UNIVERSITY OF CALIFORNIA

Santa Barbara

Ordinary Object Beliefs and Scientific Theory

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

Doctor of Philosophy

in Philosophy

by

Jeffrey Neal Bagwell

Committee in charge:

Professor Daniel Z. Korman

Professor Kevin Falvey

Professor Thomas Barrett

March 2022

The dissertation of Jeffrey Neal Bagwell is approved.

_____

Kevin Falvey

_____

Thomas Barrett

_____

Daniel Z. Korman, Committee Chair

March 2022

Ordinary Object Beliefs and Scientific Theory

ACKNOWLEDGEMENTS

This dissertation developed out of a seminar on debunking arguments in the fall of 2017, run by Dan Korman, which produced the great grandparent of Chapter 1 as well as initial ideas for the subsequent chapters.

I would first like to thank my committee members, who have enriched this dissertation in countless ways. Thanks to Dan Korman for giving very generous and helpful comments on several drafts of each of these chapters, as well as crucial guidance at every stage of the dissertation. I could not have asked for a more patient, supportive, organized, or enthusiastic dissertation advisor. I also owe deep and abiding thanks to Kevin Falvey and Thomas Barrett, who as committee members gave so generously and so patiently of their time and expertise. I am especially grateful for Kevin's help with Chapter 2 on externalism and perceptual representation, and for Thomas's help with Chapter 3 on indispensability arguments.

For help with Chapter 1 specifically, I would like to thank the members of a conference audience at Johns Hopkins in the fall of 2019. In addition, two anonymous referees at *Ergo* gave very generous and constructive comments on a late-stage draft of that manuscript.

I would like to thank UCSB's Philosophy Department for admitting me as a graduate student, for funding and support over the years, and for approving my filing leave for this quarter. Thanks to C. Anthony Anderson and Aaron Zimmerman for making my initial visit to campus so warm and welcoming. Thanks to my department chairs (and co-chairs) over my time here for their competent leadership: Tom Holden, Matthew Hanser, and Voula Tsouna, I can only imagine how difficult your jobs are and were. Thanks to Teresa Robertson for being

the first non-committee member to ask me substantive questions about my qualifying paper. Thanks to Voula Tsouna and Richard McKirahan for welcoming me and my friends and fellow graduate students into your home and for making the department, and Santa Barbara, feel more like home. Finally, thanks to René Marchington, Staff Graduate Advisor, for shepherding me through my deadlines and paperwork with efficiency and good cheer.

Many thanks to my fellow graduate students for their keen eyes and helpful comments on my drafts, for listening and commenting at graduate colloquia at which I presented ideas, for sitting with me in those write-on-site meetings that were so helpful during quarantine, or for simply letting me bounce ideas off of them when I most needed it. Thanks especially to Christopher Britton, Jason Hanschmann, Kyle Dickey, Colton Heiberg, David King, Blake Kyler, Alex LeBrun, Daniel Story, Sam Zahn, Katherina Gontaryuk, Celine Geday, Sangsu Kim, Rick Lamb, Arnel Blake Batoon, Tom Costigan, Seyed Mohammad Yarandi, and Seán Pierce.

Thanks to Robert Stufflebeam, Edward Johnson, and Clarence Mark Phillips at the University of New Orleans for stoking my interest in philosophy, modeling good teaching, encouraging me to pursue graduate school, and providing letters of recommendation.

I want to give thanks to my friends and family their crucial help and support. Thanks to my friend Paul Taylor for insisting I give graduate school a try, despite my doubts, and for his continuing support. And thanks to my mother Jan Bagwell and my sister Jennifer Cilley, for giving me so much support and encouragement from afar.

Finally, I owe deep and abiding thanks to my wife, Teresa Cahalan, whose love, support, patience, encouragement, and advice throughout this long, difficult process were indispensable, and who has given up so much to help me see this through. She reminded me

of my hidden reserves of strength when I needed it most, and never let me lose sight of the larger task. Most importantly, she always believed in finches—and me.

VITA OF JEFFREY NEAL BAGWELL

March 2022

3421 South Hall

University of California Santa Barbara

Santa Barbara, CA 93106-3090

jbagwell@ucsb.edu

Education

University of California Santa Barbara, Ph.D. degree candidate, Philosophy
expected 2022

Dissertation: "Ordinary Object Beliefs and Scientific Theory"

Committee: Dan Korman (Chair), Thomas Barrett, Kevin Falvey

University of New Orleans, B.A., Philosophy, with Honors, *cum laude* 2014

Areas of Specialization: Metaphysics, Epistemology

Areas of Competence: Philosophy of Science, Ethics, Philosophy of Language, Ancient Philosophy

Publications

"Eliminativism and Evolutionary Debunking" *Ergo* 2021

Works in Progress

"Interfaces and Objects: A Critique of the Interface Theory of Perception"

"Externalism, Object Debunking, and Reliable Misrepresentation"

Dissertation

Title: "Ordinary Object Beliefs and Scientific Theory"

Abstract: My dissertation argues in favor of abundant ontologies that include ordinary objects—the things we seem to see, feel, and touch around us—and other composite objects like molecules and cells. To this end, I advance several arguments against ordinary object eliminativism. First, I argue that evolutionary debunking arguments against ordinary objects are self-defeating and cannot be rescued by employing a popular eliminativist paraphrase strategy. I then develop and evaluate several externalist strategies to resist object debunking arguments, and I argue that we should believe in ordinary objects because they are indispensable to our best scientific theories. Finally, I argue that attempts to rescue debunking arguments such as those proceeding from the Interface Theory of Perception are doomed to fail because their appeals to universal Darwinism are insufficient to justify their empirical premises.

Teaching Experience

University of California Santa Barbara

Instructor of Record, 2019-Present

Epistemology (60 students)

Introduction to Ethics (180 students, 270 students)

Introduction to Philosophy (180 students)

Teaching Assistant, 2016-Present

Critical Thinking

Epistemology

Ethics

Metaphysics

Introduction to Ancient Philosophy

Introduction to Ethics

Introduction to Philosophy

Introduction to Philosophy of Science

Certifications

Pillars of Teaching Assistantship Certificate (UCSB) 2020

Honors, Awards, and Fellowships

2015-2016 Regent's Fellowship, University of California, Santa Barbara

Presentations

Comments on Drew Johnson's "Evolutionary Debunking Arguments and Proper Function"

- Meeting of the Pacific APA (upcoming) April 2022

Comments on Keshav Singh's "There is No Pragmatic Encroachment"

- Meeting of the Pacific APA April 2021

Comments on Uphaar Dooling's "Nativism and Bayesian Models"

- Mountain-Plains Philosophy Conference October 2020

"Eliminativism and Evolutionary Debunking Arguments Against Ordinary Objects"

- Johns Hopkins Philosophy Conference October 2019

Professional Service and Leadership:

Minorities and Philosophy & Pedagogy Reading group, co-leader	2019-2020
Minorities and Philosophy member	2018-present
Graduate Student Representative, University of California Santa Barbara	2017-2019

Courses Taken and Audited
(* indicates audited course)

Metaphysics and Epistemology: Propositions (Anderson), Identity (Salmon)*, Ordinary Objects (Korman)*, Persons (Korman)*, Debunking Arguments and Reliability Challenges (Korman), Easy Epistemology (Korman)*, Action Theory (Falvey)

Philosophy of Science: Scientific Realism and Anti-Realism (Barrett)*, Field's *Science Without Numbers* (Barrett)*

Philosophy of Mind: Knowing our Own Minds (Falvey), Philosophical Psychology (Falvey)

Language and Logic: Symbolic Logic (Salmon), Intermediate Modern Logic (Robertson), Advanced Logic: Gödel (Anderson), Meaning and Reference (Salmon)*

Ethics/Value Theory: Moral Psychology (Zimmerman)*, Ethics of Harm & Justification (Hanser), Aesthetic Theory (McMahon), Feminist Philosophy (Mason)*

History: Hellenistic Philosophy (Tsouna), Aristotle (McKirahan)*, Plato's *Theaetetus* (Tsouna), Plato's *Sophist* (Tsouna)*, Frege (Salmon), Kant's *Critique of Pure Reason* (Elizondo)

Pedagogy: College and University Teaching: From Theory to Practice (Berry and Kuntzmann)

Professional Memberships

American Philosophical Association	2013-Present

References

Dr. Thomas Barrett
5708 South Hall
University of California Santa Barbara
Santa Barbara, CA 93106-3090
tbarrett@philosophy.ucsb.edu

Dr. Kevin Falvey
5709 South Hall
University of California Santa Barbara
Santa Barbara, CA 93106-3090
falvey@philosophy.ucsb.edu

Dr. Daniel Z. Korman
5716 South Hall
University of California Santa Barbara
Santa Barbara, CA 93106-3090
dkorman@philosophy.ucsb.edu

Dr. Voula Tsouna
5722 South Hall
University of California Santa Barbara
Santa Barbara, CA 93106-3090
vtsouna@philosophy.ucsb.edu

ABSTRACT

Ordinary Object Beliefs and Scientific Theory

by

Jeffrey Neal Bagwell

I defend abundant ontologies that include ordinary midsize composite objects—the things we seem to see, feel, and touch around us—and other composite objects like molecules and cells. To this end, I argue that in general, appeals to our best scientific theories support and do not undermine our beliefs in these objects. My main targets are those who appeal to the results of our best scientific theories to foster object skepticism, or who argue that object-free scientific theories are better than the original theories. I mount my defense in two ways. First, I defend our object beliefs against skeptical arguments rooted in appeals to evolutionary biology and the evolution of our own perceptual systems. Second, I advance an original argument that we should believe in composite objects because they are indispensable to our best scientific theories.

In Chapter 1 I argue that eliminativists running evolutionary debunking arguments face a self-defeat problem: their conclusion undermines the scientific support for one of their premises, because evolutionary biology depends on our object beliefs. Using work on reductionism and multiple realizability from the philosophy of science, I argue that it will not suffice for an eliminativist debunker to simply appeal to a paraphrased version of evolutionary theory that does not mention or predict composite objects. In fact, the debunker

must pay a high price in terms of parsimony to recoup the generality of the original, object-laden theory.

An object debunker's skeptical conclusion rests on the claim that our object beliefs are not best explained by the object facts, but rather by our evolved predispositions to perceptually represent the world as containing composite objects even if they don't exist. In Chapter 2, I show that a hybrid externalist view of perceptual representation can provide a composite-friendly explanation of our object beliefs that meets the object debunker's challenge. Such a view also avoids certain objections sometimes raised against externalist views, such as the possibility of illicit *a priori* reasoning about the external world or the inability to accommodate the possibility of reliable misrepresentations.

In Chapter 3, I argue that we should believe in some composite objects because they are indispensable to our best sciences. This argument is based on arguments put forth in the philosophy of mathematics to support of beliefs in mathematical objects. I compare conventional theories like evolutionary biology to their object-free rivals in terms of the virtues involved in theory choice, and I conclude that because there are no scientific reasons for preferring such a composite-free theory, composites are indispensable for our best scientific theories, and we would need other, non-scientific reasons for rejecting them.

In Chapter 4, I raise objections to a debunking argument put forth by the cognitive scientist Donald Hoffman and others, based on Hoffman's Interface Theory of Perception. Appealing to experimental results in evolutionary game theory, Hoffman argues that our perceptual faculties evolved to guide fitness-enhancing behavior without giving us veridical perceptions of ordinary objects, spacetime, or causal interactions. I show that Hoffman's argument is self-defeating in a similar manner to what I described in Chapter 1, and his

attempts to get around this problem by appealing to a substrate-neutral Universal Darwinism that does not assume the existence of physical objects leads to a fatal dilemma.

TABLE OF CONTENTS

Chapter 1:
Eliminativism and Evolutionary Debunking[1]

1. Introduction

Eliminativists sometimes invoke evolutionary debunking arguments against ordinary object beliefs, either to help them establish skepticism about such objects or to break down one's resistance to abandoning common sense ontology.[2] My purpose in this chapter is to show that the eliminativist debunker faces a self-defeat problem. Her premises appeal to the theory of evolution by natural selection and her conclusion is skepticism about ordinary objects. However, evolutionary theory is *about* ordinary objects; it systematically appeals to our object beliefs. I argue that simply converting each scientific proposition about some ordinary object *K* into a proposition about *simples arranged K-wise* does not circumvent the problem.

My reasons are as follows. Attempts to recast the propositions of evolutionary theory in terms of simples arranged *K*-wise commit the eliminativist to a problematic form of reductionism about scientific theories. The eliminativist's low-level surrogate theory of evolution is ultimately unable to explain how the human perceptual system evolved because it will lack the needed generality and explanatory power. This undermines the justification for one of the debunker's premises.

Here is a bird's-eye view of the chapter. In Section 2, I state and explain an evolutionary debunking argument against ordinary objects. In Section 3, I sketch the self-

[1] This chapter was originally published in December of 2021 as "Eliminativism and Evolutionary Debunking" in *Ergo*, vol. 8, no. 17. doi: https://doi.org/10.3998/ergo.1154

[2] I have in mind Merricks (2001: 72–76) and Benovsky (2015: §2). But see Korman (2019b: 340 n7) for a more complete list of those invoking debunking arguments to support various kinds of departures from common sense ontologies.

defeat problem for object debunking arguments by exploring the object dependency of evolutionary theory. In Section 4, I sketch two popular variants of a *K*-wise conversion strategy and evaluate their prospects for running the debunking argument. In Sections 5 and 6, I show that using the *K*-wise conversion strategy commits the debunker to a form of scientific reductionism, and that the resulting surrogate, low-level theory will lack generality. In Section 7, I argue that because it lacks generality, the eliminativist's surrogate theory will be limited in its ability to predict or explain relevant phenomena and to utilize existing evidence. In Section 8, I show why attempts to recoup generality in terms of *pluralities* of simples arranged *K*-wise fail.

2. Evolutionary Debunking Arguments and Ordinary Objects

Debunking arguments target certain kinds of beliefs in order to establish some limited form of skepticism. *Evolutionary* debunking arguments rely on the fact that our evolutionary history *predisposes* us to form certain kinds of beliefs—not because these beliefs are true, but simply because they increased our ancestors' reproductive fitness. Learning that you are just hard-wired to believe that *p* under the right conditions, *regardless of whether* p *is true or not*, serves as a defeater for your normal justifications as to why you believe that *p*.

You probably believe that there are visible, medium-sized solid objects all around you because it *seems* as if there are. This *seeming* might be a sufficient normal justification for believing that ordinary objects like trees exist as you go about your day. But once the object debunker convinces you that your reason for believing in trees has nothing to do with

whether or not there are trees and everything to do with what was adaptive for your ancestors to believe, this defeats such a normal justification. Your tree beliefs are thereby debunked.[3]

Here is an evolutionary debunking argument against ordinary objects (EDO):

(EDO1) The best explanation of your ordinary object beliefs is that you only believe there are ordinary objects because you are hard-wired by evolution to believe in them in the presence of matter arranged object-wise—irrespective of whether it's *true* or not that there are ordinary objects.

(EDO2) If EDO1 is true, then you are not justified in retaining your object beliefs.[4]

(EDO3) So, you are not justified in retaining your object beliefs.[5]

EDO1 relies on one plausible interpretation of the evolutionary psychology of human perception. Modern humans believe in the existence of ordinary objects like trees based on their having sensory experiences as of trees existing. These experiences are the result of an evolved perceptual system. According to the debunker, our ancestors' perceptual systems evolved to track *adaptively relevant matter* (e.g., matter arranged food-wise, mate-wise, or predator-wise) well enough to out-compete reproductive rivals; at the same time, they may very well have evolved to have *false* beliefs about ordinary objects. Our predisposition to

[3] In metaethics, evolutionary debunkers sometimes make the plausible assumption that while moral realism is vulnerable to this kind of skeptical argument, realism about ordinary objects is safe because one can provide an evolutionary vindication for our believing that ordinary objects are real. For instance, Sharon Street (2006:160–61n) notes that facts about salient objects in the environment such as predators, obstacles, or other hazards, could plausibly factor into our best explanations of why we form beliefs about them. Having a capacity to track these object facts would have bestowed a clear adaptive benefit on our ancestors: creatures believing that predators exist and are dangerous would tend to avoid predators and survive to reproduce. Thus, evolution seems to vindicate our object beliefs. For a more detailed counterargument to the supposed evolutionary vindication of object beliefs, see Korman (2019b: 342–45).

