyond the dispensability of the entity to prefer T^-—it explains more! So we have ordinary, empirical reasons to prefer T^-, rather than *a priori* reasons of dispensability.

us what it is like and what it will be like. Sameness of empirical consequences, or empirical equivalence, occurs when two theories make the same observations and proffer the same predictions about the world. For the time, we will operate on this intuitive notion of empirical consequences and empirical equivalence.

The CODA argues that all composite objects are dispensable to our best scientific theories. One might be sympathetic to the CODA because they consider appeal to composite objects to be merely pragmatic. Composites like iron bars may just be heuristics, allowing us to better understand complex scientific explanations but not serving a genuinely explanatory role in those explanations (Brenner (2018, 660)).[6] Instead, the complex physical phenomena are fully explained by partless microphysical entities and processes.[7] If this is persuasive, then it seems that ordinary scientific theories need not appeal to composite objects, since all we need to explain everything is microphysical. In accordance with the standard account of dispensability, the CODA provides variants for each scientific theory that appeals to composites. Instead of meticulously constructing these variants, the CODA offers a strategy to construct, for every ordinary scientific theory that appeals to composite objects, a composite-free variant that has the same empirical consequences. That strategy, which may be familiar, works like this.[8]

Consider the theory of rust *R*: rust, a reddish-brown substance that is the result of corrosion, is an iron oxide that forms on iron in the presence of oxygen together with water or air moisture. The iron is a reducing agent, giving up electrons, while the oxygen is an oxidizing agent, gaining electrons, resulting in iron oxide—rust. As an ordinary scientific theory, *R* appeals to composite objects;

[6]Cf. Osborne (2016).

[7]A similar motivation is found in Sider (2007), where he argues that a composite object does not afford a thing with any causal powers beyond those had by the parts of that thing.

[8]Here I follow Dorr (2002) and Rosen and Dorr (2002).

rust is an *iron oxide* that forms on *iron* in the presence of *oxygen* and *water*.

The CODA presents a strategy for constructing a variant of R that does not appeal to composites and has its same empirical consequences. Let's simplify and consider only one empirical consequence and try to rid R of just one composite object. Take the following observation delivered by the theory when some particular iron bar rusted after being exposed to moisture-rich air.

(*b*) This iron bar rusted.

(*b*) is an empirical consequence that appeals to a composite: this iron bar.

Proponents of the CODA think all genuine explanatory work is done by the partless microphysical particles that make up the "iron bar." If they are correct, then to reveal the actual explanatory structure of our theories, we ought to replace all appeal to *iron bars* in R with appeal to only the partless, microphysical mereological *simples* and the complex ways in which they are arranged to be iron bar-wise.[9] (We must also replace appeal to properties that are realized by composite objects with collective properties that are realized by arrangements. Though it is much more complicated than this, I will simply refer to the collective property variant of any ordinary property by appending it with the prime symbol. This way *rusting* becomes *rusting′*, where *rusting′* is realized by simples in arrangements.) This is how to construct a variant of R where appeal to iron bars is replaced with appeal to simples (and appeal to rusting is replaced with appeal to rusting′). Call this new theory R^{-}, which has the following empirical consequence:

(b^{-}) These simples arranged iron bar-wise rusted′.

[9]The 'arranged *X*-wise' locution is from van Inwagen (1990). Here I assume both that the composite objects that science appeals to are not extended simples (Cf. McDaniel (2007)) and that the world is not *gunky* (Cf. Sider (1993)).

The CODA claims that R^{-} is empirically equivalent to R and that (b) is the *same empirical consequence* as (b^{-}). The idea is this: When we observe what we'd ordinarily describe as an iron bar rusting, the two theories can equally sufficiently explain it. R will explain that the iron bar went through the process of rusting, and R^{-} will explain that the simples arranged iron bar-wise collectively went through the process of rusting′. Sure, the iron bar-free variant will be more difficult to comprehend, since it appeals to philosophical entities like simples and is cognitively cumbersome, but this is not a mark against its empirical adequacy.

Though the CODA claims that (b) and (b^{-}) are the same empirical consequence, we note that they have different "thick" metaphysical content. This notion of thick metaphysical content is to be understood as a consequence's associated underlying metaphysical picture. (b) is associated with an underlying metaphysical picture where there are composites, and (b^{-}) is associated with one where there are no composites. According to the CODA, the mere fact that (b) and (b^{-}) have different thick metaphysical content is not sufficient for them being distinct empirical consequences. As we'll see, Hofweber demurs.

R^{-} purportedly meets the conditions for showing iron bars to be dispensable. First, R^{-} does not appeal to iron bars. Second, if the CODA's reasoning is correct, R^{-} is empirically equivalent to R. Given these facts about R^{-} together with the conditions for showing an entity to be dispensable, we have shown that iron bars are dispensable to theories of rust.[10] Accordingly, the CODA concludes that we

[10]Some suggest that providing a dispensing theory also requires that the variant one provides is more attractive than the original theory. Colyvan (2001) and Field (2016) make these claims. If one accepts this, one might be tempted to reject the CODA on the grounds that the composite-free variant is not sufficiently attractive to show that composite objects are dispensable. There are rumblings of this response in the objects literature already. For example, Parsons (2013, 332) denies composite-free scientific theories because "composite objects play a crucial role in the best explanations of my experience." I take this to mean that showing an entity to be dispensable requires offering an entity-free theory that *best* (or better) explains the phenomena, and not offering

ought to reject the existence of iron bars. This is the general strategy for dispensing with composite objects that appear in different theories, which the CODA takes to show that all composite objects in all scientific theories are dispensable.

3.1.2 The Empirical Significance of Thick Content

Hofweber straightforwardly asserts that our empirical evidence decides in favor of the existence of composite objects. The idea seems to be that our theories' empirical consequences come "pre-loaded" with a particular mereological picture and that this mereological picture, which presents the world as containing composite objects, is representationally significant. In brief, Hofweber is committed to the thesis that thick mereological content is empirically significant.

There are philosophers who are committed to similar theses. They presuppose or otherwise argue that some other thick metaphysical content is empirically significant. It would behoove us to see a few of these other arguments in metaphysics. Here's a test for whether a philosopher is committed to the thesis that some thick content is empirically significant: if someone claims that a theory's evidence is incompatible with one but not all sides of some distinctively metaphysical debate, then they believe that the metaphysical content of that debate, when it appears in theories, is empirically significant. If one was an evidential quietist, who does not think that thick content is empirically significant, they would deny that a theory's evidence could be compatible with one but not all sides of a distinctively metaphysical debate. Note that adherence to the empirical significance of some thick content (e.g., mereological) does not entail that one adheres to the empirical significance of any other thick content (e.g., identity).

an entity-free theory that simply explains the phenomena.

Here are three instances of philosophers who are committed to the empirical significance of thick metaphysical content.

First, in the literature on personal identity, Blatti (2012) argues from evolutionary theory to the thesis that human persons are identical to organisms. His idea is that any non-organism metaphysical position will be inconsistent with the empirical consequence of evolutionary biology that *my ancestor is an organism*. Blatti here straightforwardly presupposes that a theory's thick content regarding *identity* is empirically significant; our empirical evidence for the theory of evolution apparently settles the debate over personal identity. Second, Williamson (2007, 223) argues that scientific theories that are composite-free will not be supported by the same evidence that our current theories are; this is because the evidence for our current theories is committed to the existence of composite objects.[11] E.g., the evidence for *R* consists of claims like the *hygrometer* measured such-and-such humidity levels, which appeals to a composite object. Williamson is committed to the empirical significance of mereological content in a manner weaker than Hofweber's. He is presupposing that there is a *prima facie* evidential problem for composite-free scientific theories, whereas Hofweber argues that there is an open-and-shut evidential problem for composite-free theories. Finally, Lowe (2003, 2005) gives an argument against composite-free theories, where one interpretation of this argument is that such theories which do not appeal to properties like *mass* and *momentum* trivially cannot explain what ordinary physics theories explain because they do not appeal to the exact properties of *mass* and *momentum* as such. For Lowe, the structure of mass and momentum as such are empirically significant.

[11] Cf. Bagwell (2021).

In each of these examples, it is natural to think that philosophers are committed to the empirical significance of some thick metaphysical content. They are giving arguments which proceed from considerations of empirical evidence to some conclusion about purely metaphysical matters.[12]

3.1.3 The Trivial Response

Let's return to the trivial response. Hofweber claims that our empirical evidence favors the existence of composite objects. Here we will examine how this claim entails a rejection of the CODA and unravel its commitments.

We focus on the claim of the CODA that ordinary and composite-free theories are empirically equivalent. The trivial response reasons as follows. Among the empirical consequences of R is that *this iron bar* rusted, whereas R^{-} has no such empirical consequence. R^{-} has the empirical consequence (b^{-}):

(b^{-}) These simples arranged iron bar-wise rusted'.

But this is not the *same* empirical consequence, according to the trivial response. Even if we can construct a composite-free variant in the way provided above, and bypass any other objections, Hofweber thinks the predictions and observations of the two theories are trivially different. This is because he claims that our empirical evidence favors a particular mereological picture. If our empirical evidence supports the existence of composite objects, then trivially any empirical evidence that does not support the existence of composite objects is not the *same* empirical evidence. For if our observations confirm (b) but not (b^{-}), then (b) and (b^{-}) are distinct empirical consequences. The trivial response thus rejects

[12] See Bailey and Brenner (2020) for additional, similar examples.

the claims of the CODA that R and R^{-} are empirically equivalent. It is trivially impossible to provide an empirically equivalent, but composite-free, alternative to any ordinary scientific theory, says the trivial response.

Hofweber's trivial response requires a particular thesis about the individuation conditions of empirical consequences (as does any other variety of the empirical significance thesis like those outlined in §2.2). The trivial response succeeds only if we can empirically distinguish between R and R^{-}, and presupposes a position on how finely empirical consequences are individuated. In particular, the only difference Hofweber points to as a distinguishing feature between (b) and (b^{-}) is the thick metaphysical content—the underlying mereological picture associated with each empirical consequence.

Let us formally define this thesis on empirical consequence individuation. For any theories T_A and T_B, where T_A has the empirical consequences (a_1), (a_2), ... and T_B has the empirical consequences (b_1), (b_2), ..., the trivial response is committed to the following:

Fine Grained If the underlying mereological picture associated with (a_1), (a_2), ... is different from the underlying mereological picture associated with (b_1), (b_2), ..., then T_A and T_B are not empirically equivalent.

Fine Grained entails that empirical consequences may be individuated by the particular mereological pictures associated with those empirical consequences; in this sense, it is a *fine-grained* understanding of the individuation conditions of empirical consequences. (b) paints a picture where there is a composite object, an iron bar, that behaved in a certain manner; it rusted. (b^{-}) is not associated such a picture; instead, is is only associated with there being simples in arrangements that rusted'. Because these two empirical consequences are associated with different mereological pictures—i.e., they have different thick mereological content—it follows from Fine Grained that they are not the same empirical consequence. Thus the trivialist can give a simple argument for the empirical inequivalence of R and R^{-}.

If Fine Grained is correct, then the CODA is trivially unsound. The observations and predictions of ordinary scientific theories are about, and thereby appeal to, composites like planets, iron bars, and organisms. These empirical consequences present a particular mereological picture—one where there *are* planets, iron bars, and organisms. And were we to construct theories with empirical consequences associated with different mereological pictures—no planets, iron bars, nor organisms as such—then Fine Grained entails that these empirical consequences are necessarily distinct from those of ordinary scientific theories. Because the CODA attempts to achieve precisely this, the claim that composite-free theories are empirically equivalent to ordinary scientific theories is trivially false if Fine Grained is true.

3.2 Rejecting Fine Grained

The trivial response claims that composite-free theories are trivially empirically inequivalent to their ordinary counterparts. This is because it presupposes Fine Grained, and Fine Grained entails such empirical inequivalence. To me, Fine Grained is neither obviously true nor obviously false. To adjudicate the dispute about empirical equivalence, we must look into theories of empirical consequences that would vindicate Fine Grained. If we find that there is no good, obvious way to vindicate Fine Grained, then we can tentatively reject the trivial response to the CODA. I argue that we find this and more: any candidate theory of empirical consequences will not vindicate Fine Grained. Accordingly, we ought to reject the trivialist response to the CODA.

In this section, I will propose and reject two theories of the individuation conditions of empirical consequences that entail Fine Grained. Because we are determining standards of empirical equivalence, we cannot rely upon *intuitions* of empirical equivalence. Otherwise we are at the stalemate indicated at the outset. Instead, I rely on the possibility of successful *dispensing*, and I argue that from these cases of dispensing we can infer facts about empirical equivalence. This is because a dispensing theory, per the standard account of dispensability, must be empirically equivalent with the original theory. The basic idea of my argument is that Fine Grained precludes the possibility of dispensing with entities that might be dispensable.

3.2.1 Semantic Individuation

The most straightforward way to vindicate Fine Grained is a theory of empirical consequences where they are individuated according to their semantic content. Though many would find this independently implausible, it is instructive to see why it fails. Here is Semantic Individuation:

Semantic Individuation Empirical consequences (c_1) and (c_2) are the same empirical consequence if and only if the expressions of (c_1) and (c_2) have the same semantic content.

With Semantic Individuation, we have a test for whether two empirical consequences are the same—and, accordingly, a test for whether two theories have the same empirical consequences. Namely, whether their expressions are synonymous.

We can show how Semantic Individuation will entail Fine Grained. Recall R and R^{-} and the expressions of their empirical consequences in (b) and (b^{-}). Under any usual standard of synonymy, (b) and (b^{-}) are not synonymous. Given Semantic Individuation, (b) and (b^{-}) are not the same empirical consequences. Moreover, there are no empirical consequences of R^{-} that are synonymous with (b), since R^{-} is explicitly formulated in terms that are not synonymous with composite-terms. Accordingly, R and R^{-} are empirically inequivalent. More generally, semantic content itself is finely discriminating, and so individuation according to semantic content will be finely discriminating too. This is how Semantic Individuation entails Fine Grained.

But we have reason to think that Semantic Individuation is independently problematic: it entails the trivial impossibility of dispensing with entities that are intuitively dispensable.

Consider an alternate history of astronomy. As we currently think of them, constellations are apparent groupings of stars seen only according to the Earth's relative position to them. For the astronomer, there aren't constellations; there are only stars. But suppose that modern astronomers, after they fully understood that stars are Sun-like entities that are often light-years away from each other, nonetheless still thought that stars sometimes formed *a constellation*. And suppose they still believed in the independent existence of constellations for no good reason—just a superstitious holdover from antiquity. Semantic Individuation entails that we could not show that constellations are dispensable. This is because Semantic Individuation bars the possibility that constellation-free theories could have the same empirical consequences.

For consider what a dispensability argument would look like. First, we would take the current alternate theory of astronomy C that has the following empirical consequence, referring to a particular constellation (together with some background conditions):

(*c***)** This constellation is visible in August.

Suppose someone thought that constellations are dispensable to our theories. This constellation dispenser might suggest that we adopt the constellation-free theory C^- that has a different empirical consequence.

(c^-**)** These stars are visible in August

C^- seems to dispense with constellations. We can explain everything we want in astronomy without appealing to constellations if we accept C^- instead of C. But if Semantic Individuation is true, then because (c) and (c^-) are not synonymous—and because there is no constellation-free empirical consequence that is synonymous with (c)—the constellation-free theory is empirically inequivalent to the alternate astronomy theory.

This is a problem for Semantic Individuation. Constellations *are* dispensable to C. And they're dispensable in the usual way, where we conduct an *a priori* investigation into what parts of a theory are necessary to explain what the theory explains, and we realize that constellations are just a vestigial aspect of C. So, constellations are dispensable to C, and we can provide a constellation-free theory C^-, but Semantic Individuation entails that the C^- is trivially empirically inequivalent to C. Thus, because empirical equivalence is a necessary condition for a dispensing theory, and because we have a dispensing theory, there is a counterexample to Semantic Individuation and we ought to reject it.

It seems as though Semantic Individuation is *too* fine-grained. Note here that we are not consulting our intuitions about whether (c) and (c^{-}) are the same empirical consequence. The argument is more general than that. The argument is that Semantic Individuation entails that *no* entity is dispensable if that entity appears in our empirical consequences. For to dispense with that entity, we must provide an alternative that has synonymous empirical consequences. And if the empirical consequences are synonymous then we haven't dispensed with the entity. C is just an illustrative instance of the restrictions that Semantic Individuation places on which entities are possibly dispensable. Semantic Individuation implausibly entails that no entity which appears in our empirical consequences is dispensable. Accordingly, we should reject it for placing overly restrictive boundaries on the kinds of entities that can be dispensed with.[13]

3.2.2 Representational Individuation

One might naturally think that empirical consequences are individuated according to their representational contents. Scientists approach the external world through their perceptual capabilities, and it is reasonable to think that the empirical consequences of scientific theories are (at least informed by) the deliverances of our perceptual systems; call these representational contents. Representational contents are data about the world that are present to an organism. When the representational contents of two empirical consequences are different, that seems to be the mark of individuation we are seeking.

Here we explore such a theory of the individuation conditions of empirical consequences, which I will call Thick Representational Individuation (TRI). The

[13]Thanks to an anonymous referee from *Synthese* for helping me clarify the import of this argument.

basic idea is that, with the addition of some theses about humans' representational contents, there are ways to distinguish between the empirical consequences of ordinary scientific theories and their object-free variants because of the differences in representational content. In §3.4, I argue that TRI fails because of its adherence to one of the additional theses about the representational contents of humans.

One small caveat about my approach. A representational theory of individuation claims that empirical consequences are individuated according to the representational contents of those empirical consequences. This articulation leaves open the question of how representational contents are related to empirical consequences. It is unspecified whether empirical consequences are *constituted* by representational contents, or *partially constituted* by representational contents, or simply *informed* by representational contents. It is consistent with the present theory that there is more to empirical consequences than representational contents. The only commitment is that differences in representational contents is sufficient for differences in empirical consequences. I will refer to *an empirical consequence's representational contents* as a stand-in for whatever particular relationship one wishes to commit to.

Many philosophers have argued that the representational contents of perception, when that perception is veridical, bears a non-representational or external relationship to the thing being represented.[14] Call this thesis External:

External All veridical representational contents bear external, non-representational relationships to what is being represented.

[14]This conception of the contents of perception includes any externalist theory. See, e.g., Dretske (1997) and Stalnaker (2003).

The idea here is that in order to account for a variety of phenomena regarding perception, it must be that veridical representations bear non-purely-representational relations to the things being represented. (Hereafter, I will drop the word 'veridically', and unless otherwise specified, representational contents are veridical.) There must always be some external or worldly relation between representation and represented. Exactly what relation is disputed. Some think that the representational content of a perception of some tree is *constituted*, in part, by the tree (e.g., Fish (2009), Johnston (2004)); others think that this same content is *causally explained* by the tree (e.g., Burge (2010)), perhaps together with the evolutionary conditions of the perceiving creature. External is similar to externalist theories of semantic content, where cases of successful reference require a non-semantic relation between the referring term and the thing referenced. (Consider here Kripke's famous causal theory of reference (Kripke, 1981).)

The first step of a Representational Individuation theory is thus an adherence to the thesis of External. The idea is that the empirical consequences of a scientific theory bear a non-representational relation to the world—that, for example, the actual iron bar in (b) explains, in some way, the representational content of (b). Here, then, is a first pass at a Representational Individuation theory:

Representational Individuation Empirical consequences (c_1) and (c_2) are the same empirical consequence if and only if the representational contents of (c_1) and (c_2) are the same, where representational contents are External.

When empirical consequences differ, this is sometimes explainable by differences in the way the world is. That this particular bar of metal rusted in the presence of moisture-rich air partially explains the empirical consequence (b). Consider the empirical consequence ($\neg b$).

($\neg b$) This iron bar did not rust.

On this understanding of empirical consequences, ($\neg b$) is distinguishable from (b) because the representational contents of ($\neg b$) and of (b) are distinct. We can, for example, *perceive* whether or not the iron bar rusted. Moreover, we can chalk this difference in representational contents up to a difference in the way the world actually is—in one case, the bar of metal rusts, and in the other, it fails to rust.

The claim that representational contents are External is an elegant and powerful understanding of empirical consequences, and one which, I will suggest, might vindicate Hofweber's trivial response.

3.2.3 Thick Representational Individuation (TRI)

Representational Individuation as presented is not enough to entail Fine Grained. In particular, it is unclear whether the representational contents of perception are *robust* enough to distinguish between an iron bar and simples arranged iron bar-wise. Here, we examine a thesis that, when paired with Representational Individuation, will entail Fine Grained. This thesis is Thick:

Thick Representational contents present an underlying mereological picture of what is being represented.[15]

According to Thick, our representational contents are mereologically detailed. It claims that representational contents present a mereological picture of the thing represented. The idea here is that we can distinguish between representational contents where the only difference between two contents is a difference

[15] Thick should strictly be read as follows: All (or all relevant) empirically-relevant representational contents present an underlying mereological picture of what is being represented. We will leave Thick quantifier-less in the main body because this precise articulation adds complications about what Hofweber is committed to.

in their mereological pictures. This is because, per Thick, our representational contents actually present these two as having distinct metaphysical pictures. If Thick is true, we will only veridically perceive something *as* a composite object if it is a composite object. And when presented with a genuine composite object, so long as there are no perceptual errors, we will perceive it *as* a composite object. Our perceptual capacities have bequeathed us with the ability to not be fooled into representing simples arranged object-wise as an object, nor representing a composite object as simples arranged object-wise (so long as there are no errors). We can *perceptually* distinguish between the two. Byrne (2019), for example, holds this view.

Some may find Thick evolutionarily implausible—it seems like there is no evolutionary reason for an organism's perceptual capacities to be able to distinguish between simples in arrangements and composites. Others simply disagree with Thick, and claim that our representational content "would be the same whether or not the atoms arranged . . . [object]-wise composed something" (Merricks, 2001, 9).[16] But for the sake of argument, we will for now grant Thick alongside External.

These two features of empirical consequences and representational contents yield the following individuation conditions for empirical consequences:

Thick Representational Individuation Empirical consequences (c_1) and (c_2) are the same empirical consequence if and only if the representational content of (c_1) and (c_2) are the same, where representational contents are Thick and External.

TRI says that two empirical consequences are distinct when they have dis-

[16] See also Korman (2014, 4) and Thomasson (2014, 16, 157).

tinct representational contents, where representational contents are External and Thick. In short, the idea here is that representational contents are quite representationally dense—not only are we presented with some coarse-grained information about the world, but also with an underlying mereological picture. Thick content, in other words, is part of the total representational content of an empirical consequence.

TRI entails Fine Grained straightforwardly. If TRI is true, then empirical consequences are individuated according to their representational contents. According to Thick, representational contents present an underlying mereological picture of what is being represented. If this is so, then empirical consequences are finely mereologically individuating. Our empirical consequences can distinguish between those cases where the only difference is the mereological facts. Here we finally have a theory of individuation conditions that vindicates Fine Grained.

Here is how a trivialist would use TRI to argue against the CODA. The representational content of (b) is that some iron bar behaved in a certain way. This representational content presents a thick metaphysical picture where the iron bar is represented *as* an iron bar (and not simply simples arranged iron bar-wise). The representational content of (b^-), on the other hand, is that some simples arranged iron bar-wise behaved in a certain way. Thus, if Thick is true, we can *perceptually* distinguish between the empirical consequences of R and those of R^-. Per TRI, because we can perceptually distinguish between (b) and (b^-), they are not the same empirical consequence. Consequently, R and R^- are trivially empirically inequivalent. It is a consequence of the theory of TRI that the CODA is unsound, independent of any of the other problems with the dispensability argument. Note that Thick is the crucial premise here; in order to vindicate the trivial response, it must be that our representational contents present a mereological picture.

3.2.4 Against Thick Representational Individuation

Again, we will reject TRI because it precludes the possibility of dispensing with entities that might be dispensable. We begin the counterexample with a fact about humans. We perceive faces as faces—we do not *judge* an array of features to be a face, but instead perceive it that way.[17] In this way, humans actually have Thick representational contents when faces are involved; we recognize not only *arrangements* of facial features, but a face. This is, at least to some extent, a biological capacity, but there is some debate over the extent to which it is learned.[18]

It is a rather mundane mereological picture that is presented when we represent something as a face rather than as arrangements of facial features, but it is a mereological picture nevertheless. Moreover, here we are granting to the trivial response that there *are* cases where we perceptually distinguish things according to their underlying mereological picture.

We continue with another alternate history of astronomy. Suppose that in the night sky in August, there were some stars that looked exactly like a human woman's face—call this Phoebe's Face. When scientists look into the night sky, they seem to perceive Phoebe's face in the arrangement of stars. Their representational contents are actually as of Phoebe's Face, not merely some stars arranged face-wise. The perceptual evidence is so convincing, and the image so detailed,

[17] There is empirical evidence that suggests that this is a perceptual capacity—even if it is slightly informed by culture. Individuals can perceive faces in complex pictures or scenes in a short enough time frame that there is no time for substantial cognitive influence—that is, they didn't think before *seeing* the faces. See VanRullen and Koch (2003).

[18] Siegel (2010) argues that it is largely learned—that we can perceive doubt on a person's face if we know that person well enough and know when they doubt something. Block (2014) is unconvinced.

that everyone in this alternative history genuinely believes that Phoebe's Face exists. The theories of astronomy (together with background conditions) might well contain the following empirical consequence.

(*p*) Phoebe's Face is visible in August.

(*p*) is an empirical consequence of this alternate theory of astronomy P. And if TRI is correct, then (*p*) is committed to Phoebe's Face being a *face*, much like how (*b*) commits to that iron bar being a *composed* iron bar.

Suppose, though, that a scientist was conducting *a priori* investigations into whether there are any parts of our theories that are dispensable, and they nominate Phoebe's Face. After all, they think, we know that humans have a proclivity for seeing things as faces, and sometimes it is hyperactive. Maybe, then, our perceptions of Phoebe's Face are illusory in a sense. It seems that Phoebe's Face is dispensable to P; we can explain everything by appealing to arrangements of stars that seem to look like Phoebe's Face.

Accordingly, the would-be Phoebe's Face dispenser would construct variants of all astronomy theories without appealing to Phoebe's Face. Among the empirical consequences of a variant P^- would be the following:

(p^-) Those particular stars arranged face-wise are visible in August.

In (p^-), the particular stars referenced are those that "make up" Phoebe's Face. P^- shows that Phoebe's Face is dispensable to P. We do not need to appeal to a face in order to explain all of what P explains, since P^- can explain everything just as well.

If Phoebe's Face is dispensable to P in the way just described, then according to the standard account of dispensability, P^- is empirically equivalent to P.

And that would signify that (p) and (p^-) are the same empirical consequence. However, per TRI, (p) and (p^-) are *not* the same empirical consequence.

This is because Thick discriminates between different mereological pictures. And as we know, humans perceive faces *as* faces. Because of this, the representational contents of *a face* and of *Xs arranged face-wise* are distinct. We do not represent something as a face-like arrangement, but as a face. Since the representational contents of (p) present something as a face, and the representational contents of (p^-) present some things as looking like a face, these representational contents are distinct. Compare this to the case of an iron bar and simples arranged iron bar-wise. Thick entails that these two are distinct: the representational contents when perceiving an iron bar, if Thick is true, present the thing perceived *as* an iron bar. The representational contents when perceiving just simples arranged iron bar-wise, on Thick, do not present the things perceived as an iron bar. It is the same here as with Phoebe's Face. So, the same considerations that allow TRI to vindicate Fine Grained show that (p) and (p^-) are distinct empirical consequences.

However, this is a problem. Phoebe's Face *is* dispensable, and we can show this by offering P^-. And yet TRI, because it entails that the empirical consequences of a Phoebe's Face-free astronomy theory are trivially distinct from the ordinary theory, tells us that Phoebe's Face is not dispensable. So, TRI ought to be rejected. Once again, the problem is not merely that P seems empirically equivalent to P^-. Rather, the problem is that TRI entails the impossibility of dispensing with mereologically-rich entities that appear in our empirical consequences. If there is a mereologically-rich entity that is doing no genuine explanatory work in our theory, but we can perceptually distinguish that theory from one that is identical except it lacks the mereologically-rich entity, then TRI entails that the entity is not dispensable. But by hypothesis the entity is not doing any genuine explanatory work. Any entity that appears in our theories that is not genuinely explanatory should be dispensable, but TRI precludes the possibility of dispensing with some such entities. That is why we reject it.

3.2.5 Rejecting Trivial Responses

The trivial response to the CODA, as it stands, cannot be defended. For, to do so, one must provide a theory of the individuation conditions that will entail the trivial difference between the empirical consequences of object-free and ordinary theories. And no good theory of individuation conditions vindicates Fine Grained. A theory where semantic differences individuate empirical consequences will not do, nor will a theory where thick representational content differences individuate empirical consequences. These theories fail because there are simple cases of dispensing that are not possible by fiat.

The thrust of my argument against Hofweber's response is that there might be cases where some composed entity that appears in empirical consequences is dispensable, and any theory of individuation that vindicates Fine Grained trivially rules these cases out. This is because Fine Grained entails that any two theories with empirical consequences that have different thick mereological content cannot be empirically equivalent. Fine Grained, then, precludes the *possibility* of dispensing with a composed entity that appears anywhere in a theory's empirical consequences, and this seems absurd. Surely there could be a dispensable composed entity in our empirical consequences. So we must reject Fine Grained.

We can generalize my objection to other philosophers who adhere to the position that thick metaphysical content is empirically significant. The mereological version of this position presupposes that empirical consequences can be individuated along mereological lines. But if it is in principle *possible* to dispense with some mereologically-rich entity in our empirical consequences, e.g., Phoebe's Face, then it is false that empirical consequences can be individuated along mereological lines. *Mutatis mutandis* for other versions of this position: A different version presupposes that empirical consequences can be adjudicated along some particular metaphysical line. But if it is in principle possible to dispense with some metaphysically-rich entity in our empirical consequences, then it is false that empirical consequences can be individuated along such metaphysical lines. I have not yet given counterexamples for non-mereological versions of Fine Grained, but it is intuitive that for any variety of this position, there will be an analogue to Phoebe's Face. All we need is a possible scenario where we have mistakenly inferred that some thick metaphysical content in an empirical consequence is significant and where our theories are just as good when we rid our theories of that metaphysical content.

We should reject Hofweber's trivial response to the CODA because there is no good theory of empirical consequences' individuation conditions that could serve to show that ordinary and composite-free theories are trivially empirically inequivalent. We can also give similar arguments for any other variety of the position that thick metaphysical content is empirically significant. Accordingly, I have made a case for evidential quietism: our empirical evidence is radically silent on distinctively metaphysical disputes.

3.3 Upshots for full equivalence

Let us consider in more detail the position that thick content is scientifically significant. I argued that thick content is not *empirically* significant. Here I make a tentative case that thick content is sometimes not scientifically significant at all.

In the literature on theoretical equivalence, which examines the conditions under which two theories are fully equivalent, there are philosophers who endorse a position similar to Hofweber's. These philosophers seem to presuppose that thick metaphysical content, as it appears in *non*-empirical consequences, is scientifically significant. There are cases where the only relevant difference between two theories is a difference of thick content, and philosophers claim that such a difference suffices for those two theories being inequivalent. Here I will present one recent instance and suggest that we should be careful about concluding that some thick metaphysical difference is scientifically significant.

Usually, philosophers in the full equivalence literature use equivalence as a means to understanding the significant content of a scientific theory (Barrett (2019, 1186 - 1192), Weatherall (2019a, §5)). The idea seems to be that when we have a good understanding of when two theories are equivalent, we can better gauge which parts of those theories should be taken literally from a scientific realist perspective. Here, I approach from the reverse end. I am arguing that we can learn some things about *equivalence* by examining which parts of our theories are representationally significant. We already have some reason to think that thick content is not empirically significant, and I aim to show that in some cases, thick content is not at all significant. Moreover, recall above that I examined whether thick content was empirically significant by using an alternative notion, dispensability, as a proxy. For full equivalence, there is no analogous proxy to judge whether some content is significant. Accordingly, the conclusions reached here are more tentative than the preceding, but I hope to show that there is headway to be made by approaching full equivalence from this angle.

Consider Hamiltonian and Lagrangian formulations of classical mechanics. It is commonly held among physicists and philosophers of physics that these two formulations are equivalent: the two theories say the same thing about the world and are mere notational variants. Recently, North (2009, 2021b) has pushed back against this received view. She argues that there are differences between the structures of these two theories that are significant in the sense that they show that the theories are in fact *in*equivalent. This argument has generated much discussion, and has led many philosophers to discuss the conditions under which two theories are equivalent in general.

Let me simplify the debate. Briefly, in Hamiltonian mechanics, the state of a classical physical system is specified by the particles' positions and momentum, whereas in Lagrangian mechanics, the state of a classical physical system is specified by the particles' positions and *velocity*. Lagrangian state-spaces have metric structure, whereas Hamiltonian have merely symplectic structure. This difference in structure, North argues, calls into question the equivalence between the two formulations. North claims that because these two theories do not have a structure-preserving mapping between them—the structure of one is literally not present in the other—they are not equivalent.

North claims that this structural difference is significant; it is enough to show that these theories say different things about the world from the scientific perspective. North's argument that this apparent structural difference is sufficient for inequivalence requires an answer to the following question about equivalence in general: in virtue of what can difference in structure be sufficient for inequivalence? Or, as she says, "The question is whether they are equivalent, full stop. The answer depends on whether what differences there are *matter* in any way" (North, 2021a). We need some measure for when an apparent difference in structure is significant. Surely there are cases where structural differences are sufficient for inequivalence; for example, one difference between Newtonian and Galilean spacetime is the structure of absolute rest, and this difference seems significant. On the other hand, the same theory formulated in two different natural languages, say French and English, leads to two theories that have *some* structural difference, though we think that this is a case of mere notational variance. As a result, we need some answers to what kinds (or degrees) of structural differences are significant. Regarding the question of whether what differences there are matter, North might answer that even thick structural differences matter.

The difference between Hamiltonian and Lagrangian mechanics seems to be a merely thick structural difference. One formulation presents an underlying structure that is distinct from the other's; that one theory says a system is specified by momentum and the other velocity is strictly speaking a different underlying structural picture. The structural difference between these two theories is merely in their thick structural content. This is because, in most normal cases, a metric structure can be "recovered" within Hamiltonian mechanics. (More precisely, there is a mapping from any hyperregular model of Lagrangian mechanics to a model of Hamiltonian, and *vice versa*. Cf. Barrett (2019).) If so, then while North is correct that there is not a structure-preserving mapping between the two theories, the difference between the structures of the two formulations is merely in the thick structural content. Hamiltonian mechanics is not committed to metric structure *as such*, but the role that metric structure plays within Lagrangian mechanics can be recovered within Hamiltonian.

The present interpretation of North's argument is that she is pointing to a *thick* structural difference between the two formulations and presupposing that even thick structural differences are enough for inequivalence. (North might want to push back against this interpretation, but a natural reading of her argument presents such an interpretation.) I will suggest that this answer to the above question, that even thick structural differences matter, leads to problems similarly to how Hofweber's presupposition that thick content is empirically significant leads to problems. Consequently, North must offer a precise account of when differences in structure entail inequivalence; otherwise, her objection to the standard view cannot stand.

Let us consider a case of theories which have a difference in merely thick structural content but intuitively this difference is not significant. If there is such a case, then merely thick structural differences are not sufficient for inequivalence. Consider the theory of linear orders.[19] We can formulate the theory using the concept of a nonstrict order *less than or equal to*, signified the binary predicate $\leq$, or we can formulate it using the concept of a strict order *less than*, signified by the binary predicate $<$. These formulations have different axioms; for instance, the first has the axiom that everything bears $\leq$ to itself, whereas the latter has the axiom that nothing bears $<$ to itself. There is a difference in thick structure between these two formulations. The first avails itself of the property *is less than or equal to*, whereas the latter avails itself to the property *is less than*. The sec-

[19]Cf. Winnie (1986), Barrett (2020a, 1187).

ond, strictly speaking, does not appeal to *less than or equal to* as such. This is a thick structural difference. Moreover, there is no structure-preserving mapping between them since they trivially and explicitly have different structures.

However, there's an obvious sense in which these are equivalent formulations of linear orders. They both ascribe, in some sense, the same structure to sets. We take it that the difference between formulating linear orders with either $\leq$ or $<$ is not a significant difference; it does not make a difference to the content of the theory. This is a *mere* thick structural difference that we ought not interpret as scientifically significant. We noted in §3 that there are many cases where thick differences are not scientifically significant; in particular, we cannot appeal to thick differences to conclude that two theories are empirically inequivalent. Likewise, here we cannot point to the thick difference between these two formulations of linear orders as significant; just because there is a difference in the thick structural content, we cannot infer that the two formulations are inequivalent. I take it that one can consult their intuitions to tell that the difference between *less than* and *less than or equal to* does not matter in the scientific sense. Moreover, we can easily recover the structure of $\leq$ on the theory that only has $<$ and *vice versa*.

Here we have a case where, intuitively, thick structural differences are scientifically *insignificant*. Some thick structural differences do not themselves entail that two theories are theoretically inequivalent. If so, then North's claim that the structures of Hamiltonian and Lagrangian mechanics are sufficient for their inequivalence cannot hinge only on there being a thick structural difference between the two.

This argument is not meant to be conclusive.[20] If one truly wishes to dis-

[20]The present paper had been finished by the time North published her most recent book, North (2021b). Much of what has been said in this section would have changed in light of her

tinguish between Hamiltonian and Lagrangian mechanics on the basis of thick structural differences, then they may simply accept that these two formulations of linear orders are inequivalent and that even thick structural differences matter. My argument is meant to show simply that North's presupposition about the significance of structural differences, though seemingly innocuous, leads to unpalatable consequences. It seems unintuitive to count all thick structural content as significant, since it leads us to conclude that many theories which we take (or should take) to be equivalent are trivially inequivalent. Here is the lesson we should learn: North claims that there is a significant difference between Hamiltonian and Lagrangian mechanics which entails that the two formulations are inequivalent. One way to articulate the difference between Hamiltonian and Lagrangian mechanics is a difference in thick structural content, in the underlying structural picture associated with each formulation. The latter appeals to metric structure and the former merely symplectic structure. Yet *thick* structural differences between two theories seem not to be sufficient for inequivalence. We have already seen that two theories which differ only in thick content might still be empirically equivalent. Likewise, we should think that two theories which differ only in thick content might still be fully equivalent. Moreover, we have good rea-

Chapter 7 of that book. In particular, North has provided a more general explanation of what is happening between Hamiltonian and Lagrangian mechanics. In that chapter, North considers a theory's "picture of the world", which corresponds directly with what I call a theory's thick content. She identifies multiple pairs of theories that she takes to be *informationally* equivalent but *metaphysically* inequivalent; these theories disagree on "what there is, what it is like, and how and why it behaves in certain ways to give rise to what we observe" (North, 2021b, 196). She clarifies that the difference between Hamiltonian and Lagrangian mechanics is a difference between the metaphysical pictures they present (North (2021b, 224)). North suggests that metaphysical inequivalence is sufficient for theoretical inequivalence. In this we agree, though—as has become clear throughout this paper—it is not always clear when two theories *are* metaphysically inequivalent. For we need some measure of when two theories make different metaphysical claims about the world. It seems to me that the example about linear orders is *not* a case of metaphysically inequivalent theories, though they explicitly differ in structure. Still, then, North must answer whether what differences there are matter.

son to think that there *are* theories which differ in their thick structural content but are fully equivalent. In order to conclude that Hamiltonian and Lagrangian mechanics are inequivalent, North must provide a difference-maker between the two that is more than a mere thick structural difference.

3.4 Conclusion

What I have shown is this: First, that many philosophers adhere to some thesis that thick metaphysical content is empirically significant, and that any variety of this thesis requires an indefensible presupposition about the individuation conditions of our empirical consequences. Second, that the trivial response to the CODA is not tenable. It is not permissible to infer empirical inequivalence from differences in the underlying metaphysical pictures of two theories. Third, that these arguments have some purchase in the theoretical equivalence literature. We should be suspicious of philosophers who argue for inequivalence solely on the basis of differences among thick metaphysical content. There is much more work to be done on determining when the consequences of theories are the same or different, and I have shown that there are implications across philosophy of science and metaphysics for these determinations.

Chapter 4

No Science without Composites

Philosophers have recently debated whether a responsible metaphysics of science requires that we ontologically commit to composite objects like minerals, diamonds, corks, and stars. In this chapter, I argue it does.

4.1 Preliminaries

I will provide an indispensability argument for the existence of composites that relies on a widely held principle about metaphysical commitment. This is not the first composite indispensability argument in the literature. I hope to show, though, that it should be taken seriously by philosophers of science and scientifically inclined metaphysicians.

Indispensability and dispensability arguments infer some claim about the existence of an entity (or something having some structure) from the formulations of our best scientific theories. These arguments usually appeal to general principles like the following:

Dispensability. If an entity or structure X is dispensable to all of our best scientific theories, then we (defeasibly) ought not commit to Xs (or to anything having structure X).[1]

Indispensability. If an entity or structure X is indispensable to some of our best

[1] For some appeals to this principle, see Friedman (1983, p.112), Colyvan (2001, Ch. 1), Dorr (2002), Norton (2003), North (2009, 2021b), Brenner (2018), and Barrett (2020a).

> scientific theories, then we ought to commit to Xs (or to something having structure X).[2]

The basic idea is that we should metaphysically commit to the things necessary (i.e., indispensable) for the articulation of some theory and not commit to the things that are not. Plausibly, these principles are justified by some variety of philosophical naturalism or scientific realism. For this paper, I assume both, though there are debates about how to justify them.

A relevant question, given the principles, is whether composites count among the indispensable entities. Some have given arguments that indicate an affirmative answer.[3] These arguments typically rely on taking the content of our scientific theories at face value, where reference to composites in parts of our theories suggests that we cannot dispense with them from those parts of our theories (cf. Hofweber (2016, p. 199)). But these have recently been objected to on the grounds that they overgenerate cases of indispensability (LeBrun (2021)).

Others might offer a Science Without Composites (SWC), arguing that composites are dispensable to our best scientific theories.[4] SWC provides a procedure for generating theories that dispense with composites:

> Take one of our best scientific theories T. Replace all appeal to some composite X in T with appeal to mereological simples arranged X-wise and all appeal to some singular property P in T realized by com-

[2] For some appeals to this principle, see Melia (2000, 474 - 475), Field (2016, 43), (Field, 1989, 15), Colyvan (2001, Ch. 1), Rosen and Dorr (2002), and Dorr (2010, §4).

[3] See Williamson (2007, p. 223), Hofweber (2016, p. 199), Hofweber (2018, pp. 321-322), Byrne (2019), and Byrne and Manzotti (2022).

[4] For something close to a composite dispensability argument, see van Inwagen (1990, §§9 - 11), Merricks (2001, 2022); Dorr (2002); Rosen and Dorr (2002); Sider (2013); Brenner (2018). These philosophers might not describe their arguments *as* dispensability arguments (since they may not endorse the Dispensability principle), but nonetheless one can charitably interpret some of their arguments as composite dispensability arguments. Brenner (2018, pp. 661 - 664) makes the most explicit of these dispensability arguments.

posites with appeal to a collective property P^- realized by simples arranged composite-wise, and make no further changes.

Philosophers are likely familiar with the procedure of paraphrasing away composite talk in favor of simples arranged X-wise talk, but they may not see why SWC must paraphrase singular property talk. It will be useful to walk through this carefully. There are scientific properties for which the SWC proponent need not replace. Consider, e.g., orbiting the Earth. Both a composite and an arrangement of simples can orbit the Earth. SWC need not paraphrase away these properties.

But, plausibly, there are other scientific properties for which it would be false (or otherwise incoherent) to ascribe them to a plurality of simples. For instance, diachronic properties like rusting. Something rusts when, over a period of time, that very thing undergoes a particular chemical reaction. Any time we are inclined to say that some iron bar rusted, it will not be true to say that some particular arrangement of simples underwent that very process. Similarly, some have claimed that arrangements of simples don't have the same mass, velocity, momentum, shape, size, structure, color, chemical reactivity, and tidal pull properties as their ordinary composite counterparts.[5] For example, whereas it would be true to say that some ball is red all over, it would be false or incoherent to say, e.g., that the simples arranged ball-wise are red all over. Instead, SWC proponents will introduce a new property *red*$^-$ that is had collectively by simples. This

[5]See Baker (2003); Lowe (2003); Merricks (2003). Eklund (2005) suggests (without committing) that simples in arrangements cannot have the property of *happiness* but only schmappiness (which in my terminology would be happiness$^-$). More recently, Contessa (2014) has worried about simples performing the diachronic activities that ordinary composites do, like cats shedding fur. Just as well, Long (2019, p. 465) notes that simples arranged cat-wise cannot purr (which may or may not be a diachronic activity). There seems to be agreement that simples in arrangements cannot always have the same properties that ordinary composites generally do. See also Brenner (2015).

collective property red$^-$ is a placeholder (or shorthand) for the complex relation that simples will bear to one another when something is "red all over".

One heuristic for determining whether some property is a singular property is this: given some property P that is predicated of some putative object, if we cannot appropriately or truly predicate P of any of the mereological simples that "make up" the object, then P is a singular property. If some scientific theory appeals to a singular property, then SWC's procedure will replace appeal to it with appeal to a collective property had by simples.

Let's see SWC's theory generation procedure in action. Suppose one of our ordinary scientific theories entailed the following:

Ordinary Consequence. The iron bar rusted.

This consequence appeals to composites: iron bars. But it also appeals to a singular property that is only had by composites: rusting. SWC will transform the ordinary consequence into this:

SWC Consequence. The simples arranged iron bar-wise rusted$^-$.

According to proponents of SWC, all scientific theories that appeal to composites have replacements, formed by this process, that do not appeal to composites.

Proponents of SWC are providing a composite object dispensability argument. The crucial question for SWC is whether the theories generated by this procedure successfully dispense with composites. As we saw in Chapter 2, there is a debate over exactly how to characterize dispensability. The general idea is that we dispense with some entity or structure from a scientific theory when we can provide an alternative theory that is attractive, does not appeal to that entity or structure, and preserves the important scientific content—or in my words from Chapter 2, the *privileged content*—of the original. For SWC's strategy to be successful, every theory generated by SWC must satisfy the three conditions for dispensing with the composites that appear in the original theories. It is not clear to me that SWC theories meet the first two conditions (being sufficiently attractive and failing to appeal to composites[6]). For our purposes here, however, I will assume that SWC theories are both suitably attractive and do not appeal to composites. The present consideration, then, is whether all SWC theories satisfy the following condition on dispensing with composites:

Necessary Condition on Dispensing. An entity or structure X is dispensable to a theory T only if there is a theory T^- that preserves all of the privileged (or important) scientific content of T.

SWC is the only candidate composite dispensing theory that has been introduced in the literature. I will propose and consider two other candidate dispensing theories similar to SWC. In the end, I argue that each fails to meet this necessary condition on dispensing. As a result, we should not think that composites are dispensable to our best theories. Instead, they ought to be accorded an indispensable status for the time being, until we can assess other potential dispensing theories.