[4] We can assume, if we like, that one must be aware of this defeater to lose any justification one already has for object beliefs. This will not affect my discussion, as my focus is on whether the argument is self-defeating or not.

[5] This argument is loosely adapted from Korman (2019b: 340).

believe in ordinary objects *need not* be the result of such objects existing in the ancestral environment; rather, they simply need to have conferred a reproductive advantage over rivals who inherited different perceptual predispositions (or to have introduced no substantial reproductive *disadvantage*).

The final clause, "irrespective of whether it's *true* or not that there are ordinary objects," bears some unpacking. The basic idea here is that introducing ordinary object facts *adds nothing* to the above causal explanation; rather, it makes our explanation *less* parsimonious, clear, and illuminating.[6] The parsimony concern may be simply about injecting additional objects and object facts into our ontology when we already have a complete causal explanation on hand. This usually involves the idea that ordinary objects (or facts about composition) are causally inert in themselves or are mere causal overdeterminers. Positing such an overdetermining cause may itself be objectionably unparsimonious, or it may conflict with the notion that to exist is to have causal powers.[7]

In EDO2 we are assuming that whatever our reasons for believing in ordinary objects in the first place, they only merit *continued* ontological commitment if they are essential to our best explanations of why we believe in them. But if EDO1 is true, our best explanations of why we believe in ordinary objects *don't* make any essential reference to ordinary objects. This is true even if ordinary objects happen to exist. Accordingly, EDO2 captures the fact that EDO1 is a defeater for our normal reasons for believing in ordinary objects.

---

[6] For a detailed exposition of this take on the debunking argument, developed as a Sharon-Street style Darwinian dilemma for the object realist, see Korman (2019b: 342–45). For a very different, earlier take on the causal worry in object debunking arguments, in which the causal connection between a tree and our tree belief is at best a deviant one, see Korman (2014: §5).

[7] On this latter point, see Merricks (2001: 65). For an account of how Merricks' causal overdetermination argument works as a defeater, see Merricks (2003: 738–43). For a more recent version of the overdetermination argument, see Merricks (2017). For an overview of the overdetermination argument and some replies, see Korman (2015: Ch.10)

3. The *Prima Facie* Self-Defeat Problem for EDO

Scientific theories like evolutionary biology systematically appeal to our perceptual beliefs about ordinary objects. If we reject these object beliefs, we jeopardize not only our theories' explanations and laws, but also our empirical evidence and our ability to *rank* theories based on such evidence.[8, 9] In this section, I will elaborate on each of these points, and show why they collectively spell self-defeat for the object debunker.[10]

I call the self-defeat problem *prima facie* because it will be apparent to anyone from the standpoint of commonsense ontology. However, seasoned eliminativists may already be eager to dispute the claim, armed with strategies to reinterpret or recast scientific propositions to rid them of ordinary object commitments. They and other impatient readers are free to skip ahead to Section 4.

Let us consider an example of the kind of thing evolutionary theory was developed to explain. Why do certain species of Galapagos finches endemic to a particular island have substantially bigger beaks than those of finches on neighboring islands—beaks that allow them to crack open the thick-shelled seed capsules that happen to drop from trees that flourish on their island in particular? The answer to this will inform broader theoretical questions such as: How do species come to have qualities that make them seem well-suited to their environments? How does speciation occur? What even *is* a species?

[8] For instance, observations made through a microscope all depend on some theory of how the microscope and its parts—all ordinary objects—work, and why we should trust them.
[9] Williamson (2007: 223–24) mentions several of these worries in considering the promises of reductionism and their consequences for science.
[10] Eliminativists who are instrumentalists about science may be ready to bite the bullet and accept any epistemic consequences of object skepticism. However, if they wish to convince an audience by using the debunking argument, they too need to resolve the self-defeat problem.

Note that when we formulate questions about the concrete *explananda* of evolutionary theory, we must appeal to perceptual beliefs about finches, beaks, islands, and various ordinary objects in a finch's environment such as seeds, shells, and trees. Likewise, our broader theoretical questions about how a species relates to its environment over time appeal to beliefs about *patterns* involving ordinary objects: that there are living organisms of various kinds, that organisms bear properties, that some of these properties are adaptive with respect to an environment, and that an organism's environment is made up of all kinds of ordinary objects.[11][12]

The Darwinian *explanantia* that answer these questions similarly depend on ordinary objects. Here is a rough explanation of why the finches on the island evolved bigger beaks. Over time, variation in beak size in the island's finch population gave a reproductive advantage to finches with bigger beaks, because only the finches with bigger beaks were able to eat certain difficult-to-access seeds that are abundant on their island even during times of great scarcity. The trait for bigger beaks was passed on to their offspring, who were more numerous than those of their rivals with smaller beaks. This process repeated over the course of many generations, with the result that all finches on the island now have the trait of bigger beaks.[13] Note that our explanation implicitly appeals to *patterns* exhibited by organisms, such

[11] This presents a problem for eliminativists like van Inwagen (1990), who allow an exception for organisms, but not the ordinary objects that make up their environments. Inanimate objects play important roles as selective pressures on organisms.

[12] Note that even tools and methods that allow us to look beyond ordinary objects (say, into microscopica) depend on object beliefs. How does one know how to use a microscope, or trust its deliverances, if one doesn't believe it exists? Both Merricks (2001:175) and Williamson (2007: 223) raise this point.

[13] This is a greatly simplified account of one set of dynamics drawn from a large and formidably complex ecosystem. Often, this kind of niche specialization is observed between different species of finch on the *same* island during periods of scarcity due to drought and subsequent famine. See Weiner (1994) for an in-depth picture.

as heredity, phenotypic variation, and differential reproduction.[14] This explanation supports an evolutionary law: given that variation exists regarding a specific trait (here, beak size), if one variant gives individuals possessing it a reproductive advantage because it helps its possessors cope more effectively with selective pressures in the environment (here, the scarcity of food), this variant will become more frequent in succeeding generations, eventually replacing rival variants throughout an entire reproductive population.

Let us now turn to the question of evolutionary theory's *justification*. Why is the theory better than its rivals as an explanation of the complexity, diversity, and distribution of life on earth? To do this, we will examine one theoretical virtue natural selection is thought to have in spades: its explanatory power. A theory has greater explanatory power than its rivals when, all things being equal, it leaves fewer aspects of its subject matter a mystery. The following simplified example serves to give a sense of how these comparative explanations depend on data in terms of ordinary objects.

Traditionally, evolution by natural selection has had one main rival: creationism. This is the view that the species we see all around us were individually created for their environments, as opposed to being descended with modification by natural processes from ancestral species over countless generations.[15] We will compare the way each theory handles the following sets of observations: in addition to the finches that developed big beaks, there were finches on different islands with smaller, more delicate beaks that seem well adapted

[14]Of course, not all organisms are ordinary objects (e.g., bacteria and other microscopica). But if one is ruling out ordinary objects, unless one has an exception for some composite objects like DNA strands, one will have nothing upon which to base generalized properties like heredity.

[15] Though in the minds of most biologists evolutionary theory has no serious rival, there are robust disagreements *within* evolutionary theory about, e.g., the specific mechanisms of adaptive change and the role natural selection plays in combination with other factors. These intra-theoretical disputes depend on data in terms of ordinary objects.

for the diet available in their own environments; and both groups of finches bear striking resemblances to each other and to birds on the nearby mainland of South America.

Creationism would maintain that each species of bird was specially created for its particular island environment. This explains why each finch population is particularly well-suited to its island environment but does not explain their similarity to the mainland finches. However, there is no apparent reason the creator should make these island finch species resemble those on the mainland, who are not particularly well-adapted for any of these island micro-environments. Creationism leaves this striking pattern a mystery.[16]

By contrast, evolutionary theory suggests that the finches on the mainland represent an *ancestral* species that migrated to the islands in the distant past and then diverged into sub-species, as finches on each island adapted to the selective pressures of their new environment but were cut off from interbreeding with the finches on the other islands. These considerations seem to favor the evolutionary explanation, because it can explain the larger set of observations—those about the island birds *and* the mainland birds—better than its rival.

However, if we embrace skepticism about ordinary objects, we cannot cite the presence of a common ancestor as something evolutionary theory explains better than its rivals. Both ancestors and descendants here are birds—ordinary objects—populating an environment filled with ordinary objects. An object skeptic seems to lose any reason to consider evolutionary theory to be the best explanation of its subject matter. In fact, she seems to be in no position to accept evolutionary theory at all: its very subject matter—as

[16] In fact, creationism's explanation fits the observations so loosely it would be compatible with wildly different observations: for instance, if the finches on various islands did not resemble each other—or the finches on the mainland—*at all*, or if our big-beaked finches were identical copies of some species on the opposite side of the globe with a similar micro-environment.

well as its laws, explanations, observations, and methods—depend on appeals to perceptual beliefs about ordinary objects; and it must rely on such appeals to display its virtues against *competing* theories. Object skepticism leaves evolutionary theory fundamentally unjustified.[17]

This lack of justification undermines EDO1, since we now have no reason to believe the evolutionary hypothesis that selective pressures shaped the mechanisms in our ancestors' brains responsible for converting perceived qualities into representations of three-dimensional objects.[18] And this spells self-defeat for the object debunker, because it puts the skeptical conclusion EDO3 at odds with the premise EDO1. We cannot rationally accept an argument wherein the conclusion undermines one of the premises.[19]

In the following sections, I will explore a strategy for converting propositions of evolutionary theory about any ordinary object *K* into those about simples arranged *K*-wise. I will examine how it affects the debunker's appeal to evolutionary science and show why it ultimately cannot save the debunking argument. I hope to convince the reader that the self-defeat problem is not just *prima facie*; rather, it is a deep and persistent problem for the debunker.

[17] That is, if the theory remains coherent with its very subject matter removed from discussion.
[18] Premise EDO1 could still succeed on other grounds, of course, assuming those arguments in support of it do not similarly rely on ordinary objects. However, barring arguments that culture is the sole factor responsible for biasing us toward believing in ordinary objects, evolutionary debunking arguments would lose their distinctive force as arguments for EDO1: they provide positive, empirical evidence that our object beliefs are unrelated to object facts. This crucially distinguishes them from more universal kinds of skepticism (Vavova 2015: 105–6). Cultural debunking arguments also arguably presuppose an evolutionary backstory. To be able to process language and other cultural information was an adaptation that bestowed clear reproductive advantages on our ancestors. But the relationship between Darwinian evolution and exclusively cultural predispositions to believe in ordinary objects is at best complex, indirect, and controversial.
[19] My position in this dilemma is that we should reject EDO1, because we can meet the explanatory challenge by invoking the results of perceptual psychology. Explication of this is outside the scope of this chapter.

4. Running the Debunking Argument without Objects

In this section, I will describe two strategies for an eliminativist who wants to run the debunking argument while avoiding the self-defeat problem. Both involve converting the propositions of evolutionary theory into object-free propositions, and both originated as solutions to the problem of explaining why most people can be reasonable, though they hold many false, object-laden perceptual beliefs. The two strategies are *compatibilism* and *incompatibilism.*[20]

In either case, to simplify our discussion let us assume that the debunker is a *nihilist* about composition. Think of this as an extreme kind of eliminativist who rejects composition altogether and believes *all* objects are *mereologically simple* (that is, partless or uncomposed).[21][22] To explain our experiences of a world apparently filled with visible objects, a nihilist holds that simples act together in various ways to cause the *appearance* of ordinary objects and those macroscopic effects we attribute to them.

Compatibilism is the view that there is no real conflict between the beliefs of ordinary non-philosophers (the folk) and those of revisionary ontologists. A compatibilist holds that

[20] In my terminology, I follow O'Leary-Hawthorne and Michael (1996), who use 'compatibilism' to describe van Inwagen's paraphrase strategy (see van Inwagen 1990: Chs.10–11; 2014). Korman (2009) develops and utilizes this distinction as a way of contrasting van Inwagen's strategy from the views of incompatibilists like Merricks (2001: Ch.7). My versions of compatibilism and incompatibilism here are loosely based on the views of van Inwagen and Merricks, and the distinction between speaking inside and outside the ontology room comes from van Inwagen (2014).

[21] Not all nihilists are eliminativists, nor are all eliminativists nihilists. Examples of non-nihilist eliminativists include van Inwagen (1990), who famously makes exceptions for living organisms, and Merricks (2001), who makes exceptions for conscious beings. An example of a non-eliminativist nihilist is Contessa (2014), who defines a kind of nihilism that resists ordinary object eliminativism. In addition, it is possible for a nihilist to hold the odd position that ordinary objects are mereologically simple.

[22]The debunker may try to be neutral about these matters and just point to finch-wise experiences being caused by something in a certain region. But if she accepts the in-principle possibility of giving a complete lower-level causal, scientific account, her options are restricted. The stuff in that region must be either simples, composites, or gunk. And (as will become clear in Section 5) gunk wouldn't support the kind of reductive causal story the eliminativist needs to tell, because in a gunky world causation does not bottom out at some specific level of explanation.

because the folk are speaking *outside* the ontology room, their sentences should be interpreted differently than those uttered *inside* the ontology room. This is because the ontology room is a different *context of utterance* from the outside world—including the world of scientists. Philosophers involved in academic debate who say, "there is a table" would be expressing a false proposition, while ordinary folk in the course of their normal lives who utter the *same* sentence would be expressing a true proposition—provided they were in the presence of some simples arranged tablewise.

Compatibilists regard folk utterances of "there is a table" as ontologically neutral, uncommitted to the existence of ordinary objects. The truth-conditions of such folk utterances are determined by generating and evaluating a *paraphrase* of the original: "there is a table" becomes "there are some simples arranged table-wise." This strategy aims to vindicate the reasonableness of folk discourse by capturing what is correct in everyday speech involving ordinary objects.

Incompatibilism is the view that there really is a conflict between folk beliefs and those of revisionary ontologists. An incompatibilist makes no distinction between what is uttered inside or outside the ontology room, holding that both philosophers and the folk are stating a false proposition when they utter "there is a table." The incompatibilist still must explain how most people can believe false things and still be reasonable—and, crucially, what makes false beliefs about tables *more reasonable* than false beliefs about unicorns.

To solve this problem, the incompatibilist adds an epistemic category here: beliefs about things like tables are *false, but nearly as good as true*, while beliefs about unicorns are merely *false*. We can identify beliefs that are nearly as good as true by employing this kind of rule: "Any folk-ontological claim of the form 'F exists' is nearly as good as true if and only if

(i) 'F exists' is false and (ii) there are things arranged F-wise" (Merricks 2001: 171–74). Beliefs that are nearly as good as true are still false, but they can serve valuable functions such as warranting other (true) beliefs. Moreover, this distinction allows the incompatibilist to hold that scientists and other ordinary folk are reasonable because their beliefs, though false, have some measure of epistemic virtue.

In their solutions to the problem of reasonableness, both compatibilists and incompatibilists make use of a similar strategy: take any (false) proposition about some ordinary object *K* and convert it into a (true) proposition about *simples arranged K*-wise.[23] The compatibilist uses this as a *truth-maker* for statements about ordinary objects made outside the ontology room. "There is a finch" is true if and only if the ontologically neutral paraphrase "there are simples arranged finch-wise" is true. For the incompatibilist, statements like "There are some simples arranged finch-wise" are *nearly-as-good-as-true-makers.* The nearby metaphysical fact that there are some simples arranged finch-wise makes "There is a finch" nearly as good as true; however, it is not assumed to be a *paraphrase* of the speaker's words (however loose), let alone a truth-maker. "There is a finch" is still false—but it's the *good* kind of false.