4.2 No Science without Composites

I argue that there are many scientific theories for which SWC fails to preserve important scientific content. I will here survey one. If I am right, we have good reasons to reject SWC theories.

[6]Some have argued that non-composite theories are too unattractive, and so offer a route for one to object to SWC theories on these grounds. Cf. (Parsons, 2013, pp. 332 - 333).

Before we continue, it is important to ask why this matters. Is the existence of composites on such delicate ground that one must build them a new foundation? I think yes. Generally, those who think that composites exist justify their position on the basis of perception: we *see* diamonds, corks, and stars.[7] Recently, though, philosophers have attempted to undermine this source of justification through evolutionary debunking arguments.[8] They argue that our perceptual system evolved to produce perceptual beliefs whose content is independent of which composites there are. And they claim that this entails that we should think our composite object beliefs are not accurate. If we think this argument is persuasive, as I do, then the dialectical ground has shifted. We now have no good reason to think there are any composites. However, an indispensability argument for composites is a non-perceptual source of justification for their existence. Here is a promissory note: the argument presented in this paper provides a source of justification for our beliefs about composite objects that is not within the scope of the debunking argument.

My argument proceeds by examining compositional explanations of properties. Long ago, Hooke (1665) examined a piece of cork under a microscope. He presented the explanation that pieces of cork are buoyant because the cell walls of a piece of cork are hydrophobic. Hooke paved the way for a kind of scientific explanation: the properties of natural composites are explained in terms of the properties of those composites' parts. My indispensability argument takes these

[7] Cf. Hawthorne (2006, p. 109) and Korman (2015, Ch. 4). See also the arguments found in Korman (2011) that are used to justify the various positions in material object metaphysics. Two notable exceptions are Merricks (2001), who justifies the existence of conscious composites because they have non-redundant causal powers, and Thomasson (2007), who justifies the existence of composites on the basis of the rules of use of terms like 'tree' together with ordinary observations about simples arranged tree-wise. For related issues, see Brenner (2023).

[8] For philosophers who are wrestling with this argument, see Korman (2014, 2019b,a), Osborne (2016), Kovacs (2019), Barker (2020), and Bagwell (2021).

compositional property explanations as crucial aspects of the scientific image.

Let's see this in action. Ferrimagnetism is the property of a material to have non-net-zero spontaneous magnetization. It differs from ferromagnetism and anti-ferromagnetism by always having a net-positive or net-negative spontaneous magnetization. As a result, ferrimagnetic materials are naturally magnetic—they do not need to be magnetized (unlike, e.g., nickel). Many different minerals are ferrimagnetic, including magnetite, greigite, yttrium iron garnet, barium ferrite, and others. As the name suggests, all ferrimagnetic minerals have iron atoms. But the arrangement of iron atoms within the structure, and the other atoms present in the mineral, differ with each ferrimagnetic mineral. Magnetite, for example, is comprised of iron and oxygen in octahedral crystals and opposing tetrahedral crystals. These opposing crystals give the mineral an unequal and opposite magnitude that result in non-net-zero magnetization. Greigite is comprised of iron and sulfur in a hexoctahedral crystalline structure.

The compositional explanation of ferrimagnetism, à la Hooke, is this: magnetite is ferrimagnetic because its iron atoms and oxygen atoms are opposed in particular crystal patterns. Similarly, greigite is ferrimagnetic because its iron atoms and sulphur atoms are opposed in particular (different) crystal patterns. Different microphysical structures give rise to ferrimagnetic substances; put another way, ferrimagnetism can be compositionally explained by different kinds of microphysical properties.

Here then is my argument:

(1) The property of ferrimagnetism is indispensable to some of our best scientific theories.

(2) If so, then composite objects are indispensable to some of our best scientific

theories.

(3) So, composite objects are indispensable to some of our best scientific theories.

The basic idea behind this argument is this. When we look to our best theories, we see scientists appeal to a variety of different kinds of magnetism to explain empirical phenomena. Some of these empirical phenomena are explained by appeal to some mineral being ferrimagnetic. At first glance, then, ferrimagnetism is indispensable. The second premise takes initial justification from the fact that most scientific properties are singular properties in the sense that they are not had by simples in arrangements but only by composites. If these two premises are true, then some composite objects (like magnetite and greigite minerals) are indispensable to our best scientific theories, and given the indispensability principle, we ought to commit to them. I think that, if this argument is sound, it opens the door to commit to natural composites that are given scientific compositional explanations, like Hooke's corks, as well as things like diamonds, organs, and stars.

A small clarification on (2). One could reject (2) by claiming that ferrimagnetism is a property that could be had *both* by composites and by simples in arrangements, like the property of orbiting. But recall the heuristic introduced earlier: if ferrimagnetism cannot be appropriately predicated of any of the simples that make up a magnetite mineral, then it is a singular property and SWC needs to replace it with a collective property. Ferrimagnetism arises out of unequal and opposing magnetic forces, and so we cannot truly predicate it of any simple. It is a singular property. Thus, (2) follows from ferrimagnetism's status as a singular property. The relevant question is then whether we have good scientific reasons for thinking that some things are ferrimagnetic in this sense. Proponents of SWC will argue that there are not. I disagree.

4.2.1 Multiple Realization

There is a striking structural similarity between the argument presented here and the multiple realization in philosophy of mind. There, philosophers argued that mental states were irreducible because very different kinds of creatures could each have the same mental states despite having very different brains.[9] Here, I am arguing that ferrimagnetism is irreducible because very different kinds of minerals can each be ferrimagnetic. While these arguments are similar—as are any arguments about reduction and emergence—there are a few reasons to explicitly consider explanatorily important higher-level properties and their relation to SWC.

First, we are here considering an indispensability argument in the context of the above indispensability principles. The upshot of this is that our metaphysical commitments track the structures and entities that are appealed to (or referred to) within the theories we accept to be true. Accordingly, the SWC theorist is prevented from giving two common responses. First, they cannot be an "easy roader" for the higher-level sciences. Easy roaders claim that we can refrain from ontologically committing to some entity even when it is appealed to within a theory. For example, Melia (2000) thinks we can affirm the truth of Newton's law of universal gravitation but consistently deny the existence of numbers.[10] Since we are assuming the indispensability principle, the SWC theories cannot respond

[9]For an overview of the debate, see Bickle (2020) and Polger and Shapiro (2016) and the citations in each.

[10]See also Balaguer (1998, Ch. 7), Azzouni (2012), Leng (2005, 2012), Bueno (2012), Saatsi (2011, 2016, 2017, 2020), Knowles and Saatsi (2019).

in this way. Second, an SWC theorist cannot rely on a fundamentality response. Some philosophers (e.g., Cameron (2007, 2010)) might think that ferrimagnetism is a property that is grounded in the properties of the mereological simples. So while ferrimagnetism *is* realized, and this entails that some composites like magnetite minerals are indispensable, these philosophers would claim that they are not ontologically committed to composite objects. This is because they would reject the indispensability principle in favor of one that says that we are ontologically committed to all the *fundamental* entities that are appealed to in our best theories. But given the indispensability principle as written, the SWC theorist cannot take this route.

Second, I worry somewhat about wholesale importing responses to the multiple realization debate. I am operating on a particular understanding of *realization* that is not shared by all versions of, and responses to, the multiple realization argument. In particular, I take it that, e.g., the structural properties *of iron atoms and oxygen atoms* realizes the ferrimagnetism *of a magnetite mineral*. Realization, in this particular compositional sense, is a relation between the properties of some entities and the property of a numerically distinct entity. This contrasts with some versions of the multiple realization argument, where for example Putnam claims that realization is a relation between properties of the same entity—"the organism" has some brain property and some mental property (Putnam, 1967b). Other conceptions of realization differ. Polger and Shapiro (2016, p. 22) often speak of realization as a relation between entities (brains and minds) and multiple realization as a relation between properties (brain states and mental states). But their conception of realization is also not without critics (see Aizawa (2022)). Exactly how one characterizes realization matters to this indispensability argument in a way that it may not in the mind debate.

Finally, even if one disagrees with me and thinks that this is just the multiple realization argument, it is still instructive to see why SWC does not have any immediately apparent solutions. For the SWC theorist has powerful tools for avoiding ontological commitment to some composite. They may paraphrase the consequences where the composite is appealed to or they may reject the consequence altogether. But the case of properties like ferrimagnetism pose an especially difficult problem for the SWC theorist.

4.3 Three attempts at dispensing

Here we will consider three theories that attempt to dispense with ferrimagnetism. Only the first of these three theories has been proposed in the literature. I will show that each fails on the same grounds: none can adequately preserve all of the relevant scientific content from our best theories. My aim here is not simply to respond to these objections. Rather, this exercise will reveal why properties like ferrimagnetism are relevant for the debate over the indispensability of composite objects. We will see that ferrimagnetism is scientifically important (§4.1), is not fully reducible (§4.2), and is a distinctively compositional variety of irreducible property (§4.3). My responses to each dispensing theory require us to rationally reconstruct the scientific explanations that ferrimagnetism features in and that explain ferrimagnetism.

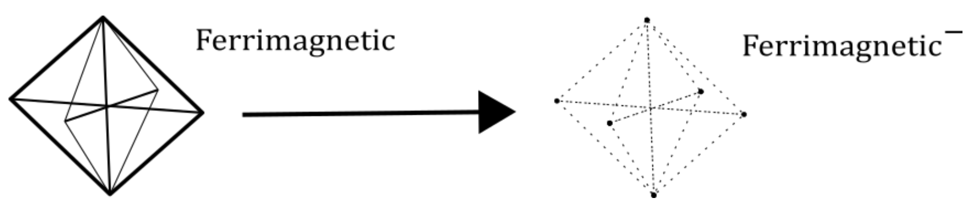
4.3.1 Rejecting (1): Eliminating Ferrimagnetism

One might object to my argument as follows:

> (1) claims that ferrimagnetism is indispensable. Only with an ultra-literal interpretation of the science might we think this. Instead, we can do away with ferrimagnetism (as a property of composites) and replace it with ferrimagnetism^{-} (a property of simples) as outlined by SWC's theory generation procedure.

I am generally sympathetic to these kinds of objections. There are many things that appear in scientific theories that apparently have metaphysical consequences. But often it is just a matter of convenience that we appeal to the properties and entities we do. Accordingly, we shouldn't simply accept the metaphysical consequences of an ultra-literal reading of our scientific theories. I will argue that *ferrimagnetism* is not like this. Ferrimagnetism is genuinely indispensable. This is because, as I will show, ferrimagnetism is a (*a*) singular property that (*b*) has multiple different manifestations at the microphysical level and (*c*) is scientifically important. If any property has these three qualities, I argue, then it is indispensable and entails the existence of composites.

Let us fill out the elimination theory from the SWC proponent first. The idea behind it is to do a simple elimination of ferrimagnetism in favor of ferrimagnetism^{-}. Consider the following diagram:



The idea is that our best scientific theories entail the image on the left, of a mineral that is ferrimagnetic. To eliminate ferrimagnetism, SWC will replace the left image with the right one, which has only mereological simples and the property of ferrimagnetism^{-}, a collective property of simples.

Let us see this in detail. Suppose our ordinary scientific theory is the following two claims:

O_1. This magnetite mineral is ferrimagnetic.

O_2. This greigite mineral is ferrimagnetic.

SWC takes ordinary scientific theories like these and *flattens* them across levels of explanation, reducing all composites to simples in arrangements and all properties of composites into properties of simples in arrangements. They face an initial problem with this theory, though. Consider the transformation of O_1. SWC replaces appeal to magnetite with appeal to simples arranged magnetite-wise, and it replaces appeal to ferrimagnetism with appeal to ferrimagnetism$^-$, a property of simples when they are arranged magnetite-wise. So, we have the following:

O_1^-. These simples arranged magnetite mineral-wise are ferrimagnetic$^-$.

Note here that *ferrimagnetic*$^-$ is a property had by simples *when they are arranged magnetite mineral-wise*. But given this, SWC cannot do the same procedure with O_2:

O_2^-. These simples arranged greigite mineral-wise are ferrimagnetic$^-$.

This second paraphrase, while it succeeds to flatten the properties and composites into simples, is false by SWC's lights. Recall that magnetite and greigite minerals have different "compositional" structures in the sense that one has iron and oxygen and the other has iron and sulphur. If so, then "ferrimagnetism" does not manifest in exactly the same way between the two. In magnetite, "ferrimagnetism" is realized by specific arrangements of iron and oxygen, and in greigite, "ferrimagnetism" is realized by different arrangements of iron and sulphur. Because of this, at the level of mereological simples, we cannot say that the realization of "ferrimagnetism" is exactly the same between the two. The simples arranged greigite mineral-wise are not *ferrimagnetic*$^{-}$, since the property of ferrimagnetism^{-} can only be had by simples when they are arranged iron and oxygen in octehedral and opposing tetrahedral crystal-wise.

SWC might then introduce a "ferrimagnetism" property that is had by simples when they are arranged greigite mineral-wise, and a different "ferrimagnetism" property that is had when simples are arranged magnetite mineral-wise. Call these g-ferrimagnetism and m-ferrimagnetism. The strategy from here is as follows. The SWC proponent should say that there is no such thing as ferrimagnetism, but there are properties of different arrangements of simples that *look just like* ferrimagnetism. When simples are arranged magnetite mineral-wise, they are m-ferrimagnetic, and when they are arranged greigite mineral-wise, they are g-ferrimagnetic. Accordingly, the proper replacements of O_1 and O_2 are these:

SWC_1**.** These simples arranged magnetite mineral-wise are m-ferrimagnetic.

SWC_2**.** These simples arranged greigite mineral-wise are g-ferrimagnetic.

SWC has a strategy for providing seemingly adequate alternatives to our ordinary scientific theory that can accommodate how different substances can be "ferrimagnetic". (For any other manifestation of "ferrimagnetism", SWC will add another property.) Accordingly, SWC claims it has dispensed with ferrimagnetism.

Consider whether this reformulation of O_1 and O_2 preserves all the important scientific content. I argued last chapter that this theory preserves the *empirical* content of the original (contra Hofweber (2016, 2018)), but I will argue that there is important non-empirical content that is not preserved. What follows is an example that shows that *ferrimagnetism*—not m-ferrimagnetism nor g-ferrimagnetism—plays an important non-empirical explanatory role in evolutionary microbiology.

Consider the following explanation in microbiology. Magnetotactic bacteria create organelles out of magnetic minerals that bias the organism's direction North-South (called magnetotaxis). They do this to travel to places with better oxygenation. For a long time, we discovered only bacteria that create these organelles out of magnetite, and scientists proposed the explanation that bacteria use magnetite because magnetite is *ferrimagnetic*. Ferrimagnetic substances retain spontaneous magnetization even after the organelle forming process, so magnetite minerals are the perfect substance for magnetotactic bacteria. We learned, in 1964, that there is a ferrimagnetic mineral that has a different compositional structure than magnetite does: greigite (Skinner et al., 1964). Greigite is used by some magnetotactic bacteria to make these magnetic organelles. Our scientific theories have concluded that what is important for the magnetotactic bacteria is not this particular substance or that, but whether the substance is ferrimagnetic.

My claim is that any adequate reconstruction of our best microbiology theories will include the explanation that these different magnetotactic bacteria species use different substances for the same process because those substances are *ferrimagnetic*. Ferrimagnetism is an indispensable part of the explanation for the phenomenon that magnetotactic bacteria use different minerals to create magnetic organelles to bias the direction of the bacteria along the Earth's magnetic axis.

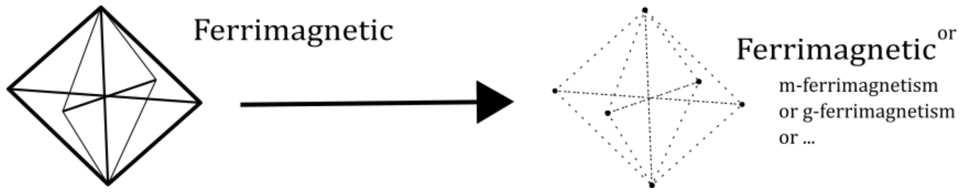
In this sense, there is important, non-empirical scientific content that cannot be preserved by the theories generated by SWC. Accordingly, we ought to reject SWC_1 and SWC_2 as retaining all of the important scientific content of the original theories. Eliminating ferrimagnetism in favor of m-ferrimagnetism and g-ferrimagnetism simply will not do. Therefore, it is not the case that SWC generates theories which show that composites are dispensable to all of our best scientific theories. Properties had only by composites play indispensable explanatory roles in some of our best scientific theories. Therefore, (1)—the claim that ferrimagnetism is indispensable—is true.

4.3.2 Rejecting (1): Disjunctifying Ferrimagnetism

Let's consider another objection. Suppose someone says,

> I agree that we cannot simply eliminate ferrimagnetism and retain all of the important scientific content. However, we should replace ferrimagnetism with a disjunctive property that has, as each of its disjuncts, one of the manifestations of "ferrimagnetism" at the level of simples. Thus, (1) is false and we can retain the original theory's scientific content.

The disjunctive strategy says that any scientifically important property that has multiple realizations at the level of simples is a disjunctive property in disguise, with each disjunct being a collective property. If this is viable, then premise (1) is false: we can replace ferrimagnetism with a disjunctive property ferrimagnetism^{or} that has as each disjunct one of the manifestations of "ferrimagnetism" at the level of simples. Consider here the following diagram:



Here the proponent eliminates ferrimagnetism in favor of a collective property of simples: ferrimagnetism^{or}.

There are many rehearsed arguments against disjunctive properties and how to characterize them.[11] One could argue that disjunctive properties are objectionably unattractive and so reject the ferrimagnetism^{or} theory as meeting the criteria for dispensing with ferrimagnetism. But I do not wish to lean too heavily on the attractiveness condition. Instead, the question is whether a ferrimagnetism^{or} theory can retain all of the important scientific content of the ordinary scientific theory. I think not, and my argument against the disjunctive strategy will reveal the explanatory role that ferrimagnetism plays in our theories.

Here is the disjunctive strategy. Ferrimagnetic^{or} is a disjunctive property, and its disjuncts are m-ferrimagnetic, g-ferrimagnetic, and so on. As I have conceded above, a theory with ferrimagnetic^{or} is empirically equivalent to a theory with ferrimagnetism (assuming there are no other differences). There are two challenges for the disjunctive strategy. First, it is often held that a disjunctive property is legitimate only if the disjuncts bear a real similarity to one another.[12] Second, we learned from the previous section that in order to preserve important scientific content, whatever we replace ferrimagnetism with must have unifying power in explaining magnetotaxis. The disjunctive strategy can seemingly answer both challenges in one breath. The real similarity that the disjuncts bear to one an-

[11] Cf. Bird (1998, 2007, 2016), Mumford (1998); Mumford and Anjum (2011b), Heil (1999, 2003, 2005), Clapp (2001), Lowe (2010), and Audi (2012, 2013).

[12] See Clapp (2001), Antony (2003, p. 10), and Tahko (2021, §§2.3 - 2.5). Cf. Audi (2012).

other, and can account for the unifying power of ferrimagnetism^{or}, is this: each of the disjuncts m-ferrimagnetic, g-ferrimagnetic, and so on all contribute the same causal powers to the simples that bear these properties.[13] It is plausible that magnetotactic bacteria are picking up on exactly this set of causal powers, and this is why they do not discriminate between magnetite and greigite for the purpose of magnetotaxis. Accordingly, the disjunctive strategy is poised to reject premise (1).

I think the disjunctive strategy as presented will not work. This is because, I will argue, the proponent of this strategy cannot identify a relevant set of causal powers of the simples that is shared by all disjuncts of ferrimagnetism^{or}. In order to argue for this, we need to carefully investigate the causal powers of the parts of ferrimagnetic minerals.

Two caveats to this argument. First, I do not know what causal powers mereological simples have. Accordingly, we will not discuss the causal powers of the simples. Instead, I will consider the causal powers of the atoms that make up ferrimagnetic minerals. If we can show that the atoms of different ferrimagnetic minerals do not share some set of causal powers that is distinctive of ferrimagnetism, then I think we can reasonably infer that the simples also do not share this set of powers. Second, I will not be considering the causal powers of the atoms that make up greigite. There is considerable debate over the magnetic structure of greigite. It was originally thought that greigite and magnetite were isostructural, but it has been discovered that they have different magnetic properties (e.g., a lower magnetic moment in greigite than magnetite, despite the iron

[13]Note that this move does not require the claim that properties are entirely characterized by their set of causal powers. Cf. Mumford and Anjum (2011b).

atoms being structurally isomorphic).[14] Because of this, I will be considering a different ferrimagnetic mineral that we better understand: barium ferrite. Barium ferrite is a common ferrimagnetic mineral found most often in credit card magnetic strips. There is no evidence that barium ferrite is used by magnetotactic bacteria to orient direction along the Earth's magnetic axis.

Ferrimagnetism, recall, is the property of an iron-based mineral to have non-net-zero spontaneous magnetization. The way this occurs, put crudely, is this: the iron atoms will be arranged in the crystal structure such that the magnetic moments of some atoms cancel out some others because they are pointed in opposite directions. The atoms pointed in one direction ↑ are unequal to the ones pointed in the opposite direction ↓, and as a result there will be net magnetization in the greater magnetic moment direction. In most ferrimagnetic minerals, there are two kinds of positive iron atoms, Fe^2 and Fe^3, which differ in the contribution each makes to the magnetic moment. Fe^2 provides 4.9 μB (Bohr magneton) in a direction per atom and Fe^3 provides 5.9 μB in a direction per atom. In ferrimagnetic minerals, it is iron atoms of one kind that make up the entirety of the net-positive magnetic moment. For example, in magnetite, there are twice as many Fe^3 atoms as Fe^2 atoms, and the Fe^3 atoms are equally distributed between the two directional sites, so that their 5.9 μB ↑ and 5.9 μB ↓ charges exactly cancel out. As a result, the Fe^2 atoms are the only ones that contribute to the magnetic moment. Here is one philosophically-relevant point: in magnetite, ferrimagnetism—its non-net-zero spontaneous magnetization—is characterized by the causal powers of the Fe^2 atoms to contribute 4.9 μB in some direction.

But not all ferrimagnetic minerals have their ferrimagnetism in virtue of the

[14] Cf. Banerjee and Moskowitz (1985) and Chang et al. (2008, p.2).

causal powers of the Fe^2 atoms. Barium ferrite, which is comprised of iron, oxygen, and barium, is one such case. In barium ferrite, it is the Fe^2 atoms that are structured so as to cancel each other out, while the Fe^3 atoms are what entirely make up the magnetic moment in some direction (Kong et al., 2018, pp. 287 - 296). Here is another philosophically-relevant point: in barium ferrite, ferrimagnetism is characterized by the causal powers of the Fe^3 atoms to contribute 5.9 μB in some direction.

I conclude that the causal powers that characterize ferrimagnetism in magnetite and in barium ferrite are distinct. At the level of atoms, the causal powers that make a mineral ferrimagnetic differ across different manifestations of ferrimagnetism. If these powers are different, then the disjunctive strategy fails. Some instances of ferrimagnetism are characterized by the causal powers of the Fe^2 atoms and others by the distinct causal powers of the Fe^3 atoms. It is just not true to say that there is one set of causal powers of iron atoms that characterizes a mineral's ferrimagnetism. I believe that we can infer, from this, that there is also not one set of causal powers of mereological simples that characterizes a mineral's ferrimagnetism. Thus, the disjunctive strategy, supplemented with a causal powers account of the similarity among disjuncts, cannot identify a real similarity between the disjuncts of $ferrimagnetism^{or}$ that can serve to unify the explanations about magnetotaxis.