In order to run the debunking argument, both compatibilists and incompatibilists need to convert the collection of all propositions necessary for evolutionary theory and its justification into *K*-wise terms.[24] Let us call this collection *E*. Included in *E* are all propositions either (i) composing the theory of evolution (propositions of law, method, and supporting explanatory discourse) or (ii) serving as evidence for that theory (propositions of

[23] See, e.g., Merricks (2001: Ch.1) and van Inwagen (1990: Ch.11) for versions of the *K-wise* strategy.
[24] Hereafter I will use 'conversion' instead of 'paraphrase' to describe what the eliminativist is doing, since 'paraphrase' implies the intent is to preserve the meaning of the original statement. This would only apply to the compatibilist strategy.

observation). Recasting the propositions of *E* according to the *K*-wise conversion strategy, we generate a *different* collection of propositions. Let us call this E_{Lite}. An eliminativist doesn't need to be skeptical about the propositions of E_{Lite}, because they are not *about* ordinary objects (or any composite objects). Let's return to our two strategies and see how they fare with E_{Lite} in hand.

Compatibilism faces a dilemma. It holds that *E* is *true* when expounded by scientists, who work outside the ontology room, because when scientists utter sentences that appear to be expressing propositions of *E*, they are really expressing propositions of E_{Lite}. So, the compatibilist has an eliminativist-friendly way of justifying EDO1. However, outside the ontology room *eliminativism* appears to be false because it entails that "There are some finches" is false.[25] But we've already established that outside the ontology room this statement is true. So, EDO3 is false. However, back *inside* the ontology room, *E* is still an unjustified theory made up of false propositions about ordinary objects. So, the compatibilist can't successfully run the debunking argument either inside or outside the ontology room.

My view is that the compatibilist is ultimately forced to abandon this distinction, and her view ultimately collapses into *incompatibilism.* First, the ontology room seems to be the appropriate place to run the debunking argument. It seems we are there right now, and anyone hearing EDO seems to be thereby ushered inside. Second, to run EDO *outside* the ontology room, the compatibilist would need some reason to reinterpret eliminativism out there such that it remains true. This move seems completely unmotivated and *ad hoc*.[26] Third, Trenton Merricks (2014) has given solid reasons why it is implausible that we should

[25] By *modus tollens*, if it's false that "There are some finches" is false, then it's false that eliminativism is true.
[26] The distinction as presented in van Inwagen doesn't offer any clues (1990: Chs. 10–11; 2014). If we try to render either eliminativism or EDO in terms of ontologically neutral paraphrase, the results would seem to be incoherent.

interpret the folk as making ontologically neutral statements when they make claims stating or presupposing ordinary objects in their ordinary lives, and in general why any revisionary ontologist should reject the compatibilist's distinction.[27] Finally, E_{Lite} cannot be a truth-maker for *E* because it is—as we will see especially in Sections 7 and 8—a *different* theory from *E*. It has different laws, *explananda,* and theoretical virtues. Accordingly, its propositions have *different* truth-conditions from those of *E*, making them unsuitable as truth-making paraphrases.

Incompatibilism, it seems, is the only viable strategy. From the outset, the incompatibilist considers E_{Lite} to be a *separate* theory from *E*, not a mere paraphrase. She uses E_{Lite} to support $EDO1_{Lite}$, the first premise of an object-free version of the debunking argument, which we'll call EDO_{Lite}.[28] She accepts that *E* is false, but believes it is *nearly as good as true*. Apart from its widespread appeals to false beliefs in ordinary objects, *E* has a certain trustworthiness that explains why it is worth invoking in EDO. This trustworthiness depends on there being some corresponding proposition of E_{Lite} about simples arranged *K*-wise for every proposition of *E* about some object *K*. Accordingly, she believes that every scientist who believed he observed a finch (and wasn't deceived, e.g., by perceptual illusion) had perceptions caused by simples arranged finch-wise. She also believes that inferences drawn from such false observational beliefs can confer *some* kind of justification or warrant.[29] E_{Lite} is a kind of *conversion* or *recasting* of the false, object-laden propositions of *E* into propositions that express these closely related truths about simples arranged *K*-wise.

[27] Among his reasons are that the ontology room doesn't seem to be a genuine context of utterance, and that the distinction is ultimately hostile to revisionary ontology and indeed to any kind of revisionism.

[28] For present purposes, I will assume there is nothing problematic about providing object-free conversions of EDO1, EDO2, and EDO3 (that is, $EDO1_{Lite,}$ $EDO2_{Lite,}$ and $EDO3_{Lite}$).

[29] Merricks is cagey about how this warranting happens. For instance, it could depend on the relation to the nearly-as-good-as-true belief, or it could depend on the relation to a nearby *truth* about simples (2001: 171–74).

We will now assume that the eliminativist is an *incompatibilist* in the above sense. In the next section, I will explore the eliminativist's scientific commitments in more depth.

5. The Eliminativist *K*-wise Strategy and Scientific Reductionism*

Despite its promise, this *K*-wise conversion strategy leaves us several reasons to be skeptical. The principal problem for an eliminativist surrogate of evolutionary theory is that the propositions of E_{Lite} do not exist. And it's not obvious that we can recast, without epistemic loss, all needed propositions of *E* into propositions that do not express commitments to ordinary objects.[30] For instance, there may be technical problems with the kind of plural reference and quantification needed for a general and systematic *K*-wise conversion strategy.[31] However, in this section I will raise a different problem: even if it turns out to be easy to convert propositions of *E* into propositions about simples arranged *K*-wise, the eliminativist is committed to a kind of scientific reductionism that ultimately limits the capabilities of E_{Lite} as a scientific theory.

Take the proposition *that this finch is brown-beaked*. The debunker could convert this into the proposition that *some of these simples arranged finch-wise are arranged brown-beak-wise*. However, to do so expresses not just a metaphysical commitment but also a physical, scientific one: the real objects of scientific study here—the things doing the causal work—are not finches but microscopic objects arranged *finch-wise* and *brown-beak-wise*.

[30] Williamson (2007: 223) briefly raises this worry.

[31] For instance, Uzquiano (2004) argues that in order to demonstrate the plausibility of the needed kind of quantification—a truly *plurally plural* quantification—one needs to supplement it with additional resources that will ultimately result in costly ontological trade-offs for an eliminativist. Uzquiano speaks of 'paraphrase' because he is criticizing van Inwagen's position (in 1990: Chs. 10–11). Hereafter, I will only speak of *converting* or *recasting* these propositions, as the incompatibilist doesn't claim her *K-wise* propositions are literal interpretations of the source statement.

What happens when we take *K*-wise propositions seriously in *physical* terms? If simples are microscopic, partless, causally efficacious objects, they must be among the smallest things scientists currently study (i.e., quarks, leptons), or else they are some as-yet-unidentified things on an even smaller scale. Whatever they turn out to be, simples would seem to belong to quantum physics.

The eliminativist implies that the story of finches can, at least in principle, be replaced by a story about finch-wise things at the level of quantum physics. Moreover, it *should* be told at the quantum level if we are to abandon talk of ordinary objects. Thus, E_{Lite} should ultimately not be composed of propositions about simples arranged *K*-wise, but of propositions that describe quantum particles and their various properties of motion, mass, charge, position, or the like that make up their being arranged *K-wise*. The eliminativist is committed to *some* kind of reductionism in science—presumably to the in-principle possibility of reducing *E* to quantum physics, with E_{Lite} being the reducing theory.

This is not a typical kind of scientific reductionism, so let us speak of reductionism* (and of reduction*, reducing*, etc.) to describe the eliminativist's commitments. Typically, a reductionist does not regard the reduced, higher-level theory as *false.* But for the eliminativist, *E* is a *false* (but *nearly as good as true*) higher-level theory that is merely a means to the *true* lower-level theory E_{Lite}. Once the reduction* is complete, we should not need or want to appeal to the higher-level theory: we climb down the ladder and kick it away.

Though false, *E* does have some measure of epistemic virtue that motivates the debunker's appealing to it in the first place. Reduction* to E_{Lite} should preserve as much of this virtue as possible. Because *E's* dependence on ordinary object beliefs is systematic, the eliminativist needs for the reduction* to *also* be systematic in nature. This ensures that it is

possible to reduce* every proposition of *E* needed to run the debunking argument to some proposition of $E_{Lite.}$ Note that the reduction* of each needed proposition of *E* about some ordinary object *K* also confirms that proposition is indeed nearly as good as true, because it establishes there *is* some nearby truth about quarks arranged *K*-wise.[32]

The standard view of scientific theory reduction involves the idea that one body of scientific knowledge can be reduced to another—specifically, that some theory T_A reduces another theory T_B if T_A logically entails T_B. This is usually understood to require *bridge principles* that establish logical relations between higher-level kinds in T_B with lower-level kinds in T_A. Of special epistemic importance is that the *laws* of the lower-level theory, combined with bridge principles, entail the *laws* of the higher-level theory. This demonstrates that the knowledge contained in the higher-level theory's generalizations is contained in the lower level, reducing theory.[33]

The eliminativist denies that any proposition S_{Lite} of E_{Lite} about elementary particles, together with bridge principles, *entails* some proposition *S* of *E* about ordinary objects. Claims about ordinary objects—and whatever entails them—are false. Rather, she would need a rule like the following: what S_{Lite} plus bridge principles entails is some proposition *S** that entails that some proposition *S* of *E* is *nearly as good as true*. Given this qualification, the eliminativist can relate propositions of E_{Lite} to those of *E* in a general, systematic manner.[34]

---

[32] Hereafter, for simplicity we will assume that all mereological simples are *quarks*, and that quarks stand in for all elementary particles. This is a convention, like calling mereological simples "atoms." I use it to stress that E_{Lite} is made up of propositions about physical particles—not merely metaphysical posits.

[33] See Nagel (1961: Ch.11) for a classic statement of the view, and Brigandt and Love (2017: §3.1) for a useful overview.

[34] The picture looks something like the following. Proposition S_{Lite} is a description of what elementary particles are doing in a particular situation. Combined with bridge principles, S_{Lite} entails proposition S*: *that there are some simples arranged finch-wise* (or perhaps *that there is a plurality of simples arranged finch-wise* or *that*

Before moving on, I want to address the objection that E_{Lite} is really an *utterly* independent theory from *E,* with an entirely independent justification. The thought runs like this: all this talk of *preserving* justification or other epistemic virtues from *E* is misplaced. After all, the theory is false, and we shouldn't worry about what we're taking away from a theory we ultimately reject anyway.

The reasons to reject this view are simple. If E_{Lite} were an entirely new science, independent from *E* in every way, it would not yet exist—nor would it be justified. Currently its laws and explanations are unwritten, its hypotheses untested. This would not meet the debunker's needs for running EDO_{Lite}.

The debunker's audience is those who believe the results of *E* but not (yet) in eliminativism. This audience would not be justified in accepting the pronouncements of an unknown, untested science as support for $EDO1_{Lite}$. Rather than appealing to the results of an established science, or some principled modification of it, the eliminativist would be appealing to the in-principle possibility of a from-scratch theory of human evolution in terms of quarks, the possibility that it would say the needed things about human perceptual beliefs in ordinary objects to support $EDO1_{Lite}$—and the possibility that we should *believe* what it says.

Ultimately, even a completely new theory would be judged by whether and how it tells the *same* story *E* tells so clearly and with such authority. At present, the only way we can get even a rough sense of how such a theory would compare to *E* is to begin with *E* and imagine what it would take to reduce* it to the lowest level in a principled way.

---

there is a finch-wise arrangement of simples). Proposition *S** entails that proposition *S* of *E*—*that there is a finch*—is nearly as good as true.

As noted above, E_{Lite} is not a reduction in the conventional sense, but a surrogate theory that replaces *E* by capturing as much content from *E's* propositions as possible in an object-free way. While E_{Lite} doesn't need to have a surrogate claim for *every* proposition of *E*, it does need to be able to reproduce *E*'s explanation of human perceptual beliefs, and to be justified enough for us to believe it over its competitors. This justification comes not from running new experiments but from *taking existing propositions of observation, law, and explanation to have been nearly as good as true*.

6. Reductionism* and the Generality of E_{Lite}

The eliminativist's reductionism* is a substantial commitment that is independent from the commitments of evolutionary biology as a special science, and there are initial reasons to think it is a liability. The first and most obvious criticism of reduction* is that a scientific reduction* from evolutionary biology to quantum physics simply has not been done. Without the propositions of E_{Lite}, it is unclear what is to take the place of *E* in the debunker's argument.[35]

Moreover, we have reasons to believe that E_{Lite} will never materialize. For instance, there are problems even partially accomplishing a reduction within biology itself. It is controversial whether classical, Mendelian genetics can be reduced to microbiology in the sense of theory reduction outlined above (see Hull 1972 and Kimbrough 1978). If there is substantial difficulty reducing one *subfield* of biology to another, it's an open question whether in some kind of grand unifying reduction of all the relevant fields of evolutionary

[35] Conventional reductions in the field of biology have been piecemeal, focused more on achieving a causal explanation of some part of the higher-level theory. Such reductions are not assumed to replace or eliminate the higher-level theory. For an overview of this kind of partial, explanatory reduction in contrast to the full-fledged theory reduction to which the eliminativist is committed, see Brigandt and Love (2017: §3.2); for a survey of the kind of methodological assumptions at work in these partial reductions, see Kaiser (2011).

biology to the smallest scale of quantum physics these difficulties might be greatly multiplied.[36]

However, my focus will not be on the lack of availability of the propositions resulting from reduction* to E_{Lite}, but on their undesirability. One major principled criticism that has been leveled against scientific reductionism is that higher-level kinds are often *multiply realizable* at the lower levels. For instance, a single phenotype in classical genetics is often realizable by multiple molecular mechanisms (see Hull 1972: §3). In such cases, a bridge principle relating a phenotype (*P*) of classical genetics to its molecular description in terms of microbiology (*M*) will be disjunctive on the side of the *reducing* theory:

$$\forall x\,(Px \leftrightarrow (M_1x \vee M_2x \vee M_3x \vee \ldots \vee M_nx))$$

This disjunctiveness becomes important when one tries to reduce the *kinds* of the higher-level theory to the *kinds* of the reducing theory. Unless the reducing theory captures the kinds of the higher-level theory in an orderly fashion, it cannot capture the full generality of the *laws* of the reduced theory. However, as Fodor and others have argued, the multiple realizability of higher-level features makes this kind of *type-type reduction* impossible. These "laws" will seem more like gerrymandered collections, not sufficiently general to do the work of real scientific laws.[37]

---

[36] For instance, it's *prima facie* unclear whether it's possible to reduce classical genetics to quantum physics without first *passing through* the level of microbiology and dealing with the aforementioned difficulties.

[37] Say we have some law of the higher-level theory that relates two kinds *Q* and *R*, such that $\forall x(Qx \rightarrow Rx)$. *Q* is realized on the lower level by the kinds *S1, S2, …Sn*, such that the bridge principle contains a disjunction:
$\forall x(Qx \leftrightarrow (S1x \vee S2x \vee S3x \vee \ldots \vee Snx))$
Let's also assume that the higher-level kind *R* is realized on the lower level by *T1, T2,…Tn*, such that it results in the bridge principle:
$\forall x(Rx \leftrightarrow (T1x \vee T2x \vee T3x \vee \ldots \vee Tnx))$
On the lower level, these realizations are related to each other in smaller laws that are instances of $\forall x(Qx \rightarrow Rx)$, such as:
$\forall x(S2x \rightarrow T3x)$, $\forall x(S3x \rightarrow T1x)$, $\forall x(S6x \rightarrow T2x)$…

To give an informal example, let's say we want to reduce the very high-level law of evolutionary biology expressed by *All species have a means of reproduction* to the lower-level theory of zoology.[38] Here *species* is multiply realized by *humans* and *corals*, and *have a means of reproduction* is multiply realized by *reproduce sexually* and *reproduce by budding*. While *Humans reproduce sexually* and *Corals reproduce by budding* express (lower level) laws that are instances of the law expressed by *All species have a means of reproduction*, we would not say of the following sentence that it expresses a general law: *all things that are humans or corals are things that reproduce sexually or reproduce by budding*.

That the lower-level kinds do not correspond neatly to the higher-level kinds means that there is a good reason to believe we will not be able to reduce the laws of the higher-level sciences in terms of the lower-level ones. We can achieve *some* kind of reduction according to the above method, but the result will not be unified laws at the lower level (Brigandt & Love 2017: §4.2; Fodor 1974: §3).

For the debunker, this means even if it is possible to recast enough propositions of E into E_{Lite} to support $EDO1_{Lite}$ in the debunking argument, the resulting laws of E_{Lite} will lack the *generality* of the laws of E. They will not even look to us like laws, being massively disjunctive. This by itself is a substantial loss in theoretical virtue.

But when these are joined to replicate the form of the law $\forall x(Qx \rightarrow Rx)$, the resulting proposition is radically disjunctive in a way that prevents it from being a unified law:
$\forall x((S1x \vee S2x \vee S3x \vee \ldots \vee Snx) \rightarrow (T1x \vee T2x \vee T3x \vee \ldots \vee Tnx))$
The argument was raised by Fodor (1974), and developed and defended in varying forms in, e.g., Gillett (2003), Aizawa (2008), and Aizawa and Gillett (2011). The formalizations are adapted from Brigandt and Love (2017: §4.2). For a dissenting view, as well as a useful summary of the multiple realizability literature, see Polger and Shapiro (2016).

[38] I won't argue for a definite position on what counts as higher-level or lower-level theories. It is plausible that zoology is lower level than evolutionary biology because the former makes up a *part* of the subject matter of the latter but is subject to its general laws.

The eliminativist may object that we don't *need* smooth, unified lower-level reductions* of higher-level theories. E_{Lite} has other qualities that still make it preferable to *E*. If so, then it is no strike against E_{Lite} that it doesn't match up neatly with the kinds of the higher-level theory: these are exactly the things about which the debunking argument urges skepticism!

The eliminativist may claim that E_{Lite} is superior to *E* because it is not *false*.[39] This would certainly be part of the story for someone already convinced of eliminativism's truth before hearing EDO_{Lite}. However, because *E*'s falsity follows from EDO3, it would be question-begging to invoke this as a reason to prefer E_{Lite}—whose purpose is to *establish* $EDO3_{Lite}$. In the absence of *independent* arguments against ordinary objects, the rest of us can safely suspend judgment on whether E_{Lite} really does possess this particular virtue.

She might also appeal to the fact that lower-level theories display *different* virtues than higher-level theories. Lower-level theories can bring out interesting and important differences between things that appear similar at a higher level; their forte is depth, detail, and precision. Exceptions in the laws of *E*, for instance, often must be explained at a lower theoretical level. Surely these distinctive lower-level virtues count in favor of E_{Lite} (Sober 1999: 560–62).

However, *E* also has access to these lower-level virtues. As *E* retains commitments to composite objects, so it retains the ability to appeal to many different levels of explanation as needed. *E* can explain patterns in entire populations of organisms over time and can relate these to microscopic changes happening in the DNA of individual members. It can take advantage of *localized* reductive explanations without giving up access to higher-level kinds.

[39] Strictly speaking, propositions of *E* may be false in places and lack a truth-value in others, depending on one's view of how false presuppositions affect the truth-value of statements that depend on them.

However, because E_{Lite} eliminates higher-level kinds as a matter of principle, it loses access to such multi-level explanations.

Even if E_{Lite} lacks generality without any clear compensating benefits, we must ask ourselves: does the loss of generality prevent E_{Lite} from supporting the debunker's argument in the needed way? In the next section, I will argue that it does.

7. Would E_{Lite} do the Work Needed by EDO_{Lite}?

The eliminativist may contend that E_{Lite}*'s* messy, gerrymandered laws and explanations would still *do the same work* as *E* in the ways that are needed for the debunking argument. Here are three things that E_{Lite} needs to be able to do, in its own low-level terms. First, E_{Lite} must be able to explain and predict the same range of phenomena as *E within evolutionary biology* by subsuming the relevant quark situations under appropriate laws; second, in order to be justified E_{Lite} must be able to utilize the existing experimental results supporting *E*; and third, it must be able to capture the content of tracking statements of *E*, whose truth depends on identity over time between higher-level entities. I will argue that E_{Lite} cannot accomplish these things.

The first problem E_{Lite} faces is that its laws don't cover the same phenomena as those of *E*. The laws of E_{Lite} lack generality because they do not appeal to higher-level kinds like composite objects. Restricted to this low level, E_{Lite}*'s* laws are necessarily incredibly particularized. Where *E* puts statements of law or observation in terms of increasingly complex *kinds* to express explanations, E_{Lite} must put them in terms of increasingly complex *propositions* about *one* kind (quarks).[40, 41] To capture even part of the content of a law (or

[40] Or one *set* of kinds—it's possible that simples are a diverse group with different properties.
[41] It also makes increasingly weak disjunctive statements as we go up the chain, as opposed to increasingly strong statements of increasing generality.

law-like generalization) of biology such as "All organisms inherit traits from their parents," E_{Lite} must disjunctively list the situations involving quarks that would realize the *atomic* kinds in order to list the situations that would make up the *molecular* kinds, etc., that would ultimately realize the kind *organism.*[42]

Because they are list-like disjunctions of the known realizations of higher-level kinds, propositions of law in E_{Lite} only cover a finite range of phenomena. By contrast, the laws of *E* quantify over general, higher-level terms, giving them a tremendous advantage: they are *open-ended*. The proposition *that all P's are Q's* (where *P* is a higher-level kind) applies to *all* things that are *P*. It makes no difference whether they have been identified or discovered yet. Perhaps some things will become *P's* in the future; our law about *P's* would cover them, too. Perhaps we haven't discovered some *P's* and never will; our law says those are also *Q's*. However, substituting for the kind *P* a *list* of things and saying these are *Q's* is a very limiting strategy.[43] Without reference to *P* as a *kind* of thing, we must just keep adding things to a list and hope we've got them all. Even assuming we're equipped with a complete list of all the known realizations of *P*, our law would still not cover *novel* cases of *P*'s we might encounter in the future. It seems that, except perhaps with some artificially restricted domain,

[42] Quarks have properties like spin, mass, charge, and position. We are to use these properties to express how individual quarks are arranged atom-wise. There are plausibly many, many ways individual quarks can be arranged atom-wise. Even if this can be specified as a mere description of spatial relationships this law will have to account for the varying structures of 118 kinds of atoms—each a different way of being arranged atom-wise. Through complex predicates, our law will specify—and this is just to establish the reference of its subject term—all the ways quarks can be arranged (in concert with other quarks) in ways that count as being arranged atom-wise. The disjointness and overall complexity of the subject terms in statements of E_{Lite} must only be compounded when the eliminativist needs to capture the content of statements of E about putative molecular kinds, so that E_{Lite} can capture the observations and laws of conventional biochemistry. In theory, we can follow this process and build E_{Lite} conversions of more and more complex putative scientific kinds, gradually fleshing out quark-level realizations of genes, cells, organs, animals, ecosystems, and environments.

[43] In the case of E_{Lite}, we would know they also have properties identified by a list of the lower-level realizations of some putative higher-level kind *Q*.

a law made up of lists of *any* length would not adequately capture the propositional content of a law *that All P's are Q's*.

What happens when practitioners using such disjunctive laws encounter some novel phenomenon that formerly would have been included under the kind '*P*'? They must *add* a disjunct somewhere in the appropriate law. This reveals a further oddity of such laws: whereas laws in terms of higher-level kinds can absorb new empirical data without changing, laws in terms of lower-level lists must change constantly to retain their predictive and explanatory power. Thus, no *single* lower-level law, not even the most up to date one, does the same work as the higher-level law it reduces.[44, 45]

In the case of E_{Lite}, we might expect this incomplete capturing of the content of higher-level laws to be compounded by the many levels of reduction* necessary to move from human perceptual psychology all the way down to quantum physics. The upshot is that E_{Lite} is crippled in its ability to explain or predict novel cases explained or predicted by *E*. The results could be catastrophic for EDO_{Lite}. For instance, is *the target audience for* EDO_{Lite} covered by these laws, assuming their quantum structure is not already spelled out in the laws' particulars? Are *our ordinary object beliefs* covered by the laws? If not, then why should we listen? It's possible to answer these questions favorably for E_{Lite}; however, to do

[44] Clearly, higher-level theories and their laws also need to be revised in light of novel data. But they are insulated from the kind of persistent reformulation described above by subsuming a wide range of potential data under general kinds. New data add supporting detail to the theoretical explanations supporting higher-level laws; but the laws themselves are stable over time, except in the rare cases where a specially designed experiment produces confuting evidence.

[45] Consider a law L_1, which covers several low-level realizations of both *P* and *Q*:
$\boldsymbol{L_1}$: $\forall x((P_1x \vee P_2x \vee P_3x) \rightarrow (Q_1x \vee Q_2x \vee Q_3x))$
Scientists later discover some new things that would have been considered realizations of the putative kinds *P* and *Q*, such that they produce a new law L_2:
$\boldsymbol{L_2}$: $\forall x((P_1x \vee P_2x \vee P_3x \vee P_4x) \rightarrow (Q_1x \vee Q_2x \vee Q_3x \vee Q_4x))$
The law L_1 does not cover P_4 or predicate Q_4 of any *P*s. Practitioners can keep producing new laws that include new realizations, but this generates a series of different laws, and no *single* law—not even the most inclusive, up-to-date version—does the work of the law *that all P's are Q's.*

so the eliminativist must find a way to recover some generality in a way that's motivated within E_{Lite} itself and that doesn't rely on illicit appeals to the higher-level kinds it rejects.

The second problem for E_{Lite} concerns the nature of the existing experimental evidence for *E*.[46] Unfortunately, all the experiments conducted and observations made to test *E*—from sciences ranging from zoology to microbiology—were not designed to measure the behavior of quarks. In fact, every experimental finding regarding *E* has been *radically* imprecise as to what the quarks were doing in the situation. Assuming that we already have some serviceable low-level law in terms of quarks, we would not know if some particular experiment supporting *E* confirmed or confuted it, or whether it represents some new quark-situation that needs to be added to our law for it to remain complete and current. Thus, in the absence of any recourse to generalities—even in terms of quark-wise things—E_{Lite}*'s* relation to the experimental evidence is unclear, as is its justification.

Finally, E_{Lite} has a problem expressing *identities over time* between higher-level entities covered by *E*. Of course, the eliminativist doesn't *believe in* these higher-level entities. But it remains a problem: for instance, the eliminativist needs to be able to express (in low-level terms) why the *identical* human organism who just had some visual experiences caused by quarks arranged object-wise now believes there is an object in front of him. This in turn depends on a story about why, of each member of a crucial set of ancestors, the *identical* ancestor that had a certain perceptual trait also had higher reproductive fitness than its rivals. Perhaps many details are dispensable for the purposes of EDO_{Lite}, but some low-level version of this central story is *not*. Likewise, the broad evolutionary picture that supports and justifies

[46] Conversely, the few experiments that have dealt directly with observing quarks have had the purpose of relating them to other subatomic particles to determine their nature and properties. These experiments were not done for the specific purpose of testing *E*, and the light they shed on *E*'s justification is correspondingly dim.

this story—from many subfields of biology—involves *tracking* individual organisms through their development, mating, and adaptive relationship with their environments. Without some way of appealing to identities between members of higher-level categories, E_{Lite} simply lacks the vocabulary to express this crucial explanation of our object beliefs. In E_{Lite}, the only thing capable of being identical to itself is a quark.

In the next section, I examine a promising strategy for solving all three of these problems by appealing to *pluralities* of quarks arranged *K*-wise, and to kinds built up in those terms.

8. Pluralities of Quarks Arranged *K*-wise, Arrangements *K*-wise, and Shmidentity

The eliminativist may raise the following objection. Surely, we—and field biologists—can say *something* about the quark-situation just based on what we can observe with our own eyes. What's causing the finch-wise experience I'm currently having? A *plurality* of quarks arranged finch-wise. I can make observations about the identical plurality over time, tracking it through changes. Similarly, I can convert propositions of observation from conventional experiments made about finches into propositions in terms of *pluralities* of quarks arranged finch-wise. These observations and experiments can then support E_{Lite} in roughly the same way they supported *E*. In addition, we can subsume all the low-level particulars about how quarks are arranged finch-wise under the *kind* "pluralities arranged finch-wise." We can also generalize to *kinds of kinds* of pluralities, and so on, using these to formulate object-free propositions of law and explanation at whatever level we please. Soon, E_{Lite} is a theory as robustly general as *E*—open-ended and covering all phenomena relevant to EDO_{Lite}. This seems to take care of the problems with lack of generality outlined in Section 7.

For this strategy to work, the propositions of E_{Lite}, in terms of pluralities of simples arranged *K*-wise, must capture the content of propositions about *individual* objects of the kind *K* in *E*. Only then can E_{Lite} generalize about kinds of pluralities, kinds of kinds of pluralities, and so on in a way that matches the attributions in *E* in the ways needed to support EDO_{Lite}. A proposition of E_{Lite} captures the content of a proposition of *E* if and only if it's true to attribute things to the plurality (or kind of plurality, etc.) of simples arranged *K*-wise in the proposition of E_{Lite} that are attributed to the object (or kind of object, etc.) *K* in the proposition of *E*. In other words, pluralities arranged finch-wise need to behave exactly like (putative) finches.

This demand for content capturing is not arbitrary: remember that the close correspondence between the content of propositions of E_{Lite} about simples arranged *K*-wise and that of propositions of *E* about some object *K* both explains the trustworthiness of *E* and allows E_{Lite} to share in *E's* epistemic virtues and justification.[47] The content of any proposition of *E* that fails to have a corresponding proposition in E_{Lite}—as well as its justifying, explanatory, or predictive value—would be lost to E_{Lite}, and so would the content of any propositions dependent upon it. If large classes of important propositions of *E* were in-principle uncapturable for $E_{Lite,}$ the results would be catastrophic for EDO_{Lite}.

I argue that proponents of E_{Lite} face a trilemma here. If they simply recast propositions of *E* in terms of pluralities arranged *K*-wise, the resulting surrogate propositions inevitably fail to capture any content that involves composite objects such as finches persisting over time. Alternately, they can supplement propositions in terms of pluralities arranged *K*-wise with a new metaphysical relation (I'll call it 'shmidentity') that obtains for

[47] Cf. Section 4 on incompatibilism and the end of Section 5 on the dependence of E_{Lite} on *E*.

pluralities arranged *K*-wise over time, allowing such propositions to capture the content related to object persistence, but at the cost of introducing a strange and unparsimonious metaphysical relation into all corners of the science; lastly, proponents of E_{Lite} can say that *arrangements K*-wise are what persist over time in its converted scientific propositions, but this introduces new entities into their ontology that have the earmarks of composite objects.

Eliminativists pursuing the strategy of generalizing in terms of pluralities arranged *K*-wise must reckon with the fact that pluralities of quarks arranged finch-wise have different persistence conditions than do (putative) finches. At time t_1, some plurality P_1 includes all and only the quarks arranged finch-wise during some scientific observation of an individual finch *F*. But change one quark and a plurality of quarks is no longer the *same* plurality. Organisms like finches are constantly changing on the microscopic level, metabolizing food into tissues and passing the rest as waste, sloughing off feathers and dead skin, sustaining small injuries, or simply growing and aging. At t_2, milliseconds later, two things have happened: first, P_1 is no longer arranged finch-wise, as some of the quarks in this plurality have passed out of finch-wise arrangement; second, the quarks of some *different* plurality P_2 are all and only the quarks populating *F*. In fact, during any observation of a single finch *F* over times $t_1 \ldots t_n$, scientists are observing a *succession* of pluralities of quarks arranged finch-wise, $P_1 \ldots P_n$.[48, 49]

Ultimately, no single plurality of quarks does the causal work of any individual finch *F*, because none remains arranged finch-wise long enough. Rather, a shifting group of quarks is involved in the causal work of a finch over time, with new sub-groups of simples being

[48] I use 'populate' or 'belong to' as an ontologically neutral way of specifying which quarks are arranged *K-wise* in any particular arrangement *K-wise* at the time in question.
[49] For economy, I will hereafter just write 'finch' or 'organism' in this section. But the reader should hear 'putative' in front of any term that presupposes commitments to ordinary objects.

shuttled in and out every millisecond. This means the proposition *that the finch that laid this clutch of eggs is the same finch that did not reproduce last year* is not captured by any corresponding proposition about identical pluralities arranged finch-wise. Even if the proposition is true—that is, if identity holds between a finch and itself—it is false when converted into a proposition about two pluralities arranged finch-wise.