It ought not be surprising that ferrimagnetism looks different at the level of atoms. One reason that our best theories include ferrimagnetism, rather than only the different manifestations of ferrimagnetism, is because it is a property that is not uniquely characterizable at the atomic level. Things can look quite different at the atomic level but all result in non-net-zero spontaneous magnetization. And non-net-zero spontaneous magnetization matters for some scientific explanations in a way that the atomic properties do not. Viz., in bacterial magnetotaxis. Ferrimagnetism is a theoretical posit of our best scientific theories, and we should expect that theoretical posits usually play a genuine explanatory role within the theory.

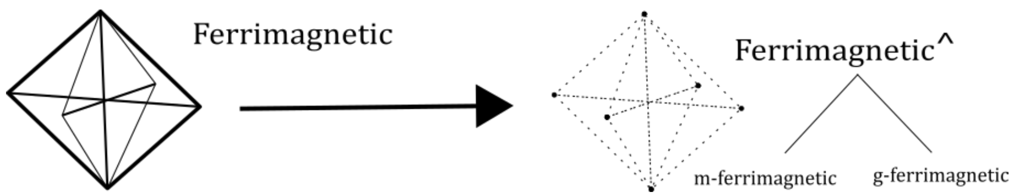
Note that I am not saying that there is no set of causal powers that characterizes ferrimagnetism. There is. And this set of causal powers also explains why magnetotactic bacteria do not discriminate between magnetite and greigite for the purposes of magnetotaxis. My point in this argument is that these causal powers are not had by the parts of the minerals. Instead, they are causal powers had by the minerals themselves: the power of magnetite to have some spontaneous magnetic moment after the formulation of magnetic organelles. The broader point is that ferrimagnetism cannot be properly characterized entirely at lower levels. We saw from the previous section that ferrimagnetism is scientifically important. And here we see that it does not admit of specification by thinking about what is shared among its manifestations. What is scientifically important, then, is that some things have non-net-zero spontaneous magnetization.

4.3.3 Rejecting (1): Irreducible properties of simples

There is another way to reject premise (1). Someone might say the following:

> The previous arguments have shown that we cannot replace ferrimagnetism with properties that are fully reducible to exact properties of simples. Accordingly, we ought to replace ferrimagnetism with ferrimagnetism^, an irreducible property of simples. With ferrimagnetism^, we can preserve the essential content that ferrimagnetism provided.

The idea behind this strategy is as follows. It first affirms that simples arranged magnetite-wise are m-ferrimagnetic and that simples arranged greigite-wise are g-ferrimagnetic. Second, it says that ferrimagnetism^, an irreducible collective property of simples, is also a property of both simples arranged magnetite-wise and of simples arranged greigite-wise. Here a diagram will be useful:



Thus, simples arranged magnetite-wise are both m-ferrimagnetic *and* ferrimagnetic^, which are both properties of simples.[15] This strategy has the advantage of accommodating the explanation that bacteria use different substances with the same property. In this case, though, it is that bacteria use substances that are ferrimagnetic^ rather than substances which are ferrimagnetic. The latter entails the existence of composites, where the former does not.

One prima facie worry for this strategy is that it abandons the spirit of SWC. The primary motivation for SWC is the vision that the physical world is fully reducible to the level of simples and ordinary properties of simples. The inclusion of irreducible properties of simples complicates this vision. It adds levels of physical reality to one's metaphysical picture. Nevertheless, if this strategy works, then premise (1) of my argument is false: ferrimagnetism would not be indispensable.

The main challenge for the ferrimagnetism^ strategy is to show that a theory with ferrimagnetism^ can preserve the unifying content of the original scientific

[15]This response to my argument resembles in spirit the reductionist movement in the multiple realization argument in philosophy of mind along the lines of Lewis (1980), Bickle (1998, Ch. 4), and Sober (1999).

theory. The proponent of this strategy can answer this challenge in the same way that I do. In particular, I say that the original scientific theory is unifying in part because ferrimagnetism is an irreducible property of minerals. The ferrimagnetism^ strategy can answer similarly: ferrimagnetism^ is unifying because it is an irreducible property of simples. It is natural to think that if one provides unifying content, the other does too. It is not enough, however, to simply assert that some property is irreducible because it is irreducible *and* multiply realizable. Many properties are irreducible and their irreducibility does not call out for explanation: charge, mass, etc. But if there is an irreducible property that is multiple realizable, it seems that we need some explanation of how it could both be irreducible and have distinct physical manifestations. So, for the ferrimagnetism^ strategy to work, a proponent must identify the source of the irreducibility. Otherwise, it is unclear whether a ferrimagnetism^ theory genuinely retains the important scientific content.

I have two arguments that the ferrimagnetism^ strategy cannot meet this challenge. First, I will argue that there is no good way to characterize the irreducibility of ferrimagnetism^. Second, I will argue that there *is* a good way to characterize the irreducibility of ferrimagnetism. I take it that if the proponent of this strategy cannot identify the source of the irreducibility of ferrimagnetism^, then they cannot simply assert that this property is unifying.

The most natural way to characterize the irreducibility of some property is to show that that property is functional. Consider the property that characterizes the kind *corkscrew*: call it corkscrewness (Shapiro, 2000). Corkscrewness is irreducible because it's essentially functional. Something has corkscrewness just in case it can serve the function of taking a wine cork out of a bottle by screwing some part of itself into the cork and removing the cork without damaging the bottle. Many different objects have this function: waiters corkscrews, double-lever corkscrews, and so on. Accordingly, one might be drawn to the view that the irreducibility of ferrimagnetism$^\wedge$ is characterized by it being a functional property.

However, ferrimagnetism$^\wedge$ is simply not a functional property in the sense needed. For this approach to work, we need to have a clear distinction between those multiply realizable properties that are functional and so irreducible and those multiply realizable properties that are non-functional and so not irreducible (in this way). Compare this to the corkscrew: its color is a non-functional property, while its corkscrewness is a functional property. But a "ferrimagnetism" property (whether ferrimagnetism proper or ferrimagnetism$^\wedge$) is not a characteristically *functional* property of this sort. It's the property of some thing(s) to have non-net-zero spontaneous magnetization. Now, certainly this property confers some dispositions or causal powers to whatever it is that is "ferrimagnetic". But if that is the mark of a functional property, then *every* natural property is functional, and we have lost the distinction between functional and non-functional that can serve to justify the irreducibility of some property on the basis of its being functional. So, it seems that the most obvious way to characterize the irreducibility of ferrimagnetism$^\wedge$—as a functional property—cannot be made to work.

The very same question might be turned on me: what can account for the irreducibility of ferrimagnetism? It cannot be, according to my reasoning, that ferrimagnetism is a functional property. Instead, the answer is that ferrimagnetism is a *compositionally determined* property. This might seem like an uninteresting or circular answer, but it is what we can say about ferrimagnetism by looking closely at the science. When iron and oxygen atoms are arranged in octahedral and opposing tetrahedral crystals, together with the magnetic properties of the iron atoms, these properties compositionally determine that the mineral is ferrimagnetic: it has non-net-zero spontaneous magnetization. What accounts for the irreducibility of ferrimagnetism is that it is a compositionally determined property that can be had by many different kinds of ferrites.

A more interesting question, though, is why our scientific theories favor positing irreducible, compositionally determined properties such that their inclusion in a theory can contribute non-empirical content. Why do our best theories include irreducible properties like ferrimagnetism? Answering this question is a much larger project, but we can say a few speculative remarks.[16] The idea, as I see it, is that an aspect of our best scientific theories is that there are dimensions or levels of description of the physical world. Our theories must be able to capture the natural processes that occur between individuals at the same level. For example, how bacteria use ferrimagnetic materials to orient themselves along the North-South axis. But our theories give explanations at other levels too. For example, our theories will explain how the Fe^2 and Fe^3 atoms are oriented within a crystal structure to have some positive magnetic moment in a direction. Both kinds of explanations are scientifically important. Moreover, it is plausible that compositional explanations in science are those explanations that hold between levels. Accordingly, our best theories include irreducible properties like ferrimagnetism because such properties permit us to explain the phenomena at one level of description. The proponent of SWC might worry that this is question-begging. I am saying that our best scientific theories include irreducible compo-

[16]For a more sustained defense of compositional explanations in science, see Gillett (2016) and Aizawa and Gillett (2016, esp. Ch. 8).

sitional properties because our scientific theories are wont to explain relations at the level where composite objects would exist. This brings us to the final objection to my argument.

4.4 Rejecting (1): Why care about bacteria?

There is one final conspicuous route for the SWC theorist to reject my argument. Suppose I have successfully shown that ferrimagnetism is indispensable to explaining that different magnetotactic bacteria use different minerals for the same behavior, and that ferrimagnetism is a singular property. Someone still might object as follows:

> I concede that ferrimagnetism (as a singular property) is indispensable to explaining bacteria behavior. But I reject the claim that ferrimagnetism is indispensable *to our best scientific theories*. I reject that our best scientific theories even need to explain these features of bacteria. All that science needs to explain is features of microphysical simples and the ways they are arranged.

I have two responses to this objection. First, as I just remarked, scientists are motivated to provide theories that explain behaviors of individuals at different levels of description. It is not question-begging against the SWC theorist to take this fact at face value. *Which* things exist or are explanatorily relevant is an open question that gets determined by the success of the theories. In particular, saying that scientists are motivated to provide explanations at different levels of description does not presuppose the existence of bacteria. It does not even presuppose the existence of composite objects. However, it turns out that positing compositional relations between levels is a historically successful way to provide explanations at different levels. And it turns out that positing irreducible compositional properties of composites, like ferrimagnetism, is successful. Accordingly, rejecting the project of providing explanations at the level of bacteria, which this objection tries to do, entails that we reject a successful practice of scientists.

Second, I do not think this objection is very convincing because of my orientation to metaphysics of science. My aim here, and the aim of one pursuing an indispensability argument, is not to produce a metaphysics that is merely consistent with our best scientific theories. Rather, it is to produce a metaphysics that is faithful to the content of our best scientific theories. We, as *responsible* metaphysicians of science,[17] ought to perform careful investigations into the content of our best scientific theories with an eye toward their metaphysical consequences. We ought not begin our metaphysics of science attempting to defend a particular metaphysical thesis we came to accept on philosophical grounds. This response to (1) fails to be a responsible metaphysics of science. It is one thing to show that portions of science can be relatively adequately preserved without appealing to composites. It is entirely another thing to show that faithful rational reconstructions of science can be done without appealing to composites. And it is the latter project that I am concerned with. It is my claim that we cannot give a faithful rational reconstruction of our best microbiology theories without appealing to ferrimagnetism and composites like minerals. Thus, a responsible examination of our best scientific theories requires that we ontologically commit to composites like minerals, diamonds, corks, and stars.

[17]Cf. Bryant (2018).

Chapter 5

Putnam on Mathematics and Ontological Commitment

In the *Stanford Encyclopedia of Philosophy*, Colyvan (2008) presents what he calls the "Quine-Putnam indispensability argument" for the existence of mathematical objects:[1]

Colyvan's version

P1. If some entity is indispensable to some of our best scientific theories, then we ought to ontologically commit to that entity.

P2. Mathematical objects are indispensable to some of our best scientific theories.

P3. So, we ought to ontologically commit to mathematical objects.

Colyvan's argument claims we must commit to mathematical objects (like sets or numbers) because such objects are indispensable to our best scientific theories. This version of the argument is seen as its canonical formulation, with nearly all commenters taking the conclusion to be about ontological commitment (Field, 2016; Maddy, 1992; Sober, 1993).

Putnam read Colyvan's article in the *SEP*. According to Putnam, he never endorsed the eponymous Quine-Putnam indispensability argument. He says that Colyvan's version "is far from right" (Putnam, 2012, p. 182). Colyvan's conclusion is that we ought to ontologically commit to mathematical objects. Putnam

[1] Colyvan's actual formulation differs in trivial ways.

claims that this is not his conclusion.[2] Instead, he says, "my "indispensability" argument was an argument for the objectivity of mathematics in a realist sense" (Putnam, 2012, p. 183). By 'the objectivity of mathematics in a realist sense' he means that there are true mathematical claims and that their truth is objective, not dependent on human activity. (Hereafter, I'll use 'true' to mean objectively true.) Moreover, he claims he argued that these true mathematical claims do not "have to be interpreted Platonistically", in the sense that we can accept their truth but refrain from ontologically committing to mind-independent mathematical objects. In his recollection, Putnam rejected the conclusion of the Quine-*Putnam* indispensability argument from the start.[3]

Instead, he recalls that he argued for the following conclusions:

C1. Some mathematical statements are objectively true.

C2. We need not be ontologically committed to mathematical objects.

While Putnam (2012) corrects Colyvan on the conclusions of his argument, he does not tell us what the argument was. It is clear he must reject at least one of Colyvan's P1 or P2. And there must be some premise that infers mathematical upshots from scientific theories. But we do not find a clear articulation of Putnam's actual indispensability argument in this recent chapter.

[2]Although Putnam himself made comments that seemingly endorse the Quine-Putnam indispensability argument. There is a notorious passage in Putnam (1971, p. 57):

> So far I have been developing an argument for realism along roughly the following lines: quantification over mathematical entities is indispensable for science, both formal and physical; therefore we should accept such quantification; but this commits us to accepting the existence of the mathematical entities in question.

Some have argued that we should take Putnam as presenting for consideration this indispensability argument. See, e.g, Bueno (2018, pp. 204 - 205).

[3]Others have noted this discrepancy, some from before Putnam published his 2006/2012 correction. See Burgess and Rosen (1997, pp. 200 - 201), Liggins (2008), Bueno (2013, 2018), Burgess (2018), Clarke-Doane (2020, p. 26), and Barrett (2020b).

We needn't remain shrouded in mystery, though. Putnam (2012, p. 183) assures us that "[o]bviously, a careful reader" of his early work on philosophy of mathematics would find his argument there. This paper attempts to cash in on Putnam's promissory note. I have two aims here. The first is to provide the definitive reconstruction of Putnam's indispensability argument for the objectivity of mathematics without mathematical objects. To do this, I show that a recent reconstruction by Bueno (2013, 2018) cannot be Putnam's argument. Then, I present my own reconstruction of Putnam's indispensability argument for C1, which employs the logic of linguistic presupposition. Once we understand Putnam's argument as one from presupposition, we can see why he so vociferously objected to Colyvan's claim that he was ontologically committing to mathematical objects. The second aim of the paper is to extract the theory of ontological commitment implicit in Putnam's argument for C2. I show that appreciating this theory of commitment can help answer a recent challenge to Putnam's mathematical metaphysics given by Burgess (2018). I also show that this theory of commitment, or something near it, might be helpful to some very recent trends in philosophy of physics.

5.1 Against the Colyvan-Bueno Reconstruction

Putnam is explicit that he does not endorse Colyvan's argument. His first comments after quoting Colyvan argument are as follows:

> From my point of view, Colyvan's description of my argument(s) is far from right. The fact is that in "What Is Mathematical Truth?" [(1975)] I argued that the internal success and coherence of mathematics is evidence that it is true under some interpretation, and that its indispensability for physics is evidence that it is true under a realist in-

> terpretation...It is true that in *Philosophy of Logic* [(1971)] I argued that at least some set theory is indispensable in physics ... but both "What Is Mathematical Truth?" and "Mathematics without Foundations" [(1967)] were published in *Mathematics, Matter and Method* together with "Philosophy of Logic," and in both of those papers I said that set theory did not have to be interpreted Platonistically. (Putnam, 2012, p. 182, emphasis removed)

Putnam here rejects Colyvan's indispensability argument as his own. He instead claims that C1 was always his conclusion and suggests that C2 was consistent with his indispensability argument. C1 is the claim that some mathematical statements are true. Putnam straightforwardly says this: "[mathematics] is true under a realist interpretation". C2 is the claim that we need not be ontologically committed to mathematical objects. In this passage, Putnam says that we need not be platonists with respect to the mathematics required by physics.

Putnam thinks some premises regarding the relationship between physics and math entail C1. Yet nowhere does he explicitly state these premises. Bueno (2013, 2018) proposes that Putnam endorsed premises similar to Colyvan's P1 and P2. Bueno's change is to weaken P1 and P2 to only guarantee that the resulting mathematical theories are *truth apt*, by which he means they are capable of truth or falsity and their truth or falsity is objective (Bueno, 2018, p. 205). Consider Bueno's reconstruction:

P1′. If any theory (existentially) quantifies over some entity that is indispensable to one of our best scientific theories, then that theory is truth apt.

P2′. Some mathematical theories (existentially) quantify over entities that are indispensable to our best theories.

C1′ So, some mathematical theories are truth apt. (Bueno, 2018, p.209)

The idea behind Bueno's version is as follows. Just like with Colyvan's, what matters for the first premise is whether quantification over mathematical objects is indispensable to our best scientific theories. With P1, indispensable quantification over mathematical objects entails that we should ontologically commit to them. With P1′, the entailment is weaker. All that is entailed by indispensable quantification over some entity is that the theories which quantify over those indispensable entities are truth apt. Bueno does not give an argument for P1′, but we can see the idea behind it. Electrons, e.g., are indispensable to our best scientific theories. Per P1′, this entails that any theory that quantifies over electrons should be interpreted as a theory that, either correctly or incorrectly, attempts to describe the world. Theories that quantify over electrons are thus truth apt in Bueno's sense.

P2′ requires some explanation. Strictly, all it says is that there are mathematical theories that *do* quantify over entities that are indispensable to our best scientific theories. Bueno clarifies that this includes pure mathematical theories (Bueno, 2018, p. 209). So we should interpret P2′ as requiring two things. First, that mathematical objects, e.g., sets or numbers, are indispensable to our best scientific theories, and second, that pure mathematical theories quantify over mathematical objects. Again, here, we see a parallel with Colyvan's P2.

Bueno's reconstruction of Putnam's indispensability argument is primarily meant to motivate the prospect of a modal mathematical picture. But as an artifact of Putnam scholarship, I think we should reject it as accurately capturing the relationship Putnam thinks mathematics bears to our best scientific theories.

The primary reason for this is that P1′ and P2′ together do not guarantee Putnam's conclusion C1. All that Bueno's version of the argument can show is that mathematical theories are truth *apt*, that they are capable of truth and falsity, not that they are true. It is consistent with Bueno's C1′ that all mathematical theories are false. And as we saw in the quote above, Putnam is explicit that he thinks his indispensability argument shows that mathematics is true. Putnam, separately in 1975, straightforwardly said as much: "Mathematical experience says that mathematics is true under some interpretation; physical experience says that that interpretation is a realistic one" (Putnam, 1975, p. 74). It is clear that Putnam endorses C1, not merely C1′, and Bueno's reconstruction wrongly claims C1′ to be the conclusion of Putnam's indispensability argument as presented in Putnam's early papers (1967, 1971, 1975).

One may wish to adjust Bueno's version of the argument in light of this objection. Suppose we replace P1′ with the following:

P1″. If some theory (existentially) quantifies over some entity that is indispensable to one of our best scientific theories, then that theory is true.

P1″ together with P2′ entail C1, that some mathematical theories are true. Bueno's version of the argument can then be simply adjusted to match Putnam's stated conclusions.

The problem with P1″ is that it entails a contradiction (on the supposition that there are any entities which are indispensable to our best scientific theories). Suppose integers are indispensable to our best scientific theories. According to P1″, any theory that existentially quantifies over integers is true. Both the claim that 1 is less than 2 and the claim that 1 is not less than 2 existentially quantify over integers. Given P1″, both are true. But this is a contradiction. So we should reject P1″ on these grounds.[4]

[4]There may be a way to infer from C1′ to C1. Namely, if one accepts certain premises about

I think what has gone wrong is this. We have been considering a version of Putnam's argument for C1 that structurally matches Colyvan's presentation, with one premise that establishes the metaphysical upshots of indispensable reference to entities and another premise that establishes the indispensable reference to mathematical entities. And we have found that there is no good way to shoehorn Putnam's conclusions into this structure. We should, instead, abandon the Colyvan-Bueno structure and look closely at Putnam's comments from those original three papers (Putnam, 1967a, 1971, 1975).

5.2 Putnam's Presupposition Argument for C1

My preferred reconstruction of Putnam's indispensability argument for C1, which I call 'Putnam's presupposition argument', aims to establish the truth of mathematical claims on the basis of the role mathematics plays in our best scientific theories. At its core, the idea is that our best theories *presuppose* the truth of mathematical claims, in the traditional sense of presupposition from philosophy of language. Here, I'll show why the following is the best reconstruction of Putnam's argument for C1:

1. Our best physical theories are truth apt.

2. Our best physical theories presuppose some mathematical statements.

3. If 1 and 2, then some mathematical statements are true.

C1. So, some mathematical statements are true.

what mathematical truth amounts to. Suppose we accept the following claim: "If a mathematical theory is consistent and truth apt, then it is true." Such a premise would, together with C1′, necessitate C1. There is, as far as I can tell, no textual evidence that Putnam accepted this premise.

The idea, as we'll see, is relatively straightforward. The definition of presupposition entails 3, and these three premises necessitate C1.

Let us define presupposition:

Presupposition. A statement S presupposes a statement p IFF for S to be truth apt, p must be true.

Roughly, one statement presupposes another statement if for the first to be *truth apt*, the second must be *true*. Strawson (1950) is the locus classicus of this sense of presupposition. Traditionally, philosophers were concerned with existential presuppositions triggered by definite noun phrases, e.g., that 'The king of France is bald' presupposes that the king of France exists. There are, though, many statements that have non-existential presuppositions. For example, the statement 'Dylan quit vaping' presupposes that Dylan once vaped. Similarly, Putnam thinks that there are statements in our best physics that presuppose the truth of mathematical claims. The relevant presupposition for Putnam is *not* the existence of mathematical objects, but the truth of mathematical claims. It is unclear whether Putnam endorsed Strawson's technical definition of presupposition.[5] Nonetheless, several comments and arguments Putnam made in the context of his indispensability argument were employing the logic of presupposition.

Consider the following, one from each of his three early works on mathematics:

> We will be justified in accepting classical propositional calculus or Peano number theory... because a great deal of science presupposes these statements. (Putnam, 1967a, p. 13)

[5] There is a passage in Putnam (1971, pp. 28 - 30) that indicates some familiarity with this technical definition. However, Putnam seems to be using 'meaningless' and 'neither true nor false' interchangeably, which makes it difficult to pin any view on him.

> [L]et us consider what is involved [in providing a mathematics-free version of Newton's theory of universal gravitation], and let us consider not only the law of gravitation itself, but also the obvious presuppositions of the law. The law presupposes, in the first place, the existence of forces, distances, and masses—not, perhaps, as real entities but as things that can somehow be measured by real numbers. (Putnam, 1971, p. 37)

> If one . . . wants to say that the Law of Universal Gravitation makes an objective statement about bodies . . . What is the statement? It is just that bodies behave in such a way that the quotient of two numbers associated with the bodies is equal to a third number associated with the bodies. But how can such a statement have any objective content at all if numbers and 'associations' (i.e. functions) are alike mere fictions? (Putnam, 1975, p. 74)

In each of these cases, Putnam seems to be pushing the claim that the relationship that physics bears to mathematics is one of *presupposition.* Though the first two mention presupposition explicitly, it is the third which is most clearly an argument from the logic of presupposition. Putnam is arguing that the Law of Universal Gravitation is neither true nor false (does not have objective content) if mathematical claims are not literally true, a straightforward implication of the thesis that the law presupposes mathematics.

With this initial justification for taking presupposition seriously, let us turn to 1 and 2 in more detail. Putnam explicitly endorses 1 as his first premise. He says,

> I shall assume here a "realistic" philosophy of physics; that is, I shall assume that one of our important purposes in doing physics is to try to state "true or very nearly true" (the phrase is Newton's) laws, and not merely to build bridges or predict experiences. (Putnam, 1971,

> 36)[6]

It is clear that Putnam took his first premise to be the statement of scientific realism. In a similar vein, the canonical argument for scientific realism, the no miracles argument, was first formulated in the very same paper. We note that the Bueno-Colyvan formulation has no premise endorsing this kind of scientific realism.

Putnam's primary argument for 2 is rather flat-footed. He considers Newton's law of universal gravitation. Simply *reading* the law shows how it presupposes mathematical claims:

> Newton's law, as everyone knows, asserts that there is a force f_{ab} exerted by any body a on any other body b. The direction of the force f_{ab} is towards a, and its magnitude F is given by
>
> $$F = \frac{gM_aM_b}{d^2} \tag{N}$$
>
> where g is a universal constant, M_a is the mass of a, M_b is the mass of b, and d is the distance which separates a and b. (Putnam, 1971, p. 36)

According to Putnam, (N) straightforwardly presupposes mathematics in the sense that it "has a mathematical structure" (*ibid.*, p. 37).

One fears that this is too easy. What is so special about the formulation of the law presented in (N)? It seems prima facie possible that there might be a reformulation of (in the sense of providing an equivalent statement of) Newton's

[6]See also Putnam (2012, 183, emphasis removed): "Nevertheless, there was a common premise in my argument and Quine's, even if the conclusions of those arguments were not the same. That premise was "scientific realism," by which I meant the rejection of operationalism and kindred forms of "instrumentalism." I believed (and in a sense Quine also believed) that fundamental physical theories are intended to tell the truth about physical reality, and not merely to imply true observation sentences."

law of universal gravitation that does *not* have a mathematical structure. If there is such a reformulation of (N) that does not have a mathematical structure, then plausibly (N) itself does not presuppose mathematics.

To this end, Putnam (1975, Ch. 5) argues that it is impossible to provide a non-mathematical reformulation of Newton's law of universal gravitation. He argues:

2a. Some true physical theories presuppose an arbitrary amount of true facts of the form "the distance between a and b is d".

2b. The only way to formulate arbitrarily many facts of the form "the distance between a and b is d" is with mathematical statements.

2c. If 2a and 2b, then some of our best physical theories presuppose some mathematical statements.

2. So, some of our best physical theories presuppose some mathematical statements.

According to Putnam, among the presuppositions of (N) is an arbitrary number of non-equivalent facts of the form "the distance between a and b is d". The argument for this is a nebulous. Here's an attempt: since (N) is universal, it necessarily permits a solution given any value of d. This seems to require an arbitrary number of distance facts. Putnam does not think anything is special about distance predicates; he also thinks that physics presupposes arbitrarily many force, mass, and charge facts. Putnam here is pointing out what has been called the *indexing* role that mathematics plays (Melia, 2000, p. 473). (Cf. Daly and Langford (2009) and Baker and Colyvan (2011).) The idea is that mathematics helps us index various concrete, physical facts. Putnam takes this indexing role of mathematics to be a presupposition of our formulations of the laws of physics. This is his defense of 2a.

Putnam then attempts to prove that the only way to countenance this presupposition is with mathematics. Premise 2b is the "indispensability" premise: in science, we must countenance these distance facts, and *the only way to do so* is with mathematics. He gives a short proof of this premise 2b. The basic idea is that, given some constraints, one cannot state all the presuppositions of (N)—pairwise inequivalent distance statements—unless one uses the apparatus of mathematics. Barrett (2020b) examines this proof in detail and shows that it is not especially convincing.

There's some more textual evidence for this interpretation of Putnam's argument for 2. Field (2016) rejected Putnam's claim 2b. Field showed that one *could* formulate arbitrarily many facts of the form "the distance between a and b is d" without the use of numbers. He did this by quantifying over an infinite number of spacetime points and using relative distances. So, if 2b was part of Putnam's argument, then we should expect that he would concede that Field's project succeeded in responding to this argument for 2.

And indeed this is exactly what happened. Putnam says,

> Hartry Field understands very well what my arguments were, and he attempts to meet them on their own terms ... I agree that, assuming the nominalistic acceptability of [his assumptions] ..., Field has shown that much or perhaps even all of classical physics can avoid any use of set theory at all. (Putnam, 2012, pp. 190 - 191)

Per this passage, Putnam concedes that Field has shown 2b to be false. However, Putnam immediately pivots and argues that Field's strategy cannot be extended to account for the mathematical presuppositions of quantum gravity theory (*ibid.*). I take this to mean that Putnam has retreated from his original argument for 2 but ultimately stands by the form of it, modulo the indexing role mathematics plays in more contemporary physics.

Let us back up. Here I have argued for a new interpretation of Putnam's argument for the objectivity of mathematics (C1) on grounds of its "indispensability to science". The basic idea, following close investigation of his original papers and his recent remarks, is that our best theories in physics presuppose the truth of mathematics: To even state our best physical theories, we must use mathematical claims. And given the reigning definition of presupposition and the truth aptness of our physical theories, this entails that these mathematical claims are *true*. I have also reconstructed Putnam's argument for the claim that physics presupposes math, connecting it to his recent concession to Field.

One may wonder how my reconstruction constitutes an "indispensability" argument. Nowhere does the term 'indispensable' appear in the premises. This is in part because, as Barrett (2020b) has shown, Putnam uses 'indispensable' in a non-standard way. Nonetheless, we see that there is a sense in which Putnam is giving an indispensability argument, in the sense of *not being able to do without*. This is in his 2b. He effectively argues that there is a theory T such that T is among our best physical theories and if any T' is equivalent to T, then T' presupposes mathematics. Or, in other words, there is no way to formulate our best physical theories without presupposing mathematics. In this sense, Putnam has shown that mathematics is "indispensable" to science.

5.3 Putnam's Metaphysical Picture Argument for C2

Naturally, one might think that Putnam's argument for C1 entails that we must be ontologically committed to mathematical objects. After all, the thought goes, if we presuppose the *truth* of 2+2=4, it is a short logical hop to committing to the existence of the number 2. Putnam disagrees; he believes we can have objectivity without objects.

This line of thought provides the motivation for Putnam's argument for C2. Here is my reconstruction of it:[7]

4. There are non-platonistic metaphysical pictures of the truth of mathematical claims.

5. If so, then we need not be ontologically committed to mathematical objects.

C2. So, we need not be ontologically committed to mathematical objects.

We will say more about metaphysical pictures, but one can think of a metaphysical picture as an explanation of a statement's truth. Putnam uses the term 'picture', but we will use the more precise 'metaphysical picture'.

Putnam clearly endorses premise 4, having both written a paper defending it (Putnam (1967a)) and continually referring to that paper as evidence for his conclusion (C2)—that we need not be committed to mathematical objects. (See, e.g., (Putnam, 1971, pp. 75 - 76), (Putnam, 1975, pp. 70 - 72), (Putnam, 2004, pp. 66 - 67), and (Putnam, 2012, pp. 182 - 183, 190).) In this section, I will reconstruct a theory of ontological commitment that entails 5, and which will permit Putnam to have his objectivity without objects.

[7] Bueno seems to think it is something similar. See (Bueno, 2018, p. 206).

First consider 4. Putnam (1967a) argued that we can provide a non-platonistic metaphysical picture of the truth of mathematical claims. The idea here is that we do not need to think that true mathematical claims are made true by a platonic realm full of mathematical objects. Rather, we can think that true mathematical claims are made true by the fact that certain mathematical structures are possibly satisfied. The idea is that, in mathematics, the existence of a given *object* is completely fungible with the possible existence of a certain *structure* (Burgess, 2018, p. 12). We can, the story goes, reformulate all true mathematical claims in a modal second-order language. Take the statement that Peano's axioms entail that there are infinitely many prime numbers. Putnam's modal reformulation would be something like the following:

Peano$_{modal}$ There are possible structures where Peano's axioms are satisfied.

Prime$_{modal}$ Any possible structure where Peano's axioms are satisfied is a possible structure where there are infinitely many prime numbers.

This second reformulation, Prime$_{modal}$, is meant to capture the same content as "Peano's axioms entail that there are infinitely many prime numbers". In this way, Putnam thinks mathematical claims like these can be understood as describing entailments among possible structures, rather than describing a realm of mathematical objects.

The tenability of Putnam's modal mathematics is controversial (Kreisel, 1972; Burgess and Rosen, 1997; Hellman, 1989; Bueno, 2018). Luckily, we need not worry about it here, as we are only concerned with the form of Putnam's argument. Let's proceed, then, on the assumption that Putnam has succeeded in giving a translation of every true mathematical claim in terms of mathematical possibility.

It is as yet unclear what Putnam thinks he accomplished by offering a modal mathematics. Crucially, Putnam does not wish to supplant ordinary mathematics with his modal picture. As he says,

> My purpose is not to start a new school in the foundations of mathematics (say, "modalism"). Even if in some contexts the modal logic picture is more helpful than the mathematical-objects picture, in other contexts the reverse is the case. Sometimes we have a clearer notion of what 'possible' means than of what 'set' means; in other cases the reverse is true; and in many, many cases both notions seem as clear as notions ever get in science. Looking at things from the standpoint of many different "equivalent descriptions," considering what is suggested by all the pictures, is both a healthy antidote to foundationalism and of real heuristic value in the study of first order scientific questions. (Putnam, 1967a, pp. 19 - 20)

We should not think of Putnam as offering a traditional alternative theory in the sense of *replacing* the original statements. Nor does Putnam think his alternative reveals the actual commitments of the original statements; he thinks the two pictures of mathematics are incompatible in some important sense.[8] In this way, Putnam is neither offering a hermeneutic nor revolutionary paraphrase: he seeks neither to supplant the original theory nor to reformulate the original theory in a way that reveals its actual commitments.[9]

What, then, is Putnam trying to do by providing non-platonistic replacements of every mathematical claim? Why does it matter to one's metaphysical picture

[8]See: "In short, if one fastens on the first picture (the "object" picture), then mathematics is wholly extensional, but presupposes a vast totality of eternal objects; while if one fastens on the second picture (the "modal" picture), then mathematics has no special objects of its own, but simply tells us what follows from what." (Putnam, 1967a, p.11).

[9]Different philosophers use different names for these paraphrases. Metaphysicians use the compatibilist-incompatibilist phrasing (O'Leary-Hawthorne and Michael, 1996; Korman, 2009; Bagwell, 2021). Philosophers of math use call these same strategies *hermeneutic* and *revolutionary* paraphrases (Burgess and Rosen, 1997). And some call them *reconciling* and *revisionary* (Keller, 2015, 2017).

whether we have objectual statements that are fungible with modal statements? Here I will provide a theory of ontological commitment that will entail 5. It is likely that Putnam would not endorse the view presented here, given his move toward anti-realism and conceptual relativity in the 1980s. But it is inspired by this early 1967 work in philosophy of mathematics.

Here is the theory of commitment: One's ontological commitments in accepting some statements as true depend on one's metaphysical picture of the truth of those statements. A metaphysical picture of the truth of a statement is an account of how that statement could be true. It is something like a metaphysical explanation.

Some examples will help clarify. Consider an easy case:

(*i*) Electrons are negatively charged.

(*i*) is objectively true. A scientific realist plausibly thinks that (*i*) is true because there actually are electrons and they actually have the property of being negatively charged. This explanation constitutes a metaphysical picture of the truth of (*i*). Call it the 'ordinary picture'. (Alternatively, we might call it the Tarskiian picture. Cf. Tarski (1956).) On the ordinary picture of the truth of some claim, the claim is true because the things the subject terms refer to actually exist and have the properties so predicated. The ordinary picture of (*i*) is that there actually are electrons and that they actually are negatively charged. Thus, if one accepts the ordinary picture of the truth of (*i*), then one is ontologically committed to electrons. For most claims we think are true, we accept the ordinary picture of their truth.

But there are statements we take to be true where we need not (and plausibly do not) accept the ordinary pictures of their truth. Consider:

(*ii*) The average star has 2.4 planets.[10]

Let's say I accept that (*ii*) is objectively true, a fact about the world. The ordinary picture says that (*ii*) is true because there actually is the average star and that it actually has the property of having 2.4 planets.[11] The ordinary picture is intuitively not our preferred explanation of its truth. We don't think that (*ii*) is true because there is some average star, floating somewhere in the universe or the platonic realm with exactly 2.4 planets orbiting it. Instead, we offer an alternative picture of the truth of this statement: the ratio of planets to stars is 2.4. What explains why (*ii*) is true is that there are 2.4 times as many planets as there are stars. This alternative picture does not commit us to the existence of some spooky *average star*. So, we can say this: (*ii*) is objectively true, and its truth is metaphysically accounted for by the ratio of planets to stars. In this way, we can accept (*ii*) as literally and objectively true but refrain from ontologically committing to such a thing as the average star. This sort of move is permitted by a theory of ontological commitment that says one's ontological commitments are determined by the metaphysical picture one accepts.

There are, then, two *admissible* pictures of the truth of (*ii*). The first is the ordinary picture, that there is an average star with the property mentioned. The second is that the ratio of planets to stars is 2.4. These two pictures are both admissible in the sense that they provide explanations for how 'The average star has 2.4 planets' can be true. We reject the first, though, because it has untoward metaphysical consequences. It commits us to the existence of strange entities like partial planets and privileged stars. There are three lessons from this exercise.

[10]Cf. Melia (1995).

[11]Kennedy and Stanley (2009) argue that 'average' is semantically not an ordinary adjective, and so 'the average star' doesn't actually serve to refer in the same way that 'electron' does. I'll ignore this complication here.

First, there can be multiple admissible pictures of a statement's truth. Second, we are rationally permitted to choose between those admissible pictures. Third, and most radically, that a given statement is true does not (always) by itself determine our metaphysical commitments in accepting its truth.

Putnam's modal mathematical view fits nicely with this theory of metaphysical commitment. For Putnam wishes to retain that the following is objectively true:

(*iii*) Peano arithmetic entails that there are infinitely many prime numbers.

Putnam accepts (*iii*), and he accepts Peano arithmetic. And as a metaphysician, Putnam must present some metaphysical picture of its truth. There is, of course, the ordinary picture, which will entail the existence of infinitely many prime numbers. Putnam, though, has a couple reasons to be suspicious of the ordinary picture; for instance, he is suspicious of the prospects of a platonistic mathematical epistemology Putnam (1975, p. 71). He has an out though: Prime_{modal} (we assume) presents an admissible picture of the objective truth of (*iii*). This alternative picture does not commit one to mathematical objects. And Putnam believes he is rationally free to choose among admissible pictures. He thinks we should look to the "many different" pictures Putnam (1967a, p. 20). He can accept that Prime_{modal} explains the truth of (*iii*), and thus that we need not be ontologically committed to mathematical objects.

This, I take it, is Putnam's argument for C2. He takes himself to have shown that there are non-platonistic pictures of the truth of mathematical claims, and his tacit appeal to this alternative theory of ontological commitment permits him to not commit to mathematical objects.

5.4 Conclusion

In this paper, I hope to have shown that Putnam endorsed an indispensability argument for the *objectivity* of mathematics that uses the logic of presupposition as a fulcrum. This is in contrast to the recent and historical interpretations (by Bueno and Colyvan respectively) that rely on a robust notion of indispensability. I have also reconstructed Putnam's argument for mathematical objectivity *without mathematical objects*, and have extracted a theory of ontological commitment that is concordant with this argument. Much of this project has been of historical interest. However, I think that this Putnam-inspired theory of ontological commitment is of interest in contemporary metaphysics of science. I shall conclude by sketching the case for this.

Let us first clearly state the metaphysical picture theory of ontological commitment:

Metaphysical picture theory of commitment. The ontological commitments of S's acceptance of a theory T are the entities appealed to in the metaphysical picture that S accepts of the truth of T.

Upon accepting some theory as true, an agent S turns to the admissible metaphysical pictures that can account for the truth of that theory. There may be multiple admissible metaphysical pictures for some theory. In these cases, S is rationally free to choose among those pictures; this is because, in such cases, the content of T does not determine a single metaphysical picture. This theory of commitment seems to provide an intuitive understanding of indispensability as well: if some entity e is appealed to in every metaphysical picture of the truth of some theory, then es are indispensable to that theory.

The nature of a metaphysical picture is still underspecified. Are they written out in a language like a theory? A list of entities and properties had by those entities? Something else entirely? I suspect that each of these options could be made to work. For our purposes here, we will proceed on the assumption that a metaphysical picture is written out in first order logic and that "appeal" to an entity is existential quantification over that entity. In this way, a metaphysical picture is a collection of statements that serves to metaphysically explain some other collection of statements. (A full treatment of the view would also have to specify this sense of metaphysical explanation.) Let's see why this theory is fruitful.

In a recent paper, Burgess (2018) presents an objection to Putnam's modal mathematical picture. In particular, he is vexed by how modal and objectual mathematics constitute "equivalent descriptions" but have different metaphysical implications. I believe this objection can be dispelled by appreciating the metaphysical picture theory of commitment. Burgess says,

> As the Council of Nicæa declared that the Father and the Son are somehow the same and yet somehow different, so Putnam declares the "mathematics as set theory" and "mathematics as modal logic" pictures . . . are somehow the same and somehow different. I find the Nicene Creed easier to understand than Putnam's notion of equivalent descriptions. (Burgess, 2018, pp. 16 - 17)

As he goes on to say, Burgess is confused what Putnam's larger project could even be. Putnam does not wish to *supplant* ordinary mathematics with his modal picture; nor does he think that the two are *completely* equivalent, for then they could not have different metaphysical consequences.

The picture theory of commitment provides an answer to Burgess's question.

Putnam is attempting to give a metaphysical picture of the truth of mathematical claims. What is "somehow the same and yet somehow different" is the ordinary picture of mathematics and the modal picture of mathematics. They are the same in the sense that they are both admissible pictures of the truth of mathematical claims; any claim of mathematics can be explained by either picture. We might say that from the perspective of *mathematics*, they are equivalent. On the other hand, they are metaphysically distinct. They appeal to different entities. One appeals to (in the sense of quantifying over) numbers, the other does not. From the perspective of *metaphysics*, they are different. The same, but different.

Some contemporary philosophers of physics are attempting to carve out a position, like Putnam, where there may be "equivalent descriptions" of some theory but where those descriptions have different metaphysical implications. This is in the literature on theoretical equivalence, which concerns the formal relations that equivalent formulations of physical theories bear to each other. In her recent book, North (2021b) summarizes a position she has argued for:

> [T]wo theories can be *informationally equivalent* [i.e., equivalent descriptions] without being *metaphysically equivalent* ... In *one* way, the two are notational variants, in that each can be used to recover the same facts: they contain the same information, coded up in different ways. But in another, physically significant way, they are not *mere* notational variants: they present different pictures of the physical world. There is *both* a sense in which they contain all the same physics, *and* a sense in which they differ with respect to the physics... (North, 2021b, p. 213)

North's position seems well-aligned with Putnam's. We can use the metaphysical picture theory of commitment to vindicate it.

Let's see how. One example in this literature is affine plane geometry, which can be equally well formulated using just lines and not points, and using just points and not lines (Barrett and Halvorson, 2016). The rough idea is that any time one is tempted to speak of a point, they can replace that with talk of two intersecting lines. (Ditto for lines.) Each of the theories can define all of the vocabulary that the other employs, and once these definitions are added to the theories, then they entail precisely the same sentences. These two formulations are equivalent descriptions of affine plane geometry because they satisfy a particular formal relationship to each other (called Morita equivalence).

One sympathetic to North's picture can use this Putnam-inspired theory of commitment to say that these two formulations present different metaphysical pictures. One a geometric world containing only lines, the other just points. Here's how: on our present specification, a metaphysical picture is a theory formulated in first order logic, and appeal is determined by existential quantification. Geometry with points is a metaphysical picture that can explain the statements of ordinary affine plane geometry like Playfair's axiom. Same with geometry with lines. These both can explain Playfair's axiom equally well. All the same, these metaphysical pictures existentially quantify over different things, and so have different metaphysical implications. If we accept the picture of geometry with points, we do not ontologically commit to lines. Mutatis mutandis for the other theory and points. In this way, geometry with points and geometry with lines can be descriptively equivalent without being metaphysically equivalent, just as one sympathetic to North's project would want.[12]

[12] A further question: what are the equivalence conditions for metaphysical pictures? Why aren't geometry with lines and with points *the same* picture? A sketch of an answer on the present conception of pictures: the conditions for equivalence between metaphysical pictures is much stricter than the conditions for equivalence between *mathematical* theories. That is, these two theories might be equivalent according to the standards of mathematical theories, but not according to metaphysical standards.

As others in the theoretical equivalence literature develop views analogous to North's in this respect, an account of ontological commitment is needed. The Putnam-inspired view introduced here seems suited to such a task.

Chapter 6

The Pragmatics of Subtraction and Weaseling

Many philosophers believe that some of our best scientific theories literally entail metaphysical claims that they wish to reject. In the most discussed instance—the indispensability argument—our best physical theories entail the existence of mathematical objects like numbers, and yet many scientific realists are wont to accept the truth of these theories but reject the existence of mathematical objects. Putnam (1975, p. 74) reports this as the "intuitive" position. Those who find themselves in the situation where they want to accept some theory without accepting some entailment of that theory are faced with a choice. They must either replace the scientific theory with one that does not have the problematic entailment, or they must argue that they are permitted to simply affirm the original theory while rejecting the entailment. Following Colyvan (2010), the former is called the *hard road* and the latter is called the *easy road*. The most famous hard roader in the debate over the existence of numbers is Field (2016) (though see Tallant (2013) for a hard road response to a newer version of the indispensability argument). The hard road presumes that if we accept some theory, we accept all the entailments of that theory.

Easy roaders claim we can be selective in which consequences of a theory we are forced to accept. For instance, Yablo (1998) argues that the mathematical portions of our best scientific theories are best interpreted as a sort of make-believe. Accordingly, he thinks, we ought not think those portions of the theory are ontologically committing (cf. Yablo (2005, 2012, 2014) for more recent versions of his position). Most easy roaders follow Yablo's approach where they identify some feature of the mathematical portion of a theory that entails that this mathematical portion can be taken at less than face value.[1]

I want to discuss a more provocative type of easy roader: the weasel. The weasel claims that we can just *prune away* certain entailments of our theories. For instance, Melia (1995, 2000, 2002) proposes that we subtract numbers out of science:

> The force between two massive objects is *proportional* to the *product* of the masses *divided* by the square of the distance, except that numbers might not exist.

The weasel *might* offer some reason for why they can subtract the mathematical portion out. But, in general, the weasel simply thinks that in presenting some picture of the world, "it is [sometimes] legitimate to take back details that were asserted earlier" (Melia, 2000, p. 467).

To justify this pruning away, weasels appeal to the practice of linguistic subtraction, which is when we amend some utterance with something along the lines of "except that not *X*" or "except maybe not *X*", where *X* is some entailment of the original utterance. We understand what someone means when they say, for example, "All my students passed the class, except for two." So, too, the weasel continues, do we understand what someone means when they subtract numbers out of science. Weaseling essentially relies on this analogy between subtraction as a linguistic practice and pruning away entailments from scientific theories. Accordingly, much of the literature on weaseling concerns linguistic subtraction.

[1]See Balaguer (1998, Ch. 7), Azzouni (2012), Leng (2005, 2012), Bueno (2012), Saatsi (2011, 2016, 2017, 2020), Knowles and Saatsi (2019).

There are two main challenges regarding linguistic subtraction that must be met in order for the weasel to rely on it.

The first challenge has roots in an objection to weaseling. In particular, it is unclear whether one has communicated *anything* significant if we subtract numbers out of scientific theories. As Colyvan says, when the weasel subtracts too much, "we no longer have a grip on what is being said" (Colyvan, 2010, p. 295).[2] Accordingly, the first challenge is to determine what content remains after subtraction. What is left when we subtract numbers out of Newton's theory of universal gravitation? Without a good answer, it is thought, the weasel fails. In response to this challenge, philosophers of language have formulated partial content semantics to give precise articulations of the remaining content. Yablo (2012, 2014, 2018) takes this challenge head on (but see also Fine (2017b,a, 2018, 2020) and Hoek (2018)).