The eliminativist could respond by introducing a new relation that applies to pluralities over time, such that a set of pluralities $P_1 \ldots P_n$ are 'shmidentical' at times $t_1 \ldots t_n$ as long as the quarks were replaced in a suitably gradual manner at each stage.[50] Even if P_n comprised an *entirely* different set of quarks at times t_1 and t_n, it could still qualify as the shmidentical plurality to P_1 if it met the condition for gradual replacement. This would seem to circumvent the problem with the above proposition about the egg-laying finch. However, these conditions are too loose. Over enough time *any* two pluralities would be shmidentical, such as a finch and the tree in which it makes its nest. Nor is it sufficient to tie shmidentity over time to *being arranged K*-wise consistently over time. For instance, a plurality of quarks arranged finch-wise might belong at t_1 to a mother about to lay a clutch of eggs and at t_n be distributed between her and her three chicks; this would not allow E_{Lite} to capture, for example, propositions exclusively about the mother during $t_1 \ldots t_n$. Nor would this strategy forbid our propositions from tracking random or uninteresting pluralities of finch-wise quarks from $t_1 \ldots t_n$, such as part of a beak or talon. We need shmidentity to apply exclusively to successions of pluralities arranged *K*-wise that are made up of all and only those quarks that populate *a K*-wise *arrangement* corresponding to a particular (putative) finch over time. So, ultimately shmidentity conditions must piggyback on our *identity* conditions for finches.

[50] See, e.g., Contessa (2014: 213–14) for a version of this strategy defending a somewhat different position.

However, note that the eliminativist has introduced a strange and unparsimonious new metaphysical relation that has to be built into E_{Lite} at every level.[51] Shmidentity holds between two *K*-wise pluralities over time whenever *identity* would hold between two composite objects of the kind *K*. This is a problematic reliance on counterpossible facts. It's true that counterpossibles occur in scientific theories quite regularly. For instance, they feature in the antecedents of counterfactual conditionals whose purpose is to explain why some actual property of something is doing what it really is doing in contrast to an another (impossible) situation that would yield a different outcome (Tan 2019). However, these are localized, limited explanations. Shmidentity is a widespread relation that features crucially in the positive propositions of law, observation, and explanation of the theory, and it can *only* obtain between nonexistent objects. This is a *radically* different kind and level of dependence on counterpossibles from what is normally encountered in the sciences. The counterpossible facts about identity conditions between nonexistent objects would seem to be fundamental, and as numerous as there are kinds of nonexistent composite objects—hence the loss of parsimony.[52, 53]

The third strategy for an eliminativist is the simplest: jettison the notion of shmidentity and claim that *arrangements K*-wise are things that can persist over time

[51]Another oddity is that shmidentity seems to depend on clear and definite identity conditions in a way that ordinary science does not. *E* proceeds unimpeded despite such identity conditions never having been specified clearly or in detail, but the very definition of shmidentity presupposes the existence of those conditions.
[52] These facts cannot be reduced, e.g., to more fundamental facts about nonexistent objects, for there are no such facts. I'm assuming here that the eliminativist would not want to say counterpossible facts about nonexistent objects are reducible to facts about impossible worlds.
[53] The contrast between the usual kinds of occurrences of counterpossibles in scientific theory and the shmidentity relation can be illustrated by comparing two cases. In the first case, explaining entropy by appealing the counterfactual: if a machine were indeed a perpetual motion machine, it would never need an infusion of energy from the outside. (This explains why *real* machines need energy to run.) In the second case, taking some property that *only* perpetual motion machines have, and attributing it to *real* groups of machines described by one's theory.

separately from any particular quarks or pluralities of quarks. An arrangement finch-wise needs has the same properties as finches do in *E*, including persisting under whatever conditions a finch would in *E*. We assume here that an arrangement has a fluctuating population of quarks and pluralities of quarks but is arranged in the right way over time to sustain these higher-level properties.

But notice that the eliminativist's ontology now looks very much as it would if it included composite objects. An arrangement is not a quark, nor is it any particular plurality of quarks. But it exists and bears attributes referenced by the propositions of law in E_{Lite}—including causal powers—that no quark or plurality of quarks could bear. Arrangements finch-wise are new entities that behave very much like composite objects. Perhaps they are finches?

Ultimately, the eliminativist seems unable to recoup generality in terms of kinds built from pluralities of quarks arranged *K*-wise without incurring great costs in the process. This implies that E_{Lite} will indeed be made up of incredibly complex, particularized propositions about quarks, and will be subject to the limitations I outlined in Sections 6 and 7. These are fatal liabilities for the view, indicating that E_{Lite} is inadequate to run the debunking argument.

This mismatch between *E* and E_{Lite} has another unattractive consequence for the eliminativist. Because a vast range of crucial propositions of *E* cannot in principle have a corresponding *K*-wise situation to be captured by a proposition of E_{Lite}, they are not *false but nearly as good as true.* Rather, they and the substantial chunk of evolutionary biology that depends on them are simply *false*—as false as the belief that unicorns are right now trotting across the rainbow.

9. Conclusion

If my argument has been successful, I have shown four things:

1.*That there is a self-defeat problem facing the evolutionary debunker of ordinary objects.* Evolutionary theory and its body of evidence depend on ordinary objects, and debunkers will need to reckon with this problem. I am not optimistic about the prospects for an eliminativist solution. I believe this argument generalizes even to more nuanced kinds of eliminativism that establish exceptions for certain kinds of objects, such as organisms or conscious beings. Evolutionary theory seems to require ordinary objects on a very wide scale to tell its story; the inanimate, unconscious objects making up organisms' environments are an indispensable part of that story.

2. *That eliminativists who utilize* K-*wise conversion strategies, believe in a complete low-level causal story of the world, and appeal to the results of the special sciences commit themselves to some form of scientific reductionism.* The alternative is to appeal to a completely unknown, untested theory. This applies to eliminativists who run EDO_{Lite} but are, for example, instrumentalists about science and claim only to be pointing out a conflict between conventional scientific realism and beliefs about ordinary objects. Without such a reduction, the scientific realist has no grounds for accepting the argument. Eliminativists who run debunking arguments against other kinds of beliefs (e.g., moral or aesthetic) face no self-defeat problem, but must reckon with the tension between their ontologies and the claims of evolutionary biology.

3. *That to recast any propositions of* E *referencing ordinary objects of some kind* K *as propositions about* pluralities *of quarks arranged* K-*wise in* E_{Lite} *is problematic.* To capture

the needed content of the propositions of *E*, the proponent of E_{Lite} must find a way out of my trilemma as presented in Section 8.

4. *That a theory* E_{Lite} *resulting from a systematic, eliminative reduction* of* E *would have insufficient justification and explanatory power to support the debunking argument.* As a theory on the level of quarks without recourse to generality in terms of *pluralities* of quarks—let alone any higher kinds—E_{Lite} must have laws that are incredibly particularized. Thus, it sacrifices not only necessary breadth and power in the form of general laws and explanations, but a critical range of observations as well. As a result, it cannot express relevant evolutionary explanations in support of EDO_{Lite}, and its justification is in serious jeopardy.

My essay has said little about permissivists, but they sometimes use debunking arguments to establish that there is no reason to believe that *only* ordinary objects exist. Given all the ways the universe *could* be carved up into objects, if our object beliefs happen to be true and all and only the ordinary ones exist, this could only be the result of incredible luck. Addressing this kind of debunking argument will have to wait for a future work, but much of what I've said here will apply to permissivists who accept Composition as Identity or some weaker whole-part reductionism; when appealing to *E*, they will have to deal with some of the same problems I've described here for the eliminativist debunker.[54]

[54] See Korman (2015: Ch.7) for a useful overview of various versions of the debunking arguments, as well as Kovacs (2019) and Barker (2019) for recent, related discussions of commitments among revisionary ontologists.

Chapter 2:
Externalism, Object Debunking, and Reliable Misrepresentation

1. Introduction

Certain evolutionary debunking arguments pose an explanatory challenge to the realist about composite objects. Their skeptical conclusion rests on the claim that our object beliefs are not best explained by the object facts, but rather by our evolved predispositions to perceptually represent the world as containing composite objects even if they don't exist. Some argue that externalism about natural kinds renders the debunker's challenge harmless because it ties our object beliefs to natural kind members in the outside world via causal-historical connections that partially determine the content of our perceptual and conceptual representations; others find externalist responses insufficient for meeting the skeptic's challenge.[1] I will develop and evaluate two hybrid externalist views that can field responses to the debunker's challenge while avoiding illicit *a priori* reasoning about the empirical world. I will then show that the conservative about objects provides the most promising response.

In its most minimal form, externalism about some mental state *F* is simply a thesis about how that state is individuated: whether a person is in state *F* depends (at least partially) on her relation to her external environment. For an externalist, two physically identical individuals can be embedded in different external environments such that one is in the mental state *F* and the other is not.

[1] For an example of the former, see Fairchild and Hawthorne (2018: 50-51). For an example of the latter, see Korman (2014: §6).

Some mental states are intentional: they have *content.* Beliefs and desires, for instance, have for their content what is believed or desired. The contents of attitudes like belief and desire are propositional and are thus determined in part by the concepts that make up these propositions. Externalism about mental content is the view that our mental (conceptual or perceptual) representations are partly determined by their causal relation to the physical or social environment. *Natural kind externalism* is the view that the content of our natural kind concepts is partly determined by a causal-historical connection to natural kind members in the external world. *Social externalism* is the view that the content of some of our concepts is partially determined by causal-historical relationships to our language-using community. This dependence of our concepts on causal-historical relationships with our environment affects the reference of our terms, and so this kind of externalism is sometimes called *semantic externalism.* Following common usage, I will simply take *content externalism* to be the view that in mental states with content, that content may be partially determined by causal-historical connection to things external to the thinker's biological body. This characterization is broad enough to include externalism about natural kinds, our social environment, and the meaning and reference of our terms.

In this chapter, I will examine three externalist strategies for responding to the debunker's explanatory challenge. Here is a road map for the journey. In Section 2, I present the object debunker's explanatory challenge, developing the role of reliable misrepresentations in her argument. In Section 3, I introduce the first possible externalist response, what I call the *a priori* route, wherein the externalist can reason a bit too easily from visual experiences of trees to the existence of trees. In Section 4, I motivate and develop a hybrid externalist view that combines an externalist causal-historical criterion for the

veridicality of representations with a separate matching condition; this forms the foundation for the next *two* views. In Section 5, I develop a response to the debunker from the first of these hybrid externalist views, non-eliminative nihilism, and show its limitations. In Section 6, I develop a response to the debunker from the second hybrid externalist view, conservatism about midsize composite objects. I show why, according to conservative view, composite object facts can figure into our best explanation of our perceptual object beliefs.

2. The Object Debunker's Explanatory Challenge

Before proceeding, let us put the debunker's argument in a form that stresses its explanatory challenge to the object realist. It will simplify our task to focus on the concrete example of trees, knowing that we can zoom back out to the broad class of composite objects as needed. As in chapter one, we will assume the object debunker is a nihilist about composition who believes the only objects are mereological simples. To distinguish it from EDO from Chapter 1, we will call the present argument "EDO*':

(EDO1*) It's false that your tree beliefs are best explained by the tree facts.

(EDO2*) If so, then you are not justified in retaining your tree beliefs.

(EDO3*) So, you are not justified in retaining your tree beliefs.

The basic idea behind EDO1* is that what *really* explains your tree beliefs are biological contingencies in your ancestors' evolution.[2] The debunker holds that you only visually carve up the world into objects the way you do because you are evolutionarily hard-wired to do so

[2] Generally speaking, EDO1* garners explanatory support from the biological *and cultural* contingencies in one's ancestors' evolution. However, as our focus will be on visual representations, I will put cultural contingencies to the side for the present chapter.

in the presence of simples arranged treewise. For instance, you are predisposed to believe there are trees because your visual system automatically carves up your visual field into trees and other reproductively relevant clumps of matter because these traits enhanced your ancestors' reproductive success. However, it is possible that the world is carved up into strange macroscopic objects very different from the ones we're predisposed to see—or that there are no composite objects at all. It is possible that our ancestors' visual systems got it right about how matter is arranged and distributed but got the facts about *composition* all wrong—wrong in a way that increased their reproductive fitness. Because we can offer this kind of explanation *without* reference to our ancestors getting the facts right about composition, composite objects like trees seem irrelevant to our explanation: even if there *were* trees, they wouldn't explain our tree beliefs.[3]

EDO2* captures the idea that learning that EDO1* is true would be a defeater for our normal reasons for holding tree beliefs. Perhaps we are normally justified in believing there's a tree when it *seems* like there's a tree in front of us. Learning EDO1* is true would undercut this justification because EDO1* gives a plausible alternate explanation of why it seems there's a tree in front of us—an explanation that has nothing to do with there *actually* being a tree in front of us.

To be more precise about the explanation motivating EDO1*, the debunker can argue that we don't have to posit the existence of composite objects like trees to explain our tree

[3] Mogensen (2015) argues that just because moral facts don't figure in the evolutionary (ultimate) explanation, that doesn't mean they don't figure in the proximate explanation. This would apply to trees here as well. I will not be pursuing this intriguing objection in the present chapter.

beliefs if we hold that apparent composite objects are a case of *reliable misrepresentation*.[4] For a representation to qualify as an instance of reliable misrepresentation, the following must hold:

i. there is a *mismatch* between what is tracked and what is represented (i.e., one thing is represented, and a *different* thing is tracked by means of the representation);
ii. nothing *exists* matching the representation; and
iii. the thing that is tracked is tracked *reliably*, such that the adaptiveness of the misrepresentation can be explained.

Let us first briefly consider how reliable misrepresentation works in a separate case. It's plausible that we perceptually represent midsize physical objects as having the *intrinsic* property of heaviness.[5] A dumbbell at the gym that weighs 100 pounds is experienced as heavy by the average person—as difficult to lift. But the dumbbell has no such intrinsic property (ii). What we are really tracking is a different, *relative* property: its mass as it relates to the earth's gravitational field and to the lifter's strength (i). Within our ancestral environment, this is a property we can reliably track (iii) because the earth's gravity and our rough lifting abilities are constant. We are just wrong about the nature of the property. Outside of this environment (e.g., on the moon), the same dumbbell may be very easy to lift.[6]

[4] This notion comes from Mendelovici (2010, 2014) and is used in an object debunking argument in Korman (2019). It's worth noting that Mendelovici considers the very possibility of reliable misrepresentation a problem for tracking theories of mental representation—and thus for externalism—generally. I will not deal with this charge in the present chapter but will focus on whether object representations specifically are a case of reliable misrepresentation.
[5] We might do so through tactile or kinesthetic representation, for instance.
[6] See Shoemaker (1994: §2) for an account of heaviness as relational rather than intrinsic. Mendelovici (2010: §5.1) gives heaviness as a potential example of reliable misrepresentation.

The debunker's claim is that there are no composite objects in the world (ii), but visually we reliably misrepresent pluralities of simples as midsize composite objects. We misrepresent the instantiated properties of a plurality of simples as being instantiated by a single composite object (i). By doing so, we *reliably track pluralities of those simples* (iii) such that we can successfully and veridically represent lots of their features: which direction they're moving, their distance from us, their collective size. The fact that we reliably track these pluralities allows us to secure representations that, though false, help us believe—and so behave—in a way that enhances our reproductive fitness.

According to the debunker, this explanation is better than the composite-dependent explanation in several ways. The story goes like this. First and foremost, it is more ontologically parsimonious: it makes no commitment to composite objects of any kind. Second, it illuminates why we represent things as being carved up into midsized objects: this is just our anthropocentric way of tracking the simples that are reproductively important to *us*. Other species with different needs might carve up the world quite differently, and still track reproductively important matter reliably. For instance, imagine a species whose reproductive needs made it advantageous to track woody, tree-trunkwise matter along with furry, dogwise matter. These creatures would carve up the world into *trogs* rather than *trees* and *dogs.* In our case, it was simply unimportant to track woody and furry matter together, but it was extremely important to track woody and *leafy* matter together. Hence, we track *trees* and *dogs* rather than *trogs.* Our object beliefs are merely a convenient way of tracking matter for our species' biologically contingent reproductive needs. Finally, the debunker's explanation is clearer than the externalist one. Given that the real explaining is done in terms of tracking the movements and distributions of reproductively important *simples*, it's obscure

how any facts about the way the world is really carved up into midsize composite objects *could* feature into such an explanation. If we were to add to this already adequate explanation of our object beliefs *that the simples arranged treewise really do compose trees,* it would not illuminate in any way why it's adaptive to divide the world into trees. This being the case, we are no longer justified in retaining our object beliefs based on how things visually *seem.*[7]

In Chapter 1, I argued that evolutionary debunking arguments are too corrosive: their conclusions undercut the nihilist's own premises because they threaten her appeal to evolutionary biology. I will not recap these concerns here, but I do want to answer the worry that the debunking argument is too corrosive in another way. One might argue that this argument can be generalized such that *no mereological facts*—not even the nihilist's own belief that composites do not exist and there are only mereological simples—bear on evolutionary explanations of our visual representations. For instance, a certain kind of creature might evolve to perceive mereological simples directly, and in such a way that they are not presented as being *part* of anything, such that it appears that only simples exist, but no composites. This kind of creature might even be wrong about this, and composition may very well happen, but the creature's own reproductive needs lead its visual system to represent the world as populated *only* by mereological simples. In this case, if such a creature were to hear the debunking argument, it too should abandon its mereological beliefs—that composition never happens—at least based on how things are represented visually.

I think that the nihilist can easily bite this particular bullet and say that, yes, everyone should avoid making any mereological conclusions based on how things visually seem to be.

[7] See Korman (2019: §13.4) for a more detailed presentation of this argument and the role of reliable misrepresentation, as well as comparisons to the case of debunking moral facts. See also Sharon Street (2005) for the version of the moral debunking argument after which this is modeled.

The nihilist would not lose much here, because for her intended audience things seem to be divided up into midsize composite objects, and it is this intuition she seeks to undermine. After all, the nihilist's own positive arguments do not appeal to how things seem, but to solutions to metaphysical puzzles about composition.

In the next section, I will examine one way that an externalist could respond the debunker's explanatory challenge in EDO1*. The externalist could attempt to reason a priori to the existence of composite objects. I will explain why this is a bad idea, and in subsequent sections I will explore more promising routes for the externalist.

3. Externalist Response #1: The *A Priori* Route

At first glance, it might seem too easy for the externalist to reply to the debunker's skepticism. After all, doesn't externalism presuppose that in thinking about natural kinds, or about samples or members of natural kinds, we already have a kind of causal-historical connection to them that entails the existence of natural kind members? In this section, I will briefly sketch this *a priori* reasoning from thoughts about natural kinds as well as a parallel, at least partially *a priori* argument identified by Mendelovici (2013) that can be made from externalist theories of perceptual representation. I will show why it is a legitimate worry for externalists, and a poor way of responding to EDO1*.

Some have argued that if externalism were true the mere possession of a natural kind concept such as *water*, plus the privileged access we seem to have to the contents of our own thoughts, allows us to reason to the existence of water simply by having thoughts about water (see Mckinsey (1991), Brown (1995), and Boghossian (1997)). The putative externalist argument for water realism goes something like this:

1) I am having a thought about water (e.g., *that water is wet*).

2) If I am having a thought about water, then water exists.

3) So, water exists.

Premise (1) is something I can know simply by introspection. Premise (2) is an entailment of content externalism. The idea is that if content externalism is true, and I do indeed possess the concept of water, then I and my linguistic community must be embedded in a world that has the right kind of causal-historical connection to the natural kind *water*. Water must be causing our water beliefs, and our concept of water is based on causal connections to samples of actual water, all of which share the same basic physical constitution.

That we can know (1) through introspection alone, without any empirical investigation, follows from its being a belief about my own mental states. This is supported by the traditional notion that we have privileged access to our own thoughts and their contents. I take this to mean that we know that we are having the thought that *p* simply by having it, and that we can identify its conceptual content. As noted above, (2) is an entailment of content of externalism, and it can be known *a priori* if one accepts content externalism. Thus, the empirical conclusion that water exists follows from introspection (1) and *a priori* reasoning (2) alone.

This is taken to be a *reductio* of externalism. We should not be able to reason *a priori* that natural kinds or their members exist, because empirical investigation is relevant to whether we can know water is a natural kind or whether samples of water actually have the qualities contained in our concept of water (such as being, in its distilled form, clear, potable, and composed of H_2O).

To avoid this consequence, the argument goes, we must give up externalism, or we must give up the doctrine of privileged access. Neither is an attractive option. Some externalists have fielded responses to this charge, arguing that externalists aren't forced into giving up either privileged access or externalism.[8] This debate is outside the scope of this essay, but I want to explore a related move that is at least *possible* for externalists in the area of perceptual representation.

Mendelovici (2013: 238-40) points out that just as externalists can make the above argument for realism about water, so can proponents of tracking theories of representation make a problematic—and at least partially *a priori*—argument from our visual representations *as of* some member of a natural kind x, to *realism about x*. By "tracking theories," she means theories of mental representation that take representation to be a matter of detecting, carrying information about, or correlating with the environment in some way. Causal-historical theories of perceptual representation would be tracking theories under Mendelovici's characterization.

Suppose an externalist extends the causal-historical theory of mental content into a causal-historical theory of visual representation. Such an externalist would hold that our visual representations as of trees have trees as both a proximal and distal cause. I am experiencing right now a visual representation *as of* a tree because a tree is present and reflecting light within my visual field (it is a *proximal* cause). However, trees also have a

[8] See Mckinsey (1991), Brown (1995), and Boghossian (1997). See also Brueckner (1992), McLaughlin and Tye (1998), and Korman (2006) for externalist responses.

deep, complex causal-historical relationship with our ancestors' visual systems going back millions of years (they are a *distal* cause).[9]

The idea is that being able to detect, distinguish, and in general to track trees was a selective pressure on the evolution of our ancestors' eyes and brains. Just as our mental states when we are thinking about trees are individuated by a causal-historical connection to actual trees, so are our visual representations *as of* trees. Long before our ancestors began having explicit thoughts about trees (or using the word "tree" or its linguistic precursors), our species' visual representations were being molded by a causal-historical connection to trees in our ancestral environment—specifically, the selective pressure to represent trees reliably for the purpose of navigating, finding food and shelter, etc.

Here is the partially *a priori* argument from a tracking theory of mental representation to the existence of trees:

1) I have visual experiences as of trees.
2) Some of these visual experiences as of trees occur in situations where they have an appropriate causal connection to something in my environment (they track something reliably).
3) If (1) and (2) are true, then trees exist.
4) So, trees exist.

Premise (1) is something to which the debunker will agree. We have visual experiences *as of* there being trees in front of us all the time. Nothing follows from (1) alone about the veridicality of these experiences. In premise (2), the appropriate causal connection could be

[9] Again, for the purposes of this chapter we are mostly bracketing the important role of the linguistic community.

spelled out in various ways. For the sake of the current argument, let us just say that the causal connection needs to be to those things that caused tree experiences under content-endowing conditions in our ancestral past—perhaps conditions under which there was selective pressure on the evolution of our ancestors' visual systems to be able to discriminate and reidentify these things. Premise (3) is an entailment of the causal-historical theory of visual representation in question. If I have visual experiences as of trees, and at least some of these experiences bear an appropriate causal connection to the content-endowing conditions for tree representations, then my visual system must be sometimes *successfully representing* trees; trees must exist. Subconclusion (4) follows from (1)-(3).[10]

It is a short step from this kind of argument to the inference that there are composite objects. If our only condition for realism about trees is that sometimes our tree experiences occur in circumstances where our causal condition is met—if this is all that is required for our representations as of trees to be *veridical*—then we can infer not only that trees exist, but that they have at least some of the properties they seem to have. We can do this without any additional checking about these properties.

It is highly plausible that compositeness is a property represented in our visual experiences as of trees. When I have a visual experience as of a tree, it is as of an object with parts I can differentiate at a glance by size, color, texture, location, number, and movement. The trunk (large, brown, rough, singular, and at the bottom) stays put, while the leaves (small, green, smooth, multifarious, and at the top) sway wildly in the breeze. If compositeness is represented in our tree experiences, then if it is an *essential* property of

[10] Compare Tyler Burge (2010: 82-87), whose account of perceptual content has a causal component, and who also appeals to content-endowing conditions. Burge's account inspired the present investigation, but his is not a strictly causal account, nor would he endorse the *a priori* route discussed here.

trees—such that something cannot be a tree without being a composite object—we can infer that it is instantiated as well. I will not argue that compositeness is an essential property of trees, except to point out that in disagreeing with this idea, the non-eliminative nihilist is staking out a new and controversial position (see Section 5).

Here, of course, the debunker would insist that it is possible that our representations as of trees are reliable misrepresentations of pluralities of simples arranged treewise. However, our causal-historical theory of visual representation does not allow for misrepresentation of trees when the causal condition is met. The debunker would say: so much the worse for your theory of representation. Mendelovici agrees, considering it a strike against all tracking theories that they close off certain empirical possibilities for misrepresentation (435).

Mendelovici argues that we should not be able to make the kind of inference from (1)-(4). Her reasoning runs as follows. We should not be able to deduce from our theory of visual representation that realism about trees is true simply because we know that sometimes our tree experiences occur in circumstances where our causal condition is satisfied. At a minimum, we should have to do some additional confirmation that there are, say, leafy and woody midsize objects out there. It is possible that the things currently causing our tree experiences, even if they bear appropriate causal connections to content-endowing conditions, are *not* trees: they might be reliably misrepresented *as* trees. Although arguments like the above are not entirely *a priori* because (2) includes some information the causal history of the representation in the external world, it still seems like illicit reasoning of some sort. What one *could* infer *a priori*, given the truth of our causal-historical theory of visual representation, is the conditional "If I have visual experiences as of trees, then either trees are

real or my tree experiences occur in conditions where the causal condition is not satisfied." A theory of mental representation alone should not allow us to infer *a priori* from visual experiences any kind of ontological or causal conclusions about the external world (439-40).

Fortunately, this is not the only option for the externalist. In the next three sections I will show that while this objection seems dire for a simple causal-historical theory of visual representation, it does not apply to hybrid theories of visual representation that include causal conditions. First, I will motivate the inclusion of a non-causal veridicality condition in an externalist theory of visual representation.

4. The Possibility of a Hybrid Route: Motivating a Separate Veridicality Condition

In Sections 5 and 6 I will examine two hybrid externalist strategies for resisting EDO1*. The first strategy, an application of non-eliminative nihilism, claims that trees just *are* pluralities of simples arranged treewise, and that they are reliably misrepresented as composite objects. The second strategy, an application of a conservative view about midsize composite objects, holds that reliable misrepresentations are possible but that the best explanation of our tree beliefs is that they are reliably represented *as* composite objects. Each strategy will require the ability to accommodate reliable misrepresentation. The focus of this section is to motivate this move and sketch how a hybrid externalist view works in broad strokes before developing the two more specific strategies mentioned above.

The first motivation for a move to a hybrid theory is, of course, that Mendelovici is right. Theories of mental representation should be able to account for cases of reliable misrepresentation, and purely causal-historical theories cannot do so. On a purely causal-historical theory, once some property P is tracked reliably via the appropriate causal-

historical connection with a representation Q, the property P just *is* what is visually represented. But it is a perfectly plausible situation that Q really is a different property than P, a putative property of things in the external world that is simply not instantiated: nothing in the world *matches* Q. There is no reason why a theory of mental representation should forbid this perfectly plausible empirical possibility.

An obvious solution would be to add a condition for matching or descriptive accuracy to the causal condition, and that is what I am going to suggest. One might object that the move to a hybrid theory of visual representation that contains both a causal condition and a descriptive condition on representation is at best an *ad hoc* response to Mendelovici's challenge, at worst incompatible with externalism. I will briefly explain why neither is the case.

Fortunately, externalism is not committed to a *purely* causal-historical theory of visual representation. Externalism only holds that where there is a relevant causal-historical connection, this may partially determine the content of our visual representations. In cases where there is an appropriate causal-historical connection with samples of a natural kind, this connection *will* at least partially determine the content of visual representations. But an externalist is free to hold that there are other conditions relevant to the veridicality of our visual representations—even representations of natural kind samples.

The following example illustrates how matching conditions and causal conditions interact and constrain what counts as a representation. A hallmark of a visual representation, whether in the human visual system or in a photograph, is that the representation matches its subject. (This marks a difference between these and abstract, symbolic representations like those in spoken or written language.) Similarity between the properties represented and the

properties of the thing represented have obvious value: they can help us identify and reidentify the represented thing and inform us of its states and properties. I can recognize my terrier, Rex, by the way he looks and acts, tell him apart from other dogs, and tell when he needs food or a walk in the park.

However, consider a photographic representation Rex that was taken last Thursday. Let's say it is a good, accurate representation of Rex, such that it meets any reasonable condition of matching between the photographic image and the real terrier. They have the same proportions and color profile; perhaps the photo is life-sized, such that it accurately represents Rex's height and width. Now imagine that Rex has a twin brother Chex who looks exactly like him. This photo is also a perfect match for Chex and his features. So, what makes it a photo *of* Rex, but not of Chex? Its causal history, to which Rex is connected in an appropriate way but Chex is not. One of the advantages of a causal theory is that it can decide cases where merely matching of the representation to the world is insufficient for something to be a representation *of* that thing in the world.

A hybrid theory that combines a causal condition with a descriptive condition in the right way would seem to combine the advantages of both. I will now list one more area where causal theories of content, considered very broadly, require descriptive elements. For instance, there are independent motivations for an externalist about reference to adopt a hybrid theory of reference.

The *Qua*-problem for the causal theory of reference shows the need for the speaker who first dubs a natural kind member using a certain term (like 'tree') to associate it with some descriptions that will specify the nature of the thing being designated. Otherwise, it is indeterminate which *aspect* of the thing is being designated by the term. For example, the

very first speaker applying the natural kind term ‘tree’ to a sample of its kind needs first to have in mind that the thing *is* a natural kind sample, and then to associate some descriptions with it that will pick out the specific nature of the term being applied. Is ‘tree’ to designate the thing *qua* (in its aspect as) a living thing? Or perhaps it is meant to designate the thing *qua* plant, *qua* birch, or even *qua* individual birch? Plausibly, the speaker will have some descriptions associated with samples of the kind based on observed characteristics, and these are what she will have in mind when applying the term ‘tree’ for the first time. For instance, perhaps the speaker’s implicit semantic intention is that the term applies to a member of a natural kind that is a large plant with a woody, elongated trunk, and a branched head.[11]

Having established that a hybrid theory is compatible with content externalism, and motivated by concerns within the causal theory of reference itself and not merely an *ad hoc* modification, we can now specify what hybrid representation conditions would look like for objects:

H1: For a visual object representation to represent some object *x*, it must be such that:

i.) It bears an appropriate causal-historical relationship to some thing (or things) in the world that it tracks reliably, and

ii.) The properties presented in it sufficiently match the properties of the things in the world that it reliably tracks.

H1 combines the causal condition we first saw in the *a priori* argument with a matching condition. This ensures that if the properties of the object are wildly misrepresented despite reliable tracking, that this will not be classified as a representation of the object.

[11] See Devitt and Sterelny (1999:79-81, 90-93) for a good explication of the qua problem for names and other terms.

However, not *every* property has to match what it reliably tracks for the representation to be *of* the object. This would set the bar for representation too high. Consider that some birds, spiders, and insects can see into the ultraviolet spectrum, representing colors that we cannot. Assuming for the moment that these are real colors that we merely misrepresent as some other color, we would not want to say that we *entirely* misrepresent objects that reflect ultraviolet light. We might very well still represent their shape, size, and location correctly. In these kinds of cases, it makes sense that we should still say that such objects exist, but that one of their properties is consistently visually misrepresented. It is also possible that there are no colors, that some form of color anti-realism is true, and therefore that the great majority of our visual representations reliably misrepresent their objects as bearing this one kind of property. It would seem unjustified to reject our visual beliefs in objects on this basis alone.

Here is the specific veridicality condition for visual representations of properties:

H2: For a visual representation as of a *property* to be veridical, it must:

i) Bear an appropriate causal-historical relationship to some property in the world that it such that it tracks this property reliably, and

ii) Match the property that it reliably tracks.