The second challenge concerning linguistic subtraction is the question of when linguistic subtraction is felicitous or permissible. One answer is that subtraction is permissible only if the remaining content is recoverable—only if we know what exactly was communicated after subtraction. The weasel, relying on the analogy, must then come up with a way of recovering the content of physical theories once we subtract the bit about mathematics.

But in this paper, I argue that there is another way to answer this second challenge by looking closely at the rules governing linguistic subtraction. I provide a purely pragmatic theory, which does not rely on partial content semantics, for predicting and explaining the permissibility of instances of subtraction. This theory follows from a corollary of Grice's maxim of relevance, which requires us

[2]Cf. Liggins (2012), Raley (2012), Knowles and Liggins (2015), Rayo (2015, §3.3), Dieveney (2020).

to make our conversational contribution relevant to the purpose at hand (Grice, 1975). In brief, my theory says that we may subtract content from some otherwise permissible utterance if and only if what we subtracted out was not crucially relevant.

This paper has a dual focus. First, I want to illuminate general conversational phenomena regarding subtraction. Second, since the weasel relies on the analogy with linguistic subtraction, I aim to understand the judge the prospects of weaseling in the mathematics case and more broadly. It turns out that context is crucial in determining when subtraction is permissible. Here, 'context' means the purpose of our scientific theories. As a result, the weasel must first settle the purpose of our scientific theories. Are scientific theories meant to be explanations of *purely empirical* phenomena, or are they meant to be explanations of *all natural* phenomena? Answering this question has profound impacts on the weasel's success.

The paper is structured as follows. In §2, I will lay out the linguistic phenomenon of subtraction, identifying the important cases discussed in the literature. In §3, I will introduce my principle of relevance that is meant to explain linguistic subtraction. This section contains the intuitive articulation of my strategy. In §4 and §5, I provide the formal pragmatic mechanism that will allow us to generate results using the principle of relevance, culminating in three tests for the permissibility of subtraction. Finally, in §6, I will apply the tests to three relevant cases in metaphysics of science.

6.1 Linguistic Subtraction

Instances of subtraction come in many varieties. Most often, an instance of subtraction is an utterance of a sentence or theory T, followed by the utterance of some subtraction clause like 'except for' or 'but' or 'minus', followed by the negation of an entailment of the sentence or theory q. We can formalize this as $T - q$. I will call T the 'original statement' and q the 'subtracted content'.

Some instances of subtraction are intuitively permissible, which I will indicate with a checkmark.

(1) ✓All my students passed the course, except that two did not.[3]

(2) ✓Triangles *a* and *b* are congruent, except that *a* and *b* are not the same size.[4]

(3) ✓Ellen wears the same hat that Sherlock Holmes does, except that Sherlock Holmes might not exist.[5]

(4) ✓Rob is six foot one, except that Rob might not be exactly between 6'.99" and 6'1.01".[6]

Each of these subtractions is a permissible contribution to the conversation.[7] From (1), we learn that it is generally permissible to subtract instances from a

[3] See von Fintel (1993), Melia (2000, p.467), Gajewski (2008), and Yablo (2014, p. 132).

[4] Yablo (2012, p. 1014), Yablo (2014, p. 158). Cf. Dan Hoek's "An autocracy is like an absolutist monarchy, except that the ruler need not be a king or queen."

[5] Hoek (2018). Cf. "A gratin is a quiche. Except it is not baked in a shell." (Fuhrmann, 1999, p. 566).

[6] Hoek (2018).

[7] Each of (1), (2), (3), and (4) has an available *paraphrase*; paraphrases, in the context of subtraction, are reformulations of the original statement that do not entail the part that one wishes to subtract. The paraphrases are as follows:

(1′) There are two students who did not pass my class, and every student who is not one of those two did pass my class.

(2′) Triangles *a* and *b* are similar.

generalization. From (2), we learn that it is generally permissible to subtract conjuncts from a conjunction. Here, the *concept* of congruence is a conjunction of similarity and sameness of size. It is difficult to draw general lessons from (3). From (4), we learn that it is generally permissible to subtract overly precise entailments from what we say.

Sometimes, the subtracted clause is 'except maybe not *q*' and other times it is 'except not *q*'. It seems that in the latter cases, like (1) and (2), the negation of what's being subtracted gets added into the remaining content. In (1), for example, the speaker isn't merely trying to distance themself from the implication that two students passed; rather, they are claiming that exactly two students did *not* pass. On the other hand, in the weaker subtraction cases like (3) and (4), the speaker seems to be merely distancing from the subtracted content. In (3), for example, the speaker seemingly wants to communicate some important content about Ellen's hat without committing to either the existence or nonexistence of Sherlock Holmes. Because my theory does not rely on generating the remaining content after subtraction, this subtlety will not affect the analysis given in this paper.

There are also many cases of subtraction that are obviously linguistically *impermissible.* These cases are nonsensical or straightforwardly contradictory in a way (1) - (4) are not. I will indicate these with a hashmark.

(3′) Ellen wears a deerstalker hat.

(4′) Rob is six foot one to the nearest inch.

Some have argued that an instance of subtraction is only permissible if there is an obvious paraphrase available (see, e.g., Colyvan (2010, pp. 295 - 296)). The idea is that (1) - (4) are okay because we *know* the paraphrases, and we are really accepting the paraphrase, not the subtraction. I disagree. However, my theory will entail that some instances of subtraction are guaranteed a paraphrase, so we can preserve at least that some subtractions come with paraphrases. See §5.1.

(5) #The tomato is scarlet, except that the tomato might not be red.[8]

(6) #There is an electron in the bubble chamber, except that electrons might not exist.[9]

(7) #I'm thirsty, except that nobody is thirsty.[10]

(8) #Snow is white, except it's not true that either snow is white or grass is green.[11]

(9) #The tomato weighs half a pound, except that the tomato doesn't weigh over an ounce.[12]

(10) #She danced badly, except she didn't dance at all.[13]

If someone utters any of these statements in an ordinary context, we get the intuition that they have said something infelicitous by subtracting out the entailment they did. Note that the original statement is an ordinary, and otherwise relevant, assertion in the course of a conversation. Because of this, we know that subtracting out the other content is what makes the entire utterance impermissible. It also does not seem that there is a pattern among these. In (5), we subtract out a determinable from a determinate. In (7), we subtract out a general existential entailment from the subject-predicate sentence. In (8), we subtract out any disjunctive entailment from the utterance of one disjunct. In all cases, though, one gets the intuition that these are impermissible because it is not clear how

[8] Yablo (2012, p. 1015), Yablo (2014, p. 155).
[9] Putnam (1975), Field (2016, p. 43), Field (1989, pp. 14 - 20), Dorr (2010).
[10] Yablo (2012, p. 1011), Hoek (2018). Cf Donnellan (1966).
[11] Yablo (2014, p. 158).
[12] Yablo (2014, p. 155).
[13] Yablo (2014, p. 155).

what the speaker says *could* contribute to the conversation, given what they've subtracted.[14]

Finally, there are contested cases of subtraction, which I indicate with a question mark.

(11) ?Willing your arm to go up is raising your arm, except that your arm may not go up.[15]

(12) ?A law-like generalization is a law, except that it might not be true.[16]

(13) ?Newtonian gravitation theory is true, except that mathematical objects may not exist.[17]

(14) ?The average star has 2.4 planets, except that average stars don't exist.[18]

(11) and (12) are instances where philosophers attempt to analyze some concept via subtraction. (13) and (14) are attempts to subtract out existential entailments from out scientific theories. Whether these are permissible subtractions is controversial.

My primary question concerns the cases under which some instance of subtraction is permissible or felicitous to utter within the context of a conversation.

[14]There are some cases in the literature that I think we can set aside for one reason or another. For example, Yablo (2014, p.155) considers the following sentence:

(i) I washed half as many tomatoes as you, except you didn't wash any tomatoes.

He claims that this is an impermissible instance of subtraction. It does not seem to me, however, that this is an instance of subtraction at all. For to subtract q from T, it must be that T entails q. And *I washed half as many tomatoes as you* does not entail *You washed some tomatoes*. The original statement surely *implicates* that you washed some tomatoes. But by denying that you washed tomatoes, we are cancelling an implicature, not subtracting an entailment.

[15]Wittgenstein (1953, 62), Yablo (2014, p. 134).

[16]Goodman (1955), Yablo (2014, p. 159).

[17]Melia (1995, 2000); Yablo (2014); Hoek (2018).

[18]Melia (1995), Yablo (1998, 2005).

The reigning strategy has been to use partial content semantics to argue for this. If we can define, e.g., the part about nobody's being thirsty from *I'm thirsty*, we can model what content remains when we subtract the former from the latter. In cases (5) - (10), the idea goes, no intelligible content remains after subtraction; accordingly, these are impermissible subtractions. I hope to show that we do not need this semantic mechanism to determine whether some instance of subtraction is permissible.

I will be assuming that, for any case of subtraction we assess, the original statement is relevant and would ordinarily be a permissible contribution in the conversation. For example, where (5) is impermissible to say, we are assuming that 'the tomato is scarlet' *would be* permissible to say. This is because, if the original statement is impermissible to utter by itself, our intuitions about the permissibility of the whole subtracted statement may be clouded by the impermissibility of the original statement.[19]

6.2 The Relevance Principle

My aim here is to provide a theory of the linguistic permissibility and impermissibility of the above examples of subtraction that can be analogized to weaseling. Given some instances of a linguistic phenomenon, the role of the theorist is to predict and explain that phenomenon in a way that integrates with other linguistic theories. As Yablo says about the variety of cases of subtraction, "it is not

[19]There are some subtleties here regarding what *would be* a permissible contribution. Consider (1). If it is true that all but two of my students passed the class, then it is strictly impermissible to utter 'All my students passed the class' because it violates the maxim to say true things. The way I am thinking of things, however, it would ordinarily be permissible to say 'All my students passed the class' in response to a question about how many students passed the class. We should understand 'would be permissible' as *would be permissible if the speaker did not want to distance themself from the subtracted entailment*.

obvious how to tell the good and bad cases apart" (Yablo, 2012, p. 1014). Reflecting on the above examples, it seems that subtraction is permissible (roughly) when what we subtract is not the essentially relevant content that we originally said. If what you subtract is all that was relevant to your original contribution, then we suspect that you didn't say something appropriate. If one subtracts out that the tomato is red from its being scarlet, has *anything* been said about the color of the tomato? Plausibly not, and so we cannot see how (5) contributes to the conversation.

To this end, I propose the following principle regarding subtraction:

Relevance. An instance of subtraction is permissible when and only when what you subtract is not the essential relevant content of your conversational contribution.

The idea is that we can predict and explain the permissibility of an instance of subtraction by examining how relevant the subtracted content is to the conversation. We do not need, I think, to examine the content left over after subtraction. In true Gricean fashion, this principle of relevance divides into several maxims. We must know whether the subtracted content is *relevant* and, if so, whether it is *essential*. There are, I will show, degrees of relevance for subtracted content. Some subtracted content is fully relevant, some partially relevant, and some not relevant at all. Each maxim corresponds to how relevant the subtracted content is.

Here I'll help myself to some intuitive understanding of relevance, and in §3 I will introduce a formal model of relevance. Here are the maxims:

Maxim 1. Do not subtract all of the relevant content from what you say.

Maxim 2. You may subtract out anything only partially relevant from what you say.

Maxim 3. Do not subtract anything from the entirely relevant things you say.

These maxims articulate the varieties of subtraction that are impermissible. If not deemed impermissible, then some instance of subtraction is permissible according to the principle of relevance. Let's see some of these in action.

Maxim 1 will predict the permissibility of (1) and (2), and the impermissibility of (5), (7), (8), and (9). The idea is this. Suppose someone asks how many students passed my class, and I answer with (1). Intuitively, the relevant content is anything that goes toward answering how many students passed my class. According to the first maxim, I may not subtract *all* of this relevant content. Accordingly, I am permitted to subtract some of the relevant content—namely, that two students passed the class. Thus, I can subtract *two students passed the class* from *all students passed the class*. On the other hand, suppose someone asks what primary color the tomato is, and I answer with (5). Intuitively, the relevant content is anything that goes towards answering what primary color the tomato is. According to this maxim, I may not subtract out all of this relevant content. In the case of (5), its redness is all the relevant content concerning the primary color of the tomato, and so I may not subtract the tomato's being red from its being scarlet.

Maxim 2 will predict the permissibility of (4). For technical reasons that we will shortly go into, Rob's exact height is merely partially relevant content. According to this maxim, then, we are permitted to subtract it. This maxim only applies in fairly rare circumstances.

Maxim 3 will predict the permissibility of (3), and the impermissibility of (6) and (10). The basic idea behind this maxim is that sometimes we utter something where its entire content is wholly relevant. If you ask me what is in the bubble chamber, and I respond, “There is an electron in the bubble chamber,” then intuitively the entirety of what I said is relevant. According to Maxim 3, we may not subtract any of the entailments out of these entirely relevant utterances. On the other hand, if you ask me what kind of hat Ellen wears, and I respond, “Ellen wears the same hat that Sherlock Holmes does,” then intuitively there are entailments of what I said that are not relevant to my answer. In particular, the existence of Sherlock Holmes is not relevant to what kind of hat Ellen is wearing. Accordingly, I may subtract Holmes’s existence out given Maxim 3.

I argue that the maxims correctly predict our intuitions on the permissibility of uttering (1) - (10), and they provide intuitively satisfying explanations of these permissibility judgments. Because of this, I think we can extend the principle of relevance to cases of weaseling in science.

6.3 Relevance as Questions Under Discussion

My principle of relevance states that we may subtract some content out just in case what we subtract is not the essential relevant content of what we said. In order to make this precise, we need an understanding of when content is relevant. I’ll be using the popular model of relevance Questions Under Discussion (QUD), introduced by Craige Roberts and developed by many for the past 25 years.[20]

Questions under discussion is a model of the relevant content for a conversation. A QUD is a set of questions which conversational participants are com-

[20] See Roberts (2012). See also the bibliography of QUD-related topics maintained by Roberts on https://www.asc.ohio-state.edu/roberts.21/QUDbib/.

mitted to resolving in a conversation. A question is a set of alternative possibilities corresponding in some way to the possible answers to the question (usually modeled with partitions of worlds). Relevant conversational contributions either provide an answer to the QUD or place another QUD on “the stack”. An answer to the QUD rules out at least one, but not all, of the possibilities. A full answer to the QUD rules in exactly one possibility by ruling out all other alternatives, and a partial answer eliminates at least one possibility.

Placing a question on the stack breaks the QUD up into more manageable chunks. Here I introduce degrees of stackability for questions. Some question is FULLY STACKABLE on the QUD just in case every answer to the question provides at least a partial answer to the QUD. Some question is PARTIALLY STACKABLE on the QUD just in case some (but not all) answers to the question provide at least a partial answer to the QUD. And a question is NOT STACKABLE on the QUD just in case there are no answers to the question that provide at least a partial answer to the QUD.

Suppose you and I are trying to determine the philosophy students’ quarterly grades to assess whether they are on track to graduate. The primary question on the stack is “What were the philosophy students’ grades?”, which induces a set of possibilities. These possibilities are mappings from each student to a grade in each class. Assuming there are 20 students and they take four classes each (with an A - F grading scale), this means there are 400 possibilities to the QUD. Figuring this out is difficult, and so you might place another question on the stack by asking “Well how many students passed your class?” Any answer to this question will start eliminating possibilities from the original QUD, so this question is fully stackable on the QUD. Let’s say I respond: “All my students passed.” Here, I have provided a full answer to the top-most question on the stack, eliminating 40 possibilities (those possibilities where any student failed all four classes). This model explains, among other things why it is that my conversational contribution was *relevant* to the context—because it provides at least a partial answer to the QUD.

6.3.1 Determining the Relevance of Subtracted Content

Before we can know whether what one subtracted was essential relevant content of what they originally said, we have to determine how relevant the subtracted content was. To do this, we use degrees of stackability. The purpose of this subsection is to map the phenomenon of subtraction onto the machinery of QUD.

First, we have to determine the original QUD. Recall my assumption that for any instance of subtraction $T - q$ we consider, the original statement T is a permissible contribution to the conversation if said on its own. This entails that T is relevant to the conversation. And given that T is relevant, it contributes at least a partial answer to the QUD. Accordingly, the first step is to determine the QUD that the original statement is answering. Consider again (1): All my students passed the course, except two. Here, the QUD is easy to reconstruct (in fact, it was explicitly asked): "How many students passed my course?" Call the QUD that the original statement T is relevant to the ORIGINAL QUD. It is likely that the original QUD can sometimes be ambiguous.

Second, we turn the subtracted content into a question. This process is straightforward. For (1), the subtracted content is that two student passed the course. The SUBTRACTED QUESTION is then "Did two students pass the course?"

Third, we ask how stackable the subtracted question is on the original QUD, and this tells us how relevant the subtracted content is. The subtracted question will be either fully relevant, partially relevant, or not relevant to the conversation; this directly tracks stackability. Let's take these in turn.

Some subtracted content q is FULLY RELEVANT just in case the subtracted question is fully stackable on the original QUD. That is, if any answer to the subtracted question entails at least a partial answer to the original QUD, then the subtracted content is fully relevant. For (1), the subtracted question is fully stackable on the original QUD because any answer to "Did exactly two students pass the course?" will provide at least a partial answer to "How many students passed the course?"

Some subtracted content q is PARTIALLY RELEVANT just in case the subtracted question is partially stackable on the original QUD. That is, if some but not all answers to the subtracted question entail at least a partial answer to the original QUD, then the subtracted content is partially relevant. Consider again (4): Rob is six foot one, except that his height is not between 6'.99" and 6'1.01" Suppose the original QUD is "How tall is Rob *to the nearest inch*?" In this case, the subtracted content is not fully relevant to the original QUD. The subtracted question is "Is Rob's height between 6'.99" and 6'1.01"?" Crucially, it is not the case that every answer to the subtracted question provides an answer to the original QUD. If the answer to the subtracted question is No, then we do not know anything about what Rob's height is to the nearest inch, for he still might be closest to 6'1" despite not being in that narrow range. Thus, the subtracted content is not fully relevant. Yet the subtracted content is still *partially* relevant because some answers to the subtracted question entail an answer to the original QUD. If the answer to the subtracted question is Yes, then Rob's height to the nearest inch *is*

6'1". Accordingly, in (4), the subtracted content is partially relevant.[21]

Some subtracted content q is NOT RELEVANT just in case the subtracted question is not stackable on the original QUD. That is, if no answers to the subtracted question entail at least a partial answer to the original QUD, then the subtracted content is not relevant.

Putting it all together, we have this:

Subtracted content relevance. Given an instance of subtraction $T-q$, an original QUD Q_1, and a subtracted QUD Q_2:

1. If Q_2 is fully stackable on Q_1, then q is wholly relevant (to the conversation),
2. If Q_2 is partially stackable on Q_1, then q is partially relevant (to the conversation), and
3. If Q_2 is not stackable on Q_1, then q is not relevant (to the conversation).

Knowing the relevance of the subtracted is necessary for determining whether it is permissible to subtract it. We now turn to the maxims in earnest.

6.4 The Maxims and Their Tests

The maxims are intuitively corollaries of the principle of relevance. Here, I will show that each maxim corresponds to a test within the QUD framework. It is

[21] Example (4) highlights how, sometimes, whether an instance of subtraction is permissible depends on what the original QUD is. In an ordinary context, the QUD is "How tall is Rob to the nearest inch?" In such a case, we deem the subtraction permissible. On the other hand, if the QUD is "Exactly how tall is Rob?", we intuitively think the subtraction is impermissible partly because the subtracted QUD *is* stackable on the original QUD. With both answers to whether Rob is between 6'.99" and 6'1.01", we learn something about the precise height of Rob.

these tests that will generate predictions about the permissibility of some particular instance of subtraction. I use these tests to vindicate the assessments of (1) - (10) and the permissibility of (14).

6.4.1 Maxim 1

We have the tools to determine whether the subtracted content is wholly *relevant* content. Maxim 1 determines whether that relevant content is also *essential.* If so, then the subtraction is impermissible; if not, then the subtraction is permissible. Moreover, if some instance of subtraction is deemed permissible by the test that corresponds to maxim 1, then—interestingly—it is guaranteed to have a paraphrase. Here's maxim 1:

Maxim 1 Do not subtract all of the relevant content from what you say.

Sometimes, we subtract relevant content, but it is intuitive that there is relevant content left over after subtraction. This is plausibly what is happening with (1). If somebody subtracts two students passing from all students passing, the question of how many students passed is still provided a univocal answer. Test 1 makes precise how these cases work:

Test 1: Given an instance of subtraction $T - q$ where q is wholly relevant, $T - q$ is permissible if and only if the subtracted content q, taken as a complete response, provides an answer to the original QUD that is incompatible with the answer T provides to the original QUD.

The basic idea behind this test is as follows. When we subtract some wholly relevant part of our original contribution, it had better not be all of the relevant content. To determine this, we see if the subtracted content, when taken as a complete response to the original QUD, leaves enough left over. If q's response to the original QUD is compatible with T's answer, then *either* we've subtracted everything out that was relevant, *or* we don't know how to interpret the subtraction of q.

Two questions are relevant for understanding this test. Why do we care if the answers are incompatible? And why are we taking q as providing a *complete* answer to the original QUD? To the first, recall that the subtracted content is an entailment of the original statement. If the answers provided by each to the QUD are incompatible, then that means they are not providing the same relevant content. They are saying different things about the QUD. And if we subtract from T some but not all of its relevant content, then there is relevant content left over. To the second question, in these tests, we consider taking q as a complete answer to the QUD because subtraction is always exacting. When we subtract 2 from 4, we subtract exactly 2, not *at least 2*. Likewise, when we subtract 'two students passed' from 'all students passed', we are subtracting exactly two students passing, not at least two students passing. Accordingly, we consider the subtracted content as a complete answer, imagining it with the proviso *and nothing else*.

Let's see how this test deems (1) permissible. Recall the original QUD is "How many students passed the class?", and the subtracted QUD is "Did two students pass the class?" Because the latter question is fully stackable on the former, the subtracted content is wholly relevant. This test tells us to examine the following. We consider the prospects of taking *Two students passed* as a complete answer, meaning that two and no more than two students passed. Thus, the subtracted content, as a complete answer to the original QUD, is *Exactly two students pass* *the class*. Then we consider the answer provided by *All the students passed* as a complete answer. These two are incompatible answers to the QUD. Accordingly, given Test 1, this instance of subtraction is deemed permissible.

Here we have a result that vindicates the intuition that subtraction sometimes comes with a paraphrase:

Result 1. If an instance of subtraction is permissible according to Test 1, then there is a paraphrase of that instance of subtraction.

Recall that a paraphrase is a reformulation of the original statement that does not entail the part that one wishes to subtract. Here's how this works. The original QUD for (1) can be answered by stating which students passed or failed the class. The original statement, that *all students passed*, provides a complete answer to this QUD. Think of this content as "ruling out" possible answers to the QUD—all those possible answers where any student failed. The subtracted content, *two students passed*, is a relevant entailment of the original statement that is distinct from the answer that the original statement gives. This subtracted content we should think of as "ruling back in" some of the possibilities ruled out by the original statement. Namely, those possibilities where just two students failed. Thus, we can recover the entirety of the content of (1) by considering what remained ruled out after the procedure of ruling back in those answers from the subtracted content. Because the answers that only two students passed and that all students passed are distinct, we are guaranteed to have a univocal answer to the original QUD left over after ruling in. In this way, we are guaranteed a paraphrase, even if it as cumbersome as articulating exactly those possibilities that remain after the ruling back in.

It makes sense that any instance of subtraction that passes Test 1 is guaranteed a paraphrase. When we subtract *relevant* content that is *distinct* from our original contribution, intuitively there's some relevant content left over. The paraphrase is then just this relevant content that is left over, and we have a procedure for generating this using the QUD model. This result vindicates the above thought that some permissible instances of subtraction come with a paraphrase. This is because those instances pass Test 1, and thus we are justified in believing that there is a paraphrase. We do not need to know the paraphrase itself in order to be justified in believing that there is a paraphrase.

Test 1 also predicts the impermissibility of those instances of subtraction where, intuitively, we have subtracted *all* the relevant content from the original statement. Consider an ordinary instance of (5): The tomato is scarlet, except the part about the tomato being red. By an *ordinary* instance, I mean one where the QUD is "What primary color is the tomato?" The subtracted content is that the tomato is red, and the subtracted QUD is "Is the tomato red?" This subtracted QUD is fully stackable on the original QUD because any answer to it will provide at least a partial answer to what primary color the tomato is. Note the answers that both the original statement and the subtracted content provide to the original QUD: 'The tomato is scarlet' provides the answer to the primary color question that the tomato is red; likewise, 'The tomato is red' provides the answer to the primary color question that the tomato is red. These answers are compatible; in fact, they're identical. The test determines that this subtraction is impermissible because you've subtracted everything relevant from what you originally said.

Moreover, Test 1 predicts why it is generally impermissible to subtract *gener-*

alizing entailments from what we say. Consider the following statement: All of my students passed the class, except that some students did not pass. Intuitively, this is impermissible; it feels more straightwardly contradictory than (1). Here, the subtracted QUD is fully stackable on the original QUD (assuming it is "How many students passed the class?"). Now, consider whether *some students passed*, if taken as a complete response to the original QUD, is compatible with *all students passed*. Yes it is. One way that some students pass is if all students passed. Accordingly, this test says this instance of subtraction is impermissible. Similar verdicts are reached with (7), (8), and (9) under certain QUDs. If the QUD is "Who is thirsty?", then (7) will fail Test 1. If the QUD is "What color is snow?", then (8) will fail this test. If the QUD is "What is the weight of the tomato?", then (9) will fail this test.[22]

6.4.2 Maxim 2

Let us now consider those cases where the subtracted content is only partially relevant. Maxim 2 tells us this:

Maxim 2. You may subtract out anything only partially relevant from what you say.

Maxim 2 tells us that for any partially relevant entailment of what we say, we are permitted to subtract it out. This seems striking. We'll see, though, that such cases are relatively rare. And once we appreciate what it takes for some content to be partially relevant, we will see why all such content may be subtracted. Here is the corresponding test:

[22]See also Dan Hoek's "The number of planets is seven, except that the number of planets is not prime." This case is deemed impermissible by Test 1 as well as well.

Test 2. Given an instance of subtraction $T - q$ where q is partially relevant, $T - q$ is permissible.

Let us first note that this test predicts the cases of partial relevance subtraction considered so far, which is only example (4). Whether Rob is between 6'.99" and 6'1.01" is partially stackable on Rob's height to the nearest inch—because an affirmative answer, but not a negative, provides at least a partial response to the question of Rob's approximate height. Accordingly, the test deems (4) permissible.

As far as I can tell, the only cases of partially relevant subtractions come from generality mismatches like subtracting precision out of a context of loose talk. Suppose someone said, "The road is flat, except that it has some tiny bumps." The original QUD is likely "Is this road closer to bumpy or flat?", and the subtracted QUD is "Does this road have some tiny bumps?" Though a negative answer to the subtracted QUD provides an answer to the original QUD, a positive answer does not. Thus, we have a case of partial relevance. Accordingly, Test 3 yields that this instance of subtraction is permissible. In contexts of loose talk, we can always subtract out the irrelevantly precise entailments of what we say. This seems correct. In fact, it is plausible that we do not even hold speakers to be committed to these entailments (Hoek, 2018, 2019).

We might think that existential subtractions fall under the Partial Relevance Test. They do not, but it is instructive to see why. The thought is this. Consider the following subtraction: Electrons are negatively charged, except there are no electrons. At first glance, we might think this is a case of partially relevant subtraction because the subtracted QUD, "Are there electrons?", has one but not all answers that bear on the original QUD. If the answer is "No", then the original

QUD "What charge do electrons have?" has alternatives ruled out. In fact, it has all alternatives ruled out. Accordingly, the thought goes, this is a case of partial relevance and the test wrongly deems it permissible.

This is not the result that the test delivers. Note that partial relevance is not determined by whether one (and only one) answer to the subtracted QUD *bears on* the original QUD. Rather, partial relevance is determined by whether exactly one answer to the subtracted QUD provides an answer to the original QUD. A complete answer to a QUD, by definition, rules in exactly one alternative by ruling out all others. Accordingly, if one answer to the subtracted QUD rules out *all* alternatives to the original QUD, then it does not provide an answer, since there must remain at least one possibility in an answer. In this case, saying there are no electrons does not provide a partial (or complete) answer to the question "What charge do electrons have?" So this case does not fall under Maxim 2.

6.4.3 Maxim 3

Finally, we come to the last maxim, the one that is most central to weasels. The cases left to be predicted are (3), (6), and (10).

Here is Maxim 3:

Maxim 3. Do not subtract anything from the entirely relevant things you say.

There is also a test corresponding to Maxim 3:

Test 3: Given an instance of subtraction $T - q$ where the subtracted QUD is not relevant, $T - q$ is impermissible if and only if T directly answers the original QUD.

The idea here is that, when we directly answer the QUD by uttering some sentence, all of the non-relevant entailments of what we said are important. (This is perhaps guaranteed by Grice's Maxim of Quantity, which tells us to say only exactly what we need to.) All these entailments are intuitively going into the way that *T* answers the QUD. We cannot then subtract any non-relevant entailments. A statement directly answers a QUD just in case that statement provides an answer no non-trivial inferences. For example, if someone asks "What did you buy at the store?", a direct answer would list items I bought at the store; a non-direct answer might be "I bought the items that were on the list." Exactly what constitutes a trivial inference plausibly depends on the context.

Let's see how this test predicts that (6) is impermissible. Plausibly, the relevant QUD is "What is in the bubble chamber?" or "What is causing the bubbles to appear in the bubble chamber?" The original statement is that there is an electron in the bubble chamber. This directly answers the QUD in the sense that there are no inferences that need to be made to have an answer to the QUD. Accordingly, any of the entailments of *T* that are not relevant to the original QUD are deemed essential; we cannot use *T* to answer the QUD without those entailments. Accordingly, we cannot subtract out the existence of electrons from the statement that there is an electron in the bubble chamber. A similar diagnosis can be made for (10).

We can also see how it predicts that (3) is permissible. Whether Sherlock Holmes exists is not relevant to the conversation when the QUD is "What kind of hat does Ellen wear?" Accordingly, the subtracted content is not relevant, and (3) meets the conditions for applying the test. Test 5 tells us that if the original statement directly answers the QUD, then we may not subtract any content that's not relevant. Plausibly, 'Ellen wears the same hat that Sherlock Holmes does' does not directly answer the QUD. In order for that statement to answer the QUD, you must also know what sort of hat Sherlock Holmes wears. This is a non-trivial inference that a speaker must make in order for the original statement to answer the QUD. Accordingly, this test says that (3) is permissible.

The controversial cases of subtraction are ones that fall under this maxim, where we are trying to subtract out existential entailments of what we say. Much ink has been spilled over determining whether any non-defective content remains after subtracting numbers out of Newtonian gravitation theory, in the effort to determine whether the subtraction is okay. If this test is correct, we need only determine whether Newtonian gravitation theory directly answers the relevant QUD. More on these questions in a moment.

6.5 Subtracting metaphysics out of science

Here, I will apply my tests to two cases of interest, and then I will draw one broader conclusion.

6.5.1 On average stars

Here I'll show that (14) is a permissible instance of subtraction.

In the context of trying to determine some facts about astronomy, we abut the question of how many orbiting planets and stars there are. Suppose, as Melia (1995, pp. 226 - 227) suggests, "there are precisely twentyfour zillion orbiting planets and ten zillion stars". For good reasons, we will likely never be justified in believing this exact specification, and so our best scientific theories will never entail the precise number of planets and stars. But scientists aren't silent on the matter; they know something about the ratio of planets to stars. In particular, they know that what 'the average star has 2.4 planets' says about how many orbiting planets and stars there are is true. However, 'the average star has 2.4 planets' entails the existence of the average star, which the ontologically squeamish might be disinclined to commit to.[23] Accordingly, one might put forward (14):

(14) The average star has 2.4 planets, except that average stars don't exist.

The question is whether this subtraction is permissible.

Let's use the tests. We first note that the subtracted content is an existential entailment of the original statement, which suggests that this will fall under Maxim 3. Nonetheless, we ought to determine how relevant the subtracted content is to the original QUD. In the context described in the previous paragraph, the original QUD is something like "How many planets and stars are there?" (or "Relatively how many planets and stars are there?"). The subtracted content is that the average star exists, and the subtracted QUD is "Does the average star exist?" Note that this subtracted QUD is not stackable on the original QUD. There is no answer to this question that would provide an answer to the question of how many planets and stars there are. Accordingly, the subtracted content is not relevant to the original QUD. This tells us that we should use Test 3 to determine whether (14) is impermissible.

According to the test, (14) is impermissible if and only 'The average star has 2.4 planets' directly answers the question "How many planets and stars are there?" It does not seem as if it does directly answer this question. There are nontrivial inferences that one must make in order to an answer for how many (or

[23]Here I am assuming, contra Kennedy and Stanley (2009), that 'The average star' is a singular noun phrase.

relatively how many) orbiting planets and stars there are. In particular, we must infer some facts about the numbers of actual concrete stars and planets from the putative existence of some average star. Accordingly, Test 3 deems (14) as a case of permissible subtraction. And, purportedly, the idea goes, accepting (14) does not ontologically commit one to the existence of average stars. We would need a theory of ontological commitment to vindicate this, but let's assume that one is supplied.

One might worry that I've gerrymandered the QUD to make (14) pass the test. For if the QUD is "How many planets does the average star have?", which an inquisitive child might ask an astronomer, then it seems that "The average star has 2.4 planets" *does* directly answer the QUD. In such a case, (14) is deemed impermissible by the Non Relevance Test.

I would agree that subtraction is not permissible in this revised context. It would be odd for the astronomer to flat-footedly utter (14). However, I think something different else is going on that would explain its infelicity. When an inquisitive child is asking an astronomer how many planets the average star has, the astronomer is under no false belief that average stars exist. The astronomer will not stop the conversation and have a frank discussion about the nonexistence of average entities. Rather, the astronomer will operate under the pretense—or something similar to pretense—that the average star exists. Accordingly, they are not within the realm of inquiry, and their utterance of 'The average star has 2.4 planets' should not be interpreted as contributing to the global QUD ("What is the world like?"). Perhaps the utterance should be interpreted as temporarily accommodating the presupposition that there is an average star.[24] The astronomer

[24]Cf. Lewis (1979) and Stalnaker (2002) on presupposition accommodation.

might discharge the accommodated presupposition at a later point, but this plausibly can't be done with a simple subtraction clause. They may have to say something like "To be clear, there are no average stars, but everything I said about how many actual planets and stars there are is true." Or they may leave it unsaid.

So much for average stars. There are still contested cases, like (11), (12), and (13). One might think that my account here allows us to easily determine whether these instances of subtraction work. While I will have some things to say about (13) below—the weasel's main project—I do not think my solution works for subtraction in robust philosophical analysis, like trying to subtract truth out of laws. This is because it is not entirely clear to me that there is a relevant QUD to assess permissibility. Whether this is a limitation depends on one's optimism regarding the success of analyzing philosophical concepts via subtraction.

6.5.2 On subtracting out of empirical consequences

There are instances where philosophers infer some metaphysical consequence from the empirical consequences of our scientific theories. Here are three examples. Hofweber (2016, p. 199) argues that a theory's ordinary observations and predictions entail the existence of composite objects. Blatti (2012) argues that evolutionary theory's empirical consequences entail that human persons are organisms. Finally, many philosophers have argued that some consequences about the life-span of cicadas entails that numbers exist (see, e.g., Bangu (2008)). Here, I'll consider this third example in detail. I will show that we can subtract the existence of numbers from a theory's empirical consequences. I take this to show that empirical consequences are relatively metaphysically neutral. As a result, we ought not infer metaphysical theses from empirical consequences.

The enhanced indispensability argument claims that there are instances of purely physical phenomena that require the existence of numbers to explain. Here is one purported instance. There is a certain subspecies of periodical cicada in North America that spends 13 years underground in a larval stage before hatching. That this year-length is irregular calls out for explanation. Biologists suggest that a 13-year length life cycle will minimize overlap with predators and other subspecies. The following calls out for explanation:

(d) The length (in years) of the life cycle of periodical cicadas is 13.

Why does a 13-year length life cycle minimize overlap with predators and competing subspecies? The proposed answer is that these are prime-numbered periods, and it is a mathematical law that prime-numbered periods minimize overlap. Accordingly, the fact that the year length of this life cycle is prime explains why it is evolutionarily advantageous (given the ecological constraints of the cicada—its predators, competitors, and facts about its environment). Primeness here plays a genuine explanatory role, so claims the enhanced indispensability argument. This is supposedly an instance of a genuine mathematical explanation of a purely physical phenomenon.

Many philosophers have worried that this argument is circular.[25] The idea is that datum (d) is already ontologically committed to numbers because it quantifies over the number 13. The circularity objection has force in the literature, and Baker (2009, 620 - 622) has taken it head on by offering a paraphrase of (d) that does not quantify over numbers. (See more recently Baker (2021).) His paraphrase is a first-order logic variant:

[25] See Baker (2005, 223) Baker (2009, 620 - 622), Leng (2005, 174), Ioan Bangu (2008), Rizza (2011, 105 - 106), Baron (2014, 473 - 476), Panza and Sereni (2016), Barrantes (2019, 252), Carbonell (2020), and Heylen and Tump (2020).

(d^*) $\exists x_1 \ldots \exists x_{13}(Fx_1 \wedge \cdots \wedge Fx_{13} \wedge x_1 \neq x_2 \wedge \cdots \wedge x_{12} \neq x_{13} \wedge \forall x_{14}(Fx_{14} \leftrightarrow (x_{14} = x_1 \vee \cdots \vee x_{14} = x_{13})))$

(d^*) is meant to be a paraphrase of (d) that is empirically equivalent but does not quantify over numbers. Baker in this way eliminates numbers from (d). Recently, though, some have criticized this paraphrase as not being adequate. Heylen and Tump (2020) argue that it becomes unclear how *primeness* enters into the explanation of (d^*) at all. In the case of (d), it was clear how the fact that prime periods minimize overlap explains how the year-length of a cicada's life cycle would be 13. But it is less clear how the fact that prime periods minimize overlap can explain (d^*). If Heylen and Tump's objection is sound, then Baker's solution to the circularity problem is in trouble. This circularity objection seems to remain for the indispensability theorist.

However, if we can subtract away the part about the number 13 existing, then the circularity objection dissolves. Let us see whether (15) passes the tests for subtraction:

(15) The length (in years) of the life cycle of periodical cicadas is 13, except that the number 13 might not exist.

The first step, as always, is to clarify what the relevant QUD is. Here, the relevant QUD seems to be "How long do periodical cicadas live?" In evolutionary biology, biologists are concerned with the phenotypes of organisms, as studying these teaches us about how evolution shaped them. Accordingly, the QUD is entirely about some ordinary empirical goings on: what is this organism's life like. It is true that in (d) abstracta are invoked to answer this QUD. But whether the number 13 exists does not go any way toward answering the QUD. The subtracted content is not relevant and we consult Test 3.

According to this test, (15) is impermissible if and only if 'the length (in years) of the life cycle of periodical cicadas is 13' directly answers the question "How long do periodical cicadas live?" It does not seem as if this directly answers the question. We must infer how long cicadas live from something (a lifespan) equalling the number 13. Accordingly, (15) is a permissible instance of subtraction. If I am right, there is no circularity objection to the enhanced indispensability argument. Baker (et al.) is not required to give a number-free variant of (*d*) in order to argue that we need primeness to explain the life cycle of cicadas. The content in (*d*) that entails the existence of numbers is superfluous, and we can subtract it out.

We see here that subtraction is a guide to the representationally significant content of a scientific theory. If there is some part of a theory that we can subtract, as with the number 13 from (*d*), it seems that that part was not representationally significant in the first place.

6.5.3 On weaseling

Let us turn to Melia's weaseling strategy in earnest. The tests I introduced for determining when an instance of subtraction is permissible indicate a way forward for flat-footed easy road strategies like Melia's where we simply take back the part about numbers. The first step in any attempted subtraction is to determine the relevant context. For the tests, these were formalized within the QUD model of at-issue content. It might be that this model can be extended to the level of our scientific theories writ large. This would be a significant project. But it seems we need not go so far as that.

Instead, we can determine the at-issue or relevant content for scientific theories by determining what the *aims* of science are. Knowing the aims of our scientific theories will tell us what is relevant content, and thus what can be subtracted.

What, at the most general level, are the aims of scientific theories? One reasonable answer is that scientific theories aim to explain the natural concrete phenomena we observe. If this is the broad aim of our scientific theories, then given any particular theory with an entailment that is irrelevant to explaining the natural concrete phenomena we observe, it seems that we can subtract that entailment given maxim 3. The weasel then must argue that the parts about numbers are irrelevant to explaining the natural concrete phenomena.

If, on the other hand, the aim is to explain natural phenomena—concrete and otherwise—then the weasel's project is much more difficult. The weasel in this case must show that the parts about numbers are irrelevant to explaining the natural phenomena.

These two different aims of scientific theories are best shown with an example. Consider again Baker's enhanced indispensability argument, which claims that *primeness* is indispensable for explaining why cicadas have life spans of 13 years. Consider first the aim of science as explaining the concrete natural world. On this approach, the aim of our scientific theories is to provide explanations of the concrete phenomena, and anything else is irrelevant. The concrete phenomenon here is that cicadas have 13 year-lengthed life spans. The reigning explanation, according to Baker, is that cicadas have 13 year-lengthed life spans because 13 years is a prime-lengthed period, and prime-lengthed periods minimize overlap with predators. Even in this short explanation, we see that what is directly explaining the QUD is that the length of time for cicadas' life spans is that it minimizes overlap with predators.[26] Accordingly, whether this period is prime is not part of what makes the explanation work. Given our sufficiently generalized tests, this suggests that we can subtract the part about primeness out of the cicada explanation if the aim of science is to explain just the concrete.

On the other hand, suppose the aim of our scientific theories is to provide explanations of nature, concrete and otherwise. Again, we have an explanation of the life cycle of cicadas in terms of both primeness and overlap-minimizing periods of time. The initial QUD "Why do cicadas have 13-year lengthed lifespans?" is answered by *because such a lifespan is overlap minimizing*. And while this is the end of the story for the concrete aims of science, for the broader aim of science, this explanation begets another question: Why is a 13-year lengthed period overlap minimizing? And the answer to this question has to do with primeness in the way that Baker explains (Baker, 2017). If this question is within the domain of questions that a scientific theory must answer, then it seems that we cannot simply subtract out the part about primeness, since primeness is essential to answering a QUD that is not present in the other aim of science.

Settling the aims of science, whether to explain simply the concrete natural world or everything about the natural world, can aid us in assessing the tenability of the weaseling project. If the aim of science is only to tell us about the concrete world, then it may be that we may subtract out any part of science that is not relevant to answering questions about the concrete world. The preceding has been quite sketchy, but we can see the kind of argument a weasel would give for broad-strokes subtraction. Such a reorientation of the debate over weaseling would begin to clarify the crux of the debate, which has increasingly become

[26] Cf. Tallant (2013).

about whether the *types* of explanations present in mathematical cases is metaphysically committing. Better, perhaps, to settle what needs to be explained.

6.6 Conclusion

Recall that the weasel faces a challenge: if weaseling is supposed to be analogous to linguistic subtraction, then we have to know when linguistic subtraction is permissible to know when weaseling is permissible. By reflecting on ordinary instances of subtraction, we see that it is permissible when what one subtracts is not essentially relevant content. And, by extension, it seems that weaseling is permissible when what one weasels away is not essentially relevant scientific content. I have argued that for some parts of our theories—the empirical consequences—weaseling is quite permissive. I have also shown that the success of the project that Melia began with, of subtracting numbers out of science, depends in part on what the aims of our scientific theories are.

Bibliography

Aizawa, K. (2022). Polger and shapiro on realization and multiple realization. *Review of Metaphysics*, 76(2):325–344.

Aizawa, K. and Gillett, C. (2016). *Scientific Composition and Metaphysical Ground*. Palgrave-Macmillan.

Antony, L. (2003). Who's afraid of disjunctive properties? *Philosophical Issues*, 13(1):1–21.

Audi, P. (2012). Properties, powers, and the subset account of realization. *Philosophy and Phenomenological Research*, 84(3):654–674.

Audi, P. (2013). How to rule out disjunctive properties. *Noûs*, 47(4):748–766.

Azzouni, J. (2004). *Deflating Existential Consequence: A Case for Nominalism*. Oxford, England: Oup Usa.

Azzouni, J. (2012). Taking the easy road out of dodge. *Mind*, 121(484):951–965.

Bagwell, J. N. (2021). Eliminativism and evolutionary debunking. *Ergo: An Open Access Journal of Philosophy*, 8:496–522.

Bailey, A. M. and Brenner, A. (2020). Why composition matters. *Canadian Journal of Philosophy*, 50(8):934–949.

Baker, A. (2005). Are there genuine mathematical explanations of physical phenomena? *Mind*, 114(454):223–238.

Baker, A. (2009). Mathematical explanation in science. *British Journal for the Philosophy of Science*, 60(3):611–633.

Baker, A. (2017). Mathematics and explanatory generality. *Philosophia Mathematica*, 25(2):194–209.

Baker, A. (2021). Circularity, indispensability, and mathematical explanation in science. *Studies in History and Philosophy of Science*, 88:156–163.

Baker, A. and Colyvan, M. (2011). Indexing and mathematical explanation. *Philosophia Mathematica*, 19(3):323–334.

Baker, L. R. (2003). Review of objects and persons. *Australasian Journal of Philosophy*, 81(4):597.

Baker, L. R. (2007). *The Metaphysics of Everyday Life: An Essay in Practical Realism*. Cambridge University Press.

Balaguer, M. (1998). *Platonism and Anti-Platonism in Mathematics*. New York: Oxford University Press.

Banerjee, S. K. and Moskowitz, B. M. (1985). Ferrimagnetic properties of magnetite. In *Magnetite Biomineralization and Magnetoreception in Organisms*, volume 5. Springer.

Bangu, S. I. (2008). Inference to the best explanation and mathematical realism. *Synthese*, 160(1):13–20.

Barker, J. (2020). Debunking arguments and metaphysical laws. *Philosophical Studies*, 177(7):1829–1855.

Baron, S. (2014). Optimisation and mathematical explanation: doing the Lévy Walk. *Synthese*, 191(3):459–479.

Baron, S. (2016). The explanatory dispensability of idealizations. *Synthese*, 193(2):365–386.

Baron, S. (2021). Unification and mathematical explanation in science. *Synthese*, pages 1–25.

Barrantes, M. (2019). Optimal representations and the Enhanced Indispensability Argument. *Synthese*, 196(1):247–263.

Barrett, T. W. (2015). Spacetime Structure. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 51:37–43.

Barrett, T. W. (2019). Equivalent and Inequivalent Formulations of Classical Mechanics. *The British Journal for the Philosophy of Science*, 70(4):1167–1199.