As with H1, here condition i) establishes the correct causal relationship with the property involved. Some nuance may be possible with ii), but for present purposes either a visual representation matches a property, or it does not. Note that a visual object representation might fulfill H1 but not all the properties represented in such a case would thereby need to fulfill H2.

The general Hybrid Argument for the existence of trees would have a revised premise (2), and would look like the following:

1) I have visual experiences as of trees.
2) My visual experiences as of a tree (here, *an object that is a tree*) are sometimes such that i) they bear an appropriate causal-historical relationship to some thing (or things) in the world that they track reliably, and ii) the properties presented in them sufficiently match the properties of the things in the world that they reliably track. (That is, condition H1 is met.)
3) If (1) and (2) are true, then trees exist.
4) So, trees exist.

Premise (1) can be known through introspection. Premise (2) captures the conditions for H1 here, because my visual tree experiences are visual representations as of objects. Note that because premise (2) now contains the conjunct asserting condition ii) of H1, it is no longer possible to reason *a priori* to a conditional like "If I have visual experiences as of trees, then either trees are real or my tree experiences occur in conditions where the causal condition is not satisfied." Now there is a genuine empirical possibility that our tree experiences might be reliably misrepresenting something else *as* trees, and in such a case condition ii) of premise 2 would not be met. Premise 3 relies on the assumption that, if an object is represented as having the property of *treeness* and *treeness* means having a certain set of properties, then if H1 is met in the case of a representation of an object that is a tree, then any properties essential to *treeness* are represented veridically. Here, we remain neutral about whether *compositeness* is essential to *treeness.*

In the next section, I will explore our first candidate way of spelling out the details of the hybrid strategy for resisting EDO1*.

5. Hybrid Externalist Response #1: Non-Eliminative Nihilism

5.1 The NEN View and Response

One way to resist EDO1* is to accept the nihilist thesis that there are no composite objects, but to deny EDO1* on the grounds that trees just *are* pluralities of simples arranged treewise. The nihilist would agree that the best explanations of our tree beliefs will involve pluralities of simples arranged treewise. If the latter just are trees, then our best explanations of our tree beliefs certainly *will* involve trees. In this section, I will explore this strategy and show why it is limited and ultimately problematic as a response to EDO*.

This take on nihilism is known as *Non-Eliminative Nihilism* (hereafter: NEN), and it is a live view that has been defended by Gabriele Contessa (2014). A mereological view sometimes supported by appeals to content externalism, NEN holds that even though there are no composite objects, the content of our thoughts about *trees* as well as the reference of the term "tree" are partly determined by a causal-historical connection to pluralities of simples arranged treewise. In general, NEN holds that where *K* is a natural kind sortal and "*K*" is a natural kind term, the mental content of our thoughts about *K* and the meaning and reference of "*K"* are partly determined by pluralities of simples arranged *K*-wise if they have the right kind of causal-historical connection to these pluralities.

Here we are taking NEN to be part of a broader strategy that helps us resist EDO1*, and to do this we are taking it to apply to the content of visual representations, not just of concepts and terms. We are also taking it to be a hybrid externalist view of the kind I

developed in Section 4, which endorses the condition H1 for object representations and H2 for property representations. (I make no claim that actual proponents of NEN would accept either of these modifications).

Understanding NEN in this way, we can meet the explanatory challenge of EDO1* by reinterpreting the perceptual and conceptual content of our beliefs about natural kinds like *tree*, whose members are conventionally believed to be composite objects. According to NEN there is no unified macroscopic composite object made up of treewise matter, but our concept *tree* has for its content pluralities of simples arranged tree-wise, so the belief *that there are trees* is true. This offers an easy explanation of how tree facts cause our tree beliefs: because trees (qua pluralities of simples) *exist*, and they cause our visual experiences as of there being trees in our environment.[12]

For NEN, the causal-historical story would run something like the following. Pluralities of simples were arranged treewise in our ancestral past and our visual systems evolved to reliably misrepresent them as composite objects because it helped our ancestors reliably track the matter that was reproductively important to them. This means that the condition H2 was not met in the case of the property of compositeness. There are no composites in the world, so there is nothing that matches the property of compositeness. However, because we accept an externalist, partially causal-historical theory of visual representation, because our tree representations have the right kind of causal-historical connection to these pluralities of simples of simples arranged treewise, and because the properties presented in our tree representations (e.g., greenness, solidity) sufficiently match

[12] The strategy can plausibly be extended to artifacts like bicycles and lamps. See Putnam's "Meaning of Meaning" (1975: 160-62) for a seminal application of the view of artifacts; see Schwartz (1978) and Thomasson (2003) for further articulation and criticism.

the properties of the things in the world that they reliably track, they are representations of trees—and these pluralities of simples arranged treewise just *are* trees. Trees, of course, are reliably misrepresented as composite objects, but the mereological status of trees is not a property essential for a visual representation of a tree to be successful.

5.2 Defending the NEN Response Against Two Bad Objections

In this subsection, I will defend the NEN response against the objection that this ontological error is too serious for our visual representations as of trees to count as veridical, and then against the objection that it *a priori* forbids reliable misrepresentations in an illicit way.

One might object that although the veridicality condition in H1 has some slack in it for some relatively minor cases of properties that might be reliably misrepresented (e.g., color in objects), mistaking a plurality of simples for a unified macroscopic object is too extreme of an error for us to ultimately say that such a visual representation is truly of a tree. My reply is that there is some reason to think such a high degree of error-tolerance is part of, and motivated within, content externalism.

One of the advantages of externalism's causal-historical theory of reference is that it allows us to say that early commenters on things like stars were in fact able to refer to those things and make *some* true statements about them—including claims that they exist or have certain observable properties—despite having sometimes deeply erroneous views about their natures. Plausibly, people were successfully talking about stars for many centuries before modern astrophysics began to make precise discoveries about their nature. This explains, for

instance, how their discussions and investigations paved the way for modern discoveries about stars.[13]

The idea here is that mistakes about one's ontological categories need not lead to reference failure. In the case of stars, people believed various false things about them over the centuries, but the term "star" was tied to actual luminous spheroids of plasma by a causal-historical chain back to the first users of the word who dubbed the phenomena "stars."[14] People in the earlier times held plenty of deeply mistaken beliefs about stars. For instance, some believed stars were powerful conscious beings who could grant favors. But they were correct to say, "Stars exist," and their references to members of the kind *star* using the word "star" were largely successful. At the very least, according to a reasonable causal-historical semantics, the real natural kind *star* is still the best candidate for the reference of their term—despite misunderstandings about its nature that would have made many attempts to specify it by description unsuccessful.

Using the following example, Maegan Fairchild and John Hawthorne (2018) argue that even if a language community has deeply erroneous beliefs about when objects of the natural kind *N* come in and out of existence, it's not necessarily true that claims of the form "There are *N*s" are false in the mouths of their speakers. Imagine a community that, because of biological and cultural contingencies, believes in *uptrees*. An uptree is an object that is essentially upright, made up of matter arranged tree-wise; if this tree-wise arrangement of

[13] A descriptivist about names for natural kinds would hold that "star" refers to things because it stands for a description of, e.g., some set of properties that stars have. If we're mistaken about those properties, our attempts at reference will fail.

[14] We'll assume for our purposes that there is an unbroken etymological descent from the Greek *"astér," "astron"* (*star*) (or some ancestor term) to the English word "star" that would count as a causal-historical connection for a content externalist.

matter falls over, the uptree ceases to exist and a *downtree* springs into existence (this is an essentially horizontal tree-wise arrangement, which happens to be made up of the same matter as some former uptree). "Uptree" here is functioning as a natural kind term. Now suppose there really are trees in the world and no uptrees. For Fairchild and Hawthorne, members of the community would still be speaking truthfully when they say, "There are uptrees," because our semantics should say "uptree" picks out the kind *tree* and this community simply has mistaken views about the essence of trees.

According to Fairchild and Hawthorne, the case of "uptree" is relevantly similar to the case of "stars." If we say statements like "There are uptrees" are false in the mouths of the imagined community, we must also think our ancestors who had incorrect views about stars were speaking falsehoods when they uttered "There are stars." They weren't speaking falsehoods when they uttered "There are stars"; thus, this community isn't speaking falsehoods when its members utter, "There are uptrees." Accordingly, existential beliefs about natural kinds or their members are more secure across modal space than other kinds of claims. In fact, the error involved with "uptrees" seems a less severe error than that for "stars" and the same principle applies in each case, there is a natural kind that is the best candidate for the reference of the term, the kind with the causal history that would best explain the origin and use of the term in question (50-51).[15]

According to NEN, though only mereological simples exist and all apparent phenomena and events at the macroscopic level are caused by these simples acting in

[15] This strategy may hold, with some modifications, in the situation where there really are uptrees and downtrees (and no trees), but where a community comes to believe there are *trees* by biological and cultural contingencies. The natural kinds that are the best candidates for *representata* of "tree" are "uptree" in one situation and "downtree" in another. But in either situation, it seems accurate by this view for one to say, "Trees exist."

concert, terms for composite objects like trees really *do* refer to something: pluralities of simples arranged tree-wise. Therefore, we can still have a true belief when we believe *that trees exist* or *that this here is a tree*. The content of the concept *tree* and the reference of "tree" here connect us by a causal-historical chain to something real—in this case, to *pluralities of simples arranged treewise*. On this view, such pluralities have a history dating back to their dubbing under the name "tree" by some ancestor in our language community. We think of a tree as a single object with parts, but it is really just a collection of tiny, partless objects arranged treewise that we refer to when we use the word "tree," and this has been the case since we started using the word.

In the above cases of uptrees and of simples arranged treewise, content externalism confers the benefit of a certain amount of ontological flexibility. Over time, we may very well have to reassign the referents of some of our natural kind terms to different fundamental ontological categories. Just as content externalism allows us to see early thinkers as achieving successful reference using words like "star," it gives future generations of speakers reason to credit us with successful reference to as-yet-undiscovered kinds that are better candidates for the referents of our terms than the kinds we currently believe in—as long as the newer kinds stand in the right kind of causal-historical relationship to our terms.

The broader point here is that content externalism's ontological flexibility applies to our terms as well as our mental content, including our concepts and visual representations. The burden would be on the opponent of NEN to say why this ontological shift was a bad one *by externalist standards* (or, of course, to say why externalism is false). I will argue below that a simple story about our visual tree representations tracking treewise pluralities

will not meet the needs of either the NEN response, or of the debunker's tracking story, but this will not depend on the degree of ontological error posited by either nihilist view.

The second objection I want to defend NEN against is that, despite the NEN response employing a hybrid causal-historical theory of visual representation, which in Section 4 allowed the externalist to avoid Mendelovici's central criticism of such causal theories, NEN entails *a priori* that some kinds of reliable misrepresentation are forbidden. According to the NEN response, it is impossible to reliably misrepresent pluralities of simples arranged treewise *as* trees because these pluralities just *are* trees. It is possible for some things to be misrepresented as trees, of course. It might turn out, for instance, that H1 is not met in the instance of our tree experiences. Perhaps, upon further investigation, nothing in the world is *green* or *solid* or a *plant*. It is possible that too many of the properties presented in our tree experiences are misrepresentations. This is not a problem for the NEN, because in such a case there are simply no trees. Because the underlying pluralities of simples involved do not have these properties, they are not really arranged treewise, but some other way. (If they *were* arranged treewise, they would be trees.)

However, there is a possible scenario where the denied kind of reliable misrepresentation could occur. Unless content externalism is necessarily true, it is possible that in some worlds trees (if they exist) just are composites, because compositeness is presented in the tree experiences of beings in that world and some sufficiently strict version of content *internalism* is true. This would entail that the criterion for the content of concepts like *tree* is strict resemblance or descriptive matching: the thing represented must match the representation in every important respect in order to qualify as a representatum. However, assuming nihilism is also true in such a world, there might be pluralities of simples arranged

treewise but no trees. In this kind of world, it *is* possible that such pluralities are reliably misrepresented as trees. Since NEN forbids this, and partly because of commitments in its theory of visual representation, one might think that it is guilty of forbidding reliable misrepresentation in the problematic way that Mendelovici warns against.

Of course, the externalist may claim simply that the aforementioned world is impossible, because externalism is an *a priori* necessarily true thesis. This means there is no possible world in which internalism is true. But even assuming for the moment that it is possible. Notice that NEN is simply arguing *a priori* that the usual candidate for what is represented—a composite object that is a tree—does not exist. It follows that the best candidate is whatever has the best causal-historical connection (for the NEN advocate, it is a plurality of simples arranged treewise). Notice that this does not forbid any *empirical* situation from occurring; it is entirely a matter of redefining natural kind sortals for *a priori* reasons. If geometers decided for *a priori* reasons that there were no polygons but only pluralities of line segments joined at the edges in certain ways, that our visual square experiences were just of some line segments joined square-wise, and that it is impossible to misrepresent segments joined square-wise as squares, they would similarly not be forbidding any *empirical* situation. For this reason, I think NEN's result is an unproblematic case of forbidding reliable misrepresentation.

5.3 Two More Serious Objections to The NEN Response

Now I will raise two objections that are more serious for the NEN response. First, assuming this strategy succeeds, it is still a very limited way of resisting the debunking argument. Unless proponents of NEN reinterpret more general terms like “composite object” such that they do *not* refer to things that are essentially singular, unified, and divisible into parts, then

our beliefs in them are still threatened by the debunking argument even if we still get to retain our beliefs in trees. Certainly, consistency with content externalism would not force such a reinterpretation and doing so seems only to further erode any substantive disagreement with the debunker in the first place.[16]

The final and most severe objection to the NEN response is that it shares a general weakness with nihilist explanations of our object beliefs. Individual pluralities of simples are an implausible candidate for what is reliably tracked by means of composite object representations, and making the necessary adjustments to the tracking story within the nihilist explanation of our object beliefs nullifies any advantages for nihilism in terms of overall ontological and ideological parsimony.

I will develop this objection as a general problem for nihilist debunkers in the next section, as it is a challenge to the idea that the property of compositeness is reliably misrepresented in our visual tree experiences. Both the nihilist debunker and the proponent of the NEN response agree about this premise, in contrast with proponents of our final hybrid externalist position: the conservative view.

6. Hybrid Externalist Response #2: Conservatism About Objects

Compared to the NEN response, the conservative view is more straightforward. The picture according to the conservative view is as follows. Our visual tree experiences meet the condition H1, and so the hybrid argument for the existence of trees presented in Section 4 is

[16] Indeed, this is the point of such moves. They are compatibilist responses to the problem of how ordinary, reasonable people can say and believe so many ontological falsehoods. See van Inwagen (1990, 2014), Contessa (2014), and Liggins (2008). In his defense of van Inwagen's view, Liggins endorses interpreting *all* talk of composite objects as involving plural reference. Contra the present view, Liggins thinks the nihilist still has ways of expressing a distinctive position against composite objects (2008: 190-192).

sound, and trees exist. In addition, the property of compositeness is *veridically* represented in our tree representations. Condition H2 is satisfied. Trees are, in fact, composite objects.

Because this is a hybrid view, it does *not* hold that reliable misrepresentations are impossible. Rather, it holds that our best explanation as to why we have visual experiences as of trees as composites is that we reliably—and veridically—represent trees as composite objects. The view depends on an appeal to biological facts as a criterion for composition, and to a particular way of reframing the debate against the nihilist debunker. I will spend the rest of this section developing the conservative view and defending it against objections.

6.1 Reframing the Debate for the Conservative View

The debunker claims that we have an adequate, ontologically minimal story of the evolution of perception. There are simples arranged treewise and dogwise and humanwise, and a long causal history where our simples arranged visionwise were shaped by selective pressures. We track the lower-level stuff competently, such that we predict and respond to it in reproductively advantageous ways. Part of our means for doing so is visually carving matter up into midsize objects. Our visual midsize object representations guide our behavior by focusing our attention on reproductively important things in our environment.

What, if anything, would *facts* about how matter is divided (or combined) into midsized objects add to this story? As Korman puts it, it is hard to see how assuming that our object reactions are accurate—that they match the object facts—*could* add to such a story (2019: 344). But, as I will now show, there is more than one way to tell the story about how compositional facts relate to lower-level facts about how matter is arranged or distributed, and to the evolutionary story of our visual object experiences. It is best to see the debunker's story and the conservative story as two entirely different stories, rather than as one layered on

top of the other. Where the debunker's story begins with the lower-level story and reasons that a composite-object level one need not be added, the conservative story begins with a multi-level explanation and holds that, properly understood, the higher and lower levels do not come apart.

It matters crucially where one begins such an account, and what one assumes at the outset. Lower-level facts about how matter is distributed or arranged play different roles in the conservative story than they do in the debunker's story. Moreover, neither evolutionary biology nor commonsense ontology *begin* with this lower-level story. Perhaps our question should be: beginning with the conventional, composite-objects view, why should we *subtract* midsize composite objects or the idea that our visual object experiences evolved to be accurate about them?

The notion of *arrangement K-wise* will play a slightly different role in the conservative explanation than the debunker's. Since "arrangement *K*-wise" is a philosophical term of art that originated in the needs of mereological nihilists, it expresses the conditions some simples must meet in order to collectively bear the properties that some composite object *K* would bear, if there were composite objects. For the conservative, arrangement *K*-wise just is the set of conditions that simples (or composite parts) must meet in order to compose a *K*. If some simples are arranged *K*-wise (and let's assume for the moment that this means *whole-K*-wise), then they compose a *K*.

Accordingly, where there are simples arranged treewise, there is a tree composed of those simples. And wherever an animal is successfully tracking treewise matter, that animal is also tracking trees. Trees have certain biologically relevant properties: they hold out reproductive benefits or carry reproductive hazards for certain organisms. There are facts

about where trees are, how they grow, what nutrients they provide, and so forth. These are certainly facts about reproductively important arrangements or distributions of matter, but whenever one is tracking these lower-level facts, one is also tracking facts about the composite objects that are trees. So understood, there is no competing lower-level explanation in terms of simples arranged treewise because this level is *part* of the explanation of the composite tree.

Another assumption of the conservative explanation is some kinds of arrangements or distributions of matter are important to any evolutionary explanation of our object beliefs. This assumption should be uncontroversial, because we are leaving aside for the moment whether any particular way of construing these facts is "indispensable" for an evolutionary explanation. The debunker appeals to such important facts, put in nihilist terms: that at the time when our ancestors' visual systems were evolving, there were some simples arranged treewise in the environment; that somewhere in the simples arranged ancestorwise, there must have been simples arranged eyewise and brainwise; and that just how these were arranged depended upon changes in simples arranged DNA-wise within previous ancestorwise arrangements of simples that were causally connected in the right ways.

The conservative explanation assumes that composites exist, and that things arranged treewise compose trees. Therefore, all these things whose arrangement *K*-wise the debunker thinks are important to the evolutionary story are things that compose *K's* that are important to the evolutionary story: there are composite trees, ancestors, eyes, brains, and strands of DNA. In general, let us assume that what is important to the evolutionary story of our object beliefs includes things the debunker would need to include, in a paraphrased version, in her

own story. Ultimately, these will just be the things referred to in conventional evolutionary explanations of perception.

Finally, for the conservative explanation, we will take what is referred to in the explanations of evolutionary biology, broadly speaking, as kinds of composite objects that exist in the world. These are not the *only* ones that exist, but some things meeting the conditions for being a *K* according to the explanations of evolutionary biology is sufficient for them to compose a *K*. The debunker should agree that each biological kind *K* referred to in evolutionary biology represents a way some simples could be arranged *K*-wise. I will refer to these as biological kinds or natural biological kinds for the rest of the chapter. (Where it is not clear from the context, I will specify whether I mean kinds of objects or kinds of arrangement of simples.) The authority of appeals to evolutionary biology should extend to the whole science, not just to what is important for perceptual explanations. Here again, the conservative is just applying the principle that arrangement *K*-wise means composing a *K*.

These, then, are the assumptions made by the conservative account of our visual object experiences. The account itself is the conventional one: there were composite objects like trees and tigers in our evolutionary past, and our ancestors faced selective pressures in favor of accurately tracking and representing the object facts about composites like trees and tigers. Therefore, our visual object experiences evolved to accurately track and represent composite objects like trees and tigers.

In the next section, I will answer whether this explanation is better than the debunker's. I will break the discussion down in terms of a few specific virtues and show that even if the debunker's explanation is no worse than the conservative's, it is no better.

6.2 The Debunker's Explanation vs. the Conservative's Explanation

In this section, I will compare the debunker's explanation with the conservative's composite-friendly explanation. I am going to assume that it *is* possible to explain our object reactions/beliefs without assuming their accuracy, and further that reliably tracking reproductively important matter in an organism's environment *is* compatible—in the broadest possible sense—with certain ways of misrepresenting which midsized objects there are. This is to say that I will assume that the absence of ordinary midsize composite objects from these explanations does not make them gibberish.

My bar for comparison is much lower. I simply want to ask: is the conservative's explanation *worse* than the debunker's? Korman (2019: 343-45) sets up a comparison between the two accounts wherein he argues that the debunker's explanation is better than the conservative's because it is more parsimonious, more illuminating, and clearer. I will address each of these points in turn.

6.3 Parsimony

The idea here is that the debunker's account is more parsimonious because it does not require the existence of ordinary objects. All things being equal, fewer ontological commitments makes a theory more parsimonious; and all things being equal, a more parsimonious theory is a better theory.

The debunker does not have to establish that the conditions for reliable misrepresentation *are* met in the case of composite object representations, but she does have to give reasons why a nihilist-friendly and mismatch-compatible explanation will be better than the conservative's composite-friendly, mismatch-incompatible explanation. To this end,

the debunker must establish plausible ways that the conditions for reliable misrepresentation might be met in a nihilist-friendly manner.

Recall that the conditions for reliable misrepresentation are: (i) mismatch, (ii) false representation, and (iii) reliable tracking. I will now show that the nihilist cannot tell a simple story about reliable misrepresentation that hinges on the claim that particular pluralities of simples are what is tracked reliably via our composite object representations. To tell a robust tracking story that will match the conservative's story in explanatory power, they must develop a less parsimonious explanation.

The debunker assumes that composite object representations are nonveridical because there are no such objects; there are only simples arranged in various ways. This parallels the conservative's assumption that there *are* composite objects and serves a similar role in the debunker's overall explanation. The implicit assumption shared by both parties is that mereological claims are to be evaluated in the context of the overall explanation. Accepting the debunker's assumptions for the sake of argument, condition (ii) is met.

That there is a mismatch between pluralities of simples and familiar objects like trees is also fairly uncontroversial to establish. Trees are unified middle-sized things that can be decomposed into little parts, while pluralities of simples are collections of lots of little partless things. Pluralities of simples arranged treewise act in concert to produce the effects we ascribe to trees, but they can survive the tree's being chopped into pieces or burned to the ground while the tree cannot. We could list many such differences that warrant, with an appeal to Leibniz's Law, that they must be different things. So, condition (i) is met.

The third requirement is that by means of the misrepresentation, *something* is tracked reliably. While this seems plausible at first glance, it proves to be impossible to establish in the case of particular pluralities of simples. As I will now argue, we cannot reliably track pluralities of simples arranged treewise by representing them as trees.

Let's say I see some treewise properties like *leafiness* and *woodiness* instantiated in front of me at time T_1, all of which I visually represent as being borne by a tree. Accordingly, at time T_1 some plurality of simples P_1 bears these treewise properties. However, over time the population of simples that are arranged treewise in a way that we would (mistakenly) associate with *the same tree* is in a great deal of flux due to metabolism, growth, aging, physical damage, and other factors.[17] Let's say that at some later time T_2, the simples of P_1 are somewhat scattered across the landscape, many of them still arranged treewise and occupying the same space as in T_1—but some of them now scattered across the landscape and arranged grasswise or soilwise, for instance. Only *some* of the simples in P_1 are now still arranged treewise and causing visible treewise properties; other simples of P_1 are now causing visible grasswise or soilwise properties or are perhaps not part of any visible arrangement. Meanwhile, some *new* plurality P_2 (whose membership overlaps somewhat with that of P_1) is now causing all and only the macroscopic treewise effects I mistakenly believe to be properties of the *same* tree I had a visual experience of at T_1. At all times after T_1, I mistakenly believe the leafiness and woodiness are properties of *the same tree I saw at* T_1. This means I have the perceptual belief that I have tracked a single tree during the time

[17] One can generalize to organismwise arrangements here.

from T_1 to T_n, but by the debunker's lights it would seem I have tracked neither a particular tree *nor* a particular plurality of simples.

Two primary adaptive advantages of reliable tracking are *re-identification* and *discrimination*. For instance, re-identification of landmarks could mean the difference between successful navigation toward ancestral feeding grounds and getting lost and starving; discriminating between one's own kin and that of reproductive rivals could pay dividends in optimal investment of resources like food and protection. But our tree example above illustrates how representing things as objects hopelessly muddles both re-identification of pluralities of simples and discrimination between them. We are not equipped to track the multitude of simples involved in telling P_1 from P_2. Rather, we will mistake a different plurality for the one we saw earlier because we can only see macroscopic properties instantiated over time. This same limitation means we cannot discriminate between pluralities that pass in and out of a single treewise arrangement over time. This problem generalizes to large classes of things that would have been important for our ancestors to track in their environments: pluralities of simples arranged matewise, kinwise, rivalwise, preywise, treewise, fruitwise—perhaps even rockwise and hillwise.[18] These considerations suggest that if we are reliably tracking something by means of our object representations, it *cannot* be particular pluralities of simples.[19]

The debunker may at this point shift to defending the idea that what is tracked via our visual object representations are *arrangements* of simples *F*-wise, which happen to have the

[18] Rockwise arrangements and other landscape features shift their particle populations as well, but at a slower pace.
[19] See Mendelovici (2013: §4.1) for a discussion of tracking, re-identification, and discrimination.

same properties—including persistence conditions—as an *F*. In Chapter 1 §8 I argue why this is just letting in composite objects through the back door.

More likely, the debunker will embrace a story like the following.[20] My visual tree representations reliably track the fact that there is usually a plurality of treewise simples every time I have a visual tree representation. My *same tree* representations track continuity relations between earlier and later collections of pluralities. Whenever I think some tree is identical to some earlier tree, what I'm really tracking is a chain of continuous minimal changes linking the earlier simples to the later simples. This enables me to track what kinds of pluralities there will be, and in which places, in a few seconds or a few days. This doesn't just mean that if I see some treewise simples in a particular spot today I know to expect *treewise* ones there tomorrow. Rather, I should expect to see *same-tree-wise* simples: those that meet the conditions for treewise arrangement and this other condition, too. This allows me to track landmarks, prey, or predators in the needed way.

However, note that this is an additional relation that must be added to the debunker's explanation for it to have the same explanatory power as the conservative explanation. In Chapter 1, I call this a "schmidentity" relation, and note that it involves some cost in parsimony. Ultimately, though the changes may be specified in low-level terms, they must match the persistence conditions for whatever kind of composite object *K* is the basis of the *K*-wise arrangement in question. (See Chapter 1 §8.)

The debunker's explanation will need to specify several other relations to match the explanatory power of the conservative explanation. It will need a *whole-K*-wise relation, for

[20] Andrew Brenner (2018: 674) explicitly embraces this kind of strategy.

when simples are arranged such that they collectively bear the properties of an entire *K* as opposed to part of a *K*. On the other hand, simples arranged such that they collectively bear properties of only part of a tree (like its trunk) are arranged *partial-tree*-wise. If I point to a branch and say that there's some treewise simples, I'm speaking the truth, but only if we distinguish between *whole*- and *partial-tree-wise*. In addition, the debunker will need a *particular*- or *countable-K*-wise relation to distinguish what we visually represent as individual trees in a whole grove of trees. Without it, if we said that the area is full of simples arranged treewise, one could be forgiven for thinking the entire area was full of simples arranged such that they collectively bear the properties of one *giant* tree.[21]

These are peculiarities of the *K*-wise arrangement paraphrase strategy, and each of these additions involves some new cost in terms of parsimony. But in Chapter 3 §4 I raise a more general parsimony worry for the nihilist. Building on the work of Bennett (2009) and Tallant (2014), I argue that the *arrangement K-wise* paraphrase and all its supporting apparatus are unique ideological commitments; that these aren't duplicated in a conservative's composite-friendly ontology + ideology; and that they represent a considerable ideological cost. I argue that nihilism's ontology + ideology therefore has no net parsimony advantage over a conservative's.

The debunker is, of course, free to adopt another sort of minimal ontology and adopt another kind of paraphrase. She may say, for instance, that we're tracking which regions of space will be filled, and by what *sort* of matter, from one time to the next. But ultimately, she would have to invoke descriptions like "tree-ish matter" and relations like "same-tree-ish

[21] If there were trees, of course. See also Chapter 3 §4.4.4.

matter," "whole-tree-ish matter," etc. She would ultimately seem to be in for the same kinds of parsimony costs as the nihilist.

I will readily grant that the conservative's explanation is not ontologically or ideologically cheap, either. There are composite trees, and the conservative is on the hook for conditions under which these are composed and under which they persist, among many other things. However, I suspect that for every extra ontological cost on the conservative side of the ledger, there is a parallel ideological cost on the nihilist side. Ultimately, the debunker has not established that her nihilist-friendly explanation is *more* parsimonious than the conservative's composite-friendly explanation simply because the latter contains composite objects.[22]

In the next section, I will compare the two explanations of our object beliefs in terms of explanatory power and clarity. I will show that, properly understood, the conservative explanation has clear advantages over the debunker's explanation.

6.4 Explanatory Power and Clarity

In this section, I will briefly recap the debunker's explanatory challenge to the conservative in terms of clarity and explanatory power. I will then show why the debunker's explanation must consider it adaptively important for organisms to get some facts right about the midsize world.

Recall that the debunker provides a positive explanation of why we visually represent the midsize world as containing trees and dogs, and not trogs. It goes as follows. Deep in our

[22] I am deliberately leaving aside the possibility that error theories like mereological nihilism are inherently less ideologically parsimonious. See Mark Johnston (1992a, 1992b, 1993) for an anti-error-theory case, and Chris Daly and David Liggins (2010) for a response.

ancestral past, our simples arranged ancestor-wise tracked various reproductively important arrangements of simples: simples arranged matewise, foodwise, treewise, tigerwise, etc. For our ancestors, visually representing simples arranged treewise as composite objects adaptively guided their behavior but did not track any facts about trees (if there were trees). Representations as of, and perceptual beliefs in, midsize composite objects functioned merely to adaptively highlight and package together chunks of an ontologically neutral landscape of facts about how simples are arranged, in a way that improved reproductive fitness. Because of our ancestors' environment and reproductive needs, it was more adaptive for them to represent things as trees and dogs instead of trogs. For troglodytes, who had a radically different environment and reproductive needs, it was more adaptive to visually represent things as trogs. In this story, the facts about composite objects, if there are any, do not play a role.

According to the debunker, this gives a good, clear, illuminating explanation of why we ended up with the sortals we have: trees and dogs, but not trogs. By contrast, the conservative account is unclear on how tracking tree *facts* would be selectively advantageous for our ancestors. Again, it is hard to see how tree facts could play any role in the above story. Korman (2019) puts the challenge like this. Say we discovered that our distant ancestors were troglodytes, who evolved in the same kind of environment as our current one, but who eventually lost out to reproductive rivals who represented the world as containing trees and dogs. Could we say why their visually representing things as trogs made them *less* fit? Maybe the conservative could appeal to factors like the cognitive inefficiency of tracking trees or dogs by visually representing them as trogs, but notice that this explanation has nothing to do with any *inaccuracy* about composite object facts (344-45).

6.5 The Argument for Composite Accuracy

In this section, I argue for the importance of accuracy about composite object beliefs, even by the nihilist debunker's lights.

Remember that the debunker's question is: even *if* there are midsize composite objects, and just the ordinary ones we believe in (like trees and dogs, but not trogs), how can accurate visual midsize object representations figure into an evolutionary story about why we believe in them? For the purposes of this argument, we are assuming that our ordinary beliefs about midsize composite objects are correct, and