Barrett, T. W. (2020a). How to count structure. *Noûs*, 56(2):295–322.

Barrett, T. W. (2020b). On putnam's proof of the impossibility of a nominalistic physics. *Erkenntnis*, 88(1):1–28.

Barrett, T. W. and Halvorson, H. (2016). Glymour and {Q}uine on Theoretical Equivalence. *Journal of Philosophical Logic*, 45(5):467–483.

Bas, C. V. F. (1980). *The Scientific Image*. New York: Oxford University Press.

Bickle, J. (2020). Multiple Realizability. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2020 edition.

Bickle, J. W. (1998). *Psychoneural Reduction: The New Wave*. Bradford.

Bird, A. (1998). Dispositions and antidotes. *Philosophical Quarterly*, 48(191):227–234.

Bird, A. (2007). *Nature's Metaphysics: Laws and Properties*. Oxford, GB: Oxford University Press.

Bird, A. (2016). Overpowering: How the powers ontology has overreached itself. *Mind*, 125(498):341–383.

Blatti, S. (2012). A new argument for animalism. *Analysis*, 72(4):685–690.

Block, N. (2014). Seeing-As in the Light of Vision Science. *Philosophy and Phenomenological Research*, 89(3):560–572.

Boyce, K. (2018). Why inference to the best explanation doesn't secure empirical grounds for mathematical platonism. *Synthese*, 198(1):1–13.

Bradley, C. (2021). The non-equivalence of einstein and lorentz. *British Journal for the Philosophy of Science*, 72(4):1039–1059.

Brenner, A. (2015). Mereological nihilism and the special arrangement question. *Synthese*, 192(5):1295–1314.

Brenner, A. (2018). Science and the special composition question. *Synthese*, 195(2):657–678.

Brenner, A. (2023). Sense perception and mereological nihilism. *forthcoming in Philosophical Quarterly*.

Brown, J. D. K. (2015). Chemical atomism: A case study in confirmation and ontology. *Synthese*, 192(2):453–485.

Bryant, A. (2018). Naturalizing grounding: How theories of ground can engage science. *Philosophy Compass*, 13(5):e12489.

Bueno, O. (2012). An easy road to nominalism. *Mind*, 121(484):967–982.

Bueno, O. (2013). Putnam and the indispensability of mathematics. *Principia: An International Journal of Epistemology*, 17(2):217.

Bueno, O. (2018). Putnam's indispensability argument revisited, reassessed, revived. *Theoria*, 33(2):201–218.

Burge, T. (2010). *Origins of Objectivity*. Oxford University Press.

Burgess, J. (2018). Putnam on Foundations: Models, Modals, Muddles. In Hellman, G. and T.~Cook, R., editors, *Hilary Putnam on Logic and Mathematics*. Springer.

Burgess, J. P. (1984). Synthetic mechanics. *Journal of Philosophical Logic*, 13(4):379–395.

Burgess, J. P. and Rosen, G. (1997). *A Subject With No Object: Strategies for Nominalistic Interpretation of Mathematics*. Clarendon Press: Oxford.

Byrne, A. (2019). Perception and ordinary objects. In Cumpa, J. and Brewer, B., editors, *The Nature of Ordinary Objects*. Cambridge University Press.

Byrne, A. and Manzotti, R. (2022). Hallucination and its objects. *Philosophical Review*, 131(3):327–359.

Cameron, R. P. (2007). The contingency of composition. *Philosophical Studies*, 136(1):99–121.

Cameron, R. P. (2010). Quantification, Naturalness, and Ontology. In *New Waves in Metaphysics*, pages 8–26. Palgrave Macmillan UK.

Carbonell, F. M. (2020). Epistemic projects, indispensability, and the structure of modal thought. *Res Philosophica*, 97(4):611–638.

Carnap, R. (1956). The methodological character of theoretical concepts. *Minnesota Studies in the Philosophy of Science*, 1(1):38–76.

Castro, E. (2013). Defending the indispensability argument: Atoms, infinity and the continuum. *Journal for General Philosophy of Science / Zeitschrift für Allgemeine Wissenschaftstheorie*, 44(1):41–61.

Chang, L., Roberts, A. P., Tang, Y., Rainford, B. D., Muxworthy, A. R., and Chen, Q. (2008). Fundamental magnetic parameters from pure synthetic greigite (fe3s4). *Journal of Geophysical Research-Solid Earth*, 113.

Clapp, L. (2001). Disjunctive properties. *Journal of Philosophy*, 98(3):111–136.

Clarke-Doane, J. (2020). *Morality and Mathematics*. Oxford, England: Oxford University Press.

Colyvan, M. (2001). *The Indispensability of Mathematics*. Oxford University Press.

Colyvan, M. (2002). Mathematics and Aesthetic Considerations in Science. *Mind*, 111(441):69–74.

Colyvan, M. (2008). Indispensability arguments in the philosophy of mathematics. In Zalta, E. N., editor, *Stanford Encyclopedia of Philosophy*.

Colyvan, M. (2010). There is no easy road to nominalism. *Mind*, 119(474):285–306.

Contessa, G. (2014). One's a Crowd: Mereological Nihilism without Ordinary-Object Eliminativism. *Analytic Philosophy*, 55(2):199–221.

Craig, W. (1953). On axiomatizability within a system. *Journal of Symbolic Logic*, 18(1):30–32.

Craig, W. (1956). Replacement of auxiliary expressions. *Philosophical Review*, 65(1):38–55.

Curiel, E. (2014). Classical mechanics is Lagrangian; it is not Hamiltonian. *The British Journal for the Philosophy of Science*, 65(2):269–321.

Daly, C. and Langford, S. (2009). Mathematical Explanation and Indispensability Arguments. *The Philosophical Quarterly*, 59(237):641–658.

Dieveney, P. (2020). The honest weasel a guide for successful weaseling. *Disputatio*, 12(56):41–69.

Donnellan, K. S. (1966). Reference and definite descriptions. *Philosophical Review*, 75(3):281–304.

Dorr, C. (2002). *The Simplicity of Everything*. PhD Thesis: Princeton.

Dorr, C. (2010). Of numbers and electrons. *Proceedings of the Aristotelian Society*, 110(2pt2):133–181.

Dretske, F. (1997). Naturalizing the mind. *Noûs*, 31(4):528–537.

Eklund, M. (2005). Fiction, indifference, and ontology. *Philosophy and Phenomenological Research*, 71(3):557–579.

Field, H. (2016). *Science Without Numbers*. Oxford University Press, second edi edition.

Field, H. H. (1989). *Realism, Mathematics & Modality*. Blackwell.

Fine, K. (2017a). A Theory of Truthmaker Content I: Conjunction, Disjunction and Negation. *Journal of Philosophical Logic*, 46(6):625–674.

Fine, K. (2017b). A Theory of Truthmaker Content II: Subject-matter, Common Content, Remainder and Ground. *Journal of Philosophical Logic*, 46(6):675–702.

Fine, K. (2018). Truthmaking and the is-ought gap. *Synthese*, 198:887–914.

Fine, K. (2020). Yablo on subject-matter. *Philosophical Studies*, 177(1):129–171.

von Fintel, K. (1993). Exceptive constructions. *Natural Language Semantics*, 1(2):123–148.

Fish, W. (2009). *Perception, Hallucination, and Illusion*. Oxford University Press.

Fodor, J. A. (1984). Observation reconsidered. *Philosophy of Science*, 51(March):23–43.

Friedman, M. (1983). *Foundations of space-time theories: Relativistic physics and philosophy of science*. Princeton University Press.

Fuhrmann, A. (1999). When hyperpropositions meet . *Journal of Philosophical Logic*, 28(6):559–574.

Gajewski, J. (2008). Npi any and connected exceptive phrases. *Natural Language Semantics*, 16(1):69–110.

Gillett, C. (2016). *Reduction and Emergence in Science and Philosophy*. Cambridge University Press.

Goodman, N. (1955). *Fact, Fiction, and Forecast*. Harvard University Press.

Goodman, N. (1957). Replacement of auxiliary expressions. *Journal of Symbolic Logic*, 22(3):317–318.

Grice, H. P. (1975). Logic and conversation. In Davidson, D. and Harman, G., editors, *The Logic of Grammar*, pages 64–75.

Hawthorne, J. (2006). *Metaphysical Essays*. Oxford University Press.

Healey, R. (2013). Physical composition. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 44(1):48–62.

Heil, J. (1999). Multiple realizability. *American Philosophical Quarterly*, 36(3):189–208.

Heil, J. (2003). *From an Ontological Point of View*. Oxford University Press.

Heil, J. (2005). Dispositions. *Synthese*, 144(3):343–356.

Hellman, G. (1989). *Mathematics Without Numbers: Towards a Modal-Structural Interpretation*. Oxford, England: Oxford University Press.

Hempel, C. G. (1958). The theoretician's dilemma: A study in the logic of theory construction. *Minnesota Studies in the Philosophy of Science*, 2:173–226.

Heylen, J. and Tump, L. A. (2020). The Enhanced Indispensability Argument, the circularity problem, and the interpretability strategy. *Synthese*.

Hilbert, D. (1930). *Grundlagen der Geometrie*. Teubner.

Hoek, D. (2018). Conversational exculpature. *Philosophical Review*, 127(2):151–196.

Hoek, D. (2019). Loose talk, scale presuppositions and qud. In Schlöder, J. J., McHugh, D., and Roelofsen, F., editors, *Proceedings of the 22nd Amsterdam Colloquium*, pages 171–180.

Hofweber, T. (2016). *Ontology and the Ambitions of Metaphysics*. Oxford University Press.

Hofweber, T. (2018). Replies to eklund and uzquiano. *Analysis*, 78(2):315–334.

Hooke, R. (1665). *Micrographia, or, Some physiological descriptions of minute bodies made by magnifying glasses with observations and inquiries thereupon*. Jo. Marten, and Ja. Allestry.

Hooker, C. A. (1968a). Craigian transcriptionism. *American Philosophical Quarterly*, 5(3):152–163.

Hooker, C. A. (1968b). Five arguments against craigian transcriptionism. *Australasian Journal of Philosophy*, 46(3):265–276.

Ioan Bangu, S. (2008). Inference to the Best Explanation and Mathematical Realism. *Synthese*, 160(1):13–20.

Johnston, M. (2004). The obscure object of hallucination. *Philosophical Studies*, 120(1-3):113–83.

Keller, J. A. (2015). Paraphrase, Semantics, and Ontology. *Oxford Studies in Metaphysics*, 9.

Keller, J. A. (2017). Paraphrase and the Symmetry Objection. *Australasian Journal of Philosophy*, 95(2):365–378.

Kennedy, C. and Stanley, J. (2009). On 'average'. *Mind*, 118(471):583–646.

Knowles, R. and Liggins, D. (2015). Good weasel hunting. *Synthese*, 192(10):3397–3412.

Knowles, R. and Saatsi, J. (2019). Mathematics and Explanatory Generality: Nothing but Cognitive Salience. *Forthcoming in Erkenntnis*.

Kong, L. B., Liu, L., Yang, Z., Li, S., Zhang, T., and Wang, C. (2018). *Theory of ferrimagnetism and ferrimagnetic metal oxides*. Elsevier, Amsterdam, Netherlands.

Korman, D. Z. (2009). Eliminativism and the challenge from folk belief. *Noûs*, 43(2):242–264.

Korman, D. Z. (2011). Ordinary objects. In Zalta, E. N., editor, *Stanford Encyclopedia of Philosophy*.

Korman, D. Z. (2014). Debunking perceptual beliefs about ordinary objects. *Philosophers' Imprint*, 14.

Korman, D. Z. (2015). *Objects: Nothing Out of the Ordinary*. Oxford University Press UK.

Korman, D. Z. (2019a). Debunking arguments. *Philosophy Compass*, 14(12):e12638.

Korman, D. Z. (2019b). Debunking arguments in metaethics and metaphysics. In Goldman, A. and McLaughlin, B., editors, *Metaphysics and Cognitive Science*, pages 337–363. Oxford University Press.

Kovacs, D. M. (2019). How to be an uncompromising revisionary ontologist. *Synthese*, 198(3):2129–2152.

Kreisel, G. (1972). Hilary putnam. mathematics without foundations. the journal of philosophy, vol. 64 , pp. 5?22. *Journal of Symbolic Logic*, 37(2):402–404.

Kripke, S. (1981). *Naming and Necessity*. Blackwell.

Lange, M. (2016). *Because Without Cause: Non-Causal Explanations in Science and Mathematics*. Oxford University Press.

Laudan, L. (1990). Demystifying underdetermination. In Savage, C. W., editor, *Scientific Theories*, pages 267–297. University of Minnesota Press.

LeBrun, A. (2021). What are empirical consequences? on dispensability and composite objects. *Synthese*, 199(5-6):13201–13223.

Leng, M. (2005). Revolutionary fictionalism: A call to arms. *Philosophia Mathematica*, 13(3):277–293.

Leng, M. (2012). Taking it easy: A response to colyvan.

Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8(1):339–359.

Lewis, D. (1980). Mad pain and martian pain. In Block, N., editor, *Readings in the Philosophy of Psychology*, pages 216–222. Harvard University Press.

Lewis, D. (1988). Statements partly about observation. *Philosophical Papers*, 17(1):1–31.

Liggins, D. (2008). Quine, {P}utnam, and the '{Q}uine-{P}utnam' indispensability argument. *Erkenntnis*, 68(1):113–127.

Liggins, D. (2012). Weaseling and the content of science. *Mind*, 121(484):997–1005.

Long, J. (2019). One's an illusion: Organisms, reference, and non-eliminative nihilism. *Philosophia*, 47(2):459–475.

Lowe, E. J. (2003). In defense of moderate-sized specimens of dry goods. *Philosophy and Phenomenological Research*, 67(3):704–710.

Lowe, E. J. (2005). How Are Ordinary Objects Possible? *Monist*, 88(4):510–533.

Lowe, E. J. (2010). On the individuation of powers. In Marmodoro, A., editor, *The Metaphysics of Powers: Their Grounding and Their Manifestations*. Routledge.

Maddy, P. (1992). Indispensability and Practice. *The Journal of Philosophy*, 89(6):275–289.

Maddy, P. (1997). *Naturalism in Mathematics*. Oxford, England: Oxford University Press.

Maxwell, G. (1962). The ontological status of theoretical entities. In Feigl, H. and Maxwell, G., editors, *Scientific Explanation, Space, and Time: Minnesota Studies in the Philosophy of Science*, pages 181–192. University of Minnesota Press.

McDaniel, K. (2007). Extended simples. *Philosophical Studies*, 133(1):131–141.

Melia, J. (1995). On What There's Not. *Analysis*, 55(4):223.

Melia, J. (2000). Weaseling away the indispensability argument. *Mind*, 109(435):455–480.

Melia, J. (2002). Response to Colyvan. *Mind*, 111(441):75–80.

Merricks, T. (2001). *Objects and Persons*. Number 3. Oxford University Press, New York.

Merricks, T. (2003). Replies. *Philosophy and Phenomenological Research*, 67(3):727–744.

Merricks, T. (2022). Do ordinary objects exist? no. In Barnes, E., editor, *Current Controversies in Metaphysics*. Routledge.

Mumford, S. (1998). *Dispositions*. Oxford, GB: Clarendon Press.

Mumford, S. and Anjum, R. L. (2010). A powerful theory of causation. In Marmodoro, A., editor, *The Metaphysics of Powers: Their Grounding and Their Manifestations*, pages 143–159. Routledge.

Mumford, S. and Anjum, R. L. (2011a). *Getting Causes From Powers*. Oxford University Press.

Mumford, S. and Anjum, R. L. (2011b). Spoils to the vector - how to model causes if you are a realist about powers. *The Monist*, 94(1):54–80.

Nagel, E. (1961). *The Structure of Science*. Routledge and Kegan Paul.

Nagel, E. (1965). A rejoinder to putnam. *Journal of Philosophy*, 62(16):429–432.

North, J. (2009). The 'structure' of physics: A case study. *The Journal of Philosophy*, 106:57–88.

North, J. (2021a). Formulations of classical mechanics. In Knox, Eleanor; Wilson, A., editor, *Forthcoming in A Companion to the Philosophy of Physics*. Routledge.

North, J. (2021b). *Physics, Structure, and Reality*. Oxford University Press.

Norton, J. D. (2003). Must evidence underdetermine theory. *The Challenge of the Social and the Pressure of Practice*, pages 17–44.

O'Leary-Hawthorne, J. and Michael, M. (1996). Compatibilist semantics in metaphysics: A case study. *Australasian Journal of Philosophy*, 74(1):117–134.

Osborne, R. C. (2016). Debunking rationalist defenses of common-sense ontology: An empirical approach. *Review of Philosophy and Psychology*, 7(1):197–221.

Panza, M. and Sereni, A. (2016). The varieties of indispensability arguments. *Synthese*, 193(2):469–516.

Parsons, J. (2013). Conceptual conservatism and contingent composition. *Inquiry: An Interdisciplinary Journal of Philosophy*, 56(4):327–339.

Polger, T. W. and Shapiro, L. A. (2016). *The Multiple Realization Book*. Oxford: Oxford University Press UK.

Putnam, H. (1965). Craig's Theorem. *Journal of Philosophy*, 62(10):251–260.

Putnam, H. (1967a). Mathematics without Foundations. *The Journal of Philosophy*, 64(1):5–22.

Putnam, H. (1967b). Psychological predicates. In Capitan, W. H. and Merrill, D. D., editors, *Art, Mind, and Religion*, pages 37–48. University of Pittsburgh Press.

Putnam, H. (1971). *Philosophy of Logic*. Routledge.

Putnam, H. (1975). What is mathematical truth? In *Mathematics, Matter, and Method*. Cambridge University Press.

Putnam, H. (2004). *Ethics Without Ontology*. Harvard University Press.

Putnam, H. (2012). Indispensability arguments in the philosophy of mathematics. In *Philosophy in an age of science: Physics, mathematics, and skepticism*, pages 181–201. Harvard University Press.

Quine, W. V. O. (1948). On What There Is. *Review of Metaphysics*, 2(5):21–38.

Quine, W. V. O. (1956). Unification of universes in set theory. *The Journal of Symbolic Logic*, 21(03):267–279.

Quine, W. v. O. (1960). Carnap and logical truth. *Synthese*, 12(4):350–74.

Quine, W. V. O. (1981). *Theories and Things*. Number 1. Harvard University Press.

Raley, Y. (2012). Why the weasel fails. *Philosophia Mathematica*, 20(3):339–345.

Rayo, A. (2015). Nominalism, trivialism, logicism. *Philosophia Mathematica*, 23(1):nku013.

Rizza, D. (2011). Magicicada, Mathematical Explanation and Mathematical Realism. *Erkenntnis*, 74(1):101–114.

Roberts, C. (2012). Information structure: Afterword. *Semantics and Pragmatics*, 5(7):1–19.

Rosen, G. and Dorr, C. (2002). Composition as a Fiction. In Gale, R., editor, *The Blackwell Companion to Metaphysics*, pages 151–174. Blackwell.

Saatsi, J. (2011). The enhanced indispensability argument: Representational versus explanatory role of mathematics in science. *British Journal for the Philosophy of Science*, 62(1):143–154.

Saatsi, J. (2016). On the 'Indispensable Explanatory Role' of mathematics. *Mind*, 125(500):1045–1070.

Saatsi, J. (2017). Structuralism with and without causation. *Synthese*, 194(7):2255–2271.

Saatsi, J. (2020). Explanation and explanationism in science and metaphysics *. In Slater, M. and Yudell, Z., editors, *Metaphysics and the Philosophy of Science: New Essays*. Oxford University Press.

Scheffler, I. (1957). Prospect of a modest empiricism; ii. *Review of Metaphysics*, 10(4):602–625.

Shapiro, L. A. (2000). Multiple realizations. *Journal of Philosophy*, 97(12):635–654.

Sider, T. (1993). Van Inwagen and the Possibility of Gunk. *Analysis*, 53(4):285–289.

Sider, T. (2007). Parthood. *Philosophical Review*, 116(1):51–91.

Sider, T. (2013). Against Parthood. In *Oxford Studies in Metaphysics*, volume 8, pages 237–293.

Siegel, S. (2010). *The Contents of Visual Experience*. Oxford University Press.

Skinner, B. J., Erd, R. C., and Grimaldi, F. S. (1964). Greigite, the thio-spinel of iron; a new mineral. *American Mineralogist*, 49(5-6):543–555.

Sober, E. (1993). Mathematics and indispensability. *Philosophical Review*, 102(1):35–57.

Sober, E. (1999). The multiple realizability argument against reductionism. *Philosophy of Science*, 66(4):542–564.

Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy*, 25(5-6):701–721.

Stalnaker, R. (2003). What might nonconceptual content be? In Gunther, Y. H., editor, *Essays on Nonconceptual Content*, pages 95–106. MIT Press.

Stanford, K. (2009). Underdetermination of Scientific Theory. In Zalta, E., editor, *Stanford Encyclopedia of Philosophy*.

Strawson, P. F. (1950). On referring. *Mind*, 59(235):320–344.

Tahko, T. E. (2021). *Unity of Science*. Cambridge: Cambridge University Press.

Tallant, J. (2013). Optimus prime: paraphrasing prime number talk. *Synthese*, 190(12):2065–2083.

Tarski, A. (1956). A General Theorem Concerning Primitive Notions of {E}uclidean Geometry. *Indagationes mathematicae*, 18.

Tarski, A. (1959). What is Elementary Geometry? In *The Axiomatic Method With Special Reference to Geometry and Physics*. North-Holland.

Thomasson, A. (2007). *Ordinary Objects*. Oxford University Press.

Thomasson, A. (2014). *Ontology Made Easy*. Oup Usa.

van Inwagen, P. (1990). *Material Beings*. Ithaca, N.Y.: Cornell University Press.

VanRullen, R. and Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Sciences*, 7(5):207–213.

Weatherall, J. O. (2019a). Classical Spacetimes. In E.~Knox and A.~Wilson, editors, *The Routledge Companion to Philosophy of Physics*. Routledge.

Weatherall, J. O. (2019b). Why Not Categorical Equivalence? In J.~Madarász and G.~Székely, editors, *Forthcoming in Hajnal Andreka and Istvan Nemeti on Unity of Science*.

Weaver, C. G. (2019). *Fundamental Causation: Physics, Metaphysics, and the Deep Structure of the World*. Routledge.

Williamson, T. (2007). *The Philosophy of Philosophy*. Wiley-Blackwell.

Winnie, J. (1986). Invariants and Objectivity: A Theory with Applications To Relativity and Geometry. In Colodny, R. G., editor, *From Quarks to Quasars*, pages 71–180. Pittsburgh: Pittsburgh University Press.

Wittgenstein, L. J. J. (1953). *Philosophical Investigations*. New York, NY, USA: Wiley-Blackwell.

Woodward, J. (2015). A Functional Account of Causation; or, A Defense of the Legitimacy of Causal Thinking by Reference to the Only Standard That Matters—Usefulness (as Opposed to Metaphysics or Agreement with Intuitive Judgment). *https://doi.org/10.1086/678313*, 81(5):691–713.

Worrall, J. (2011). Underdetermination, realism and empirical equivalence. *Synthese*, 180(2):157–172.

Yablo, S. (1998). Does Ontology Rest on a Mistake? *Proceedings of the Aristotelian Society, Supplementary Volumes*, 72:229–261.

Yablo, S. (2005). The myth of seven. In Kalderon, M. E., editor, *Fictionalism in Metaphysics*, pages 88–115. Clarendon Press.

Yablo, S. (2012). Explanation, extrapolation, and existence. *Mind*, 121(484):1007–1029.

Yablo, S. (2014). *Aboutness*. Princeton University Press.

Yablo, S. (2018). Reply to fine on aboutness. *Philosophical Studies*, 175(6):1495–1512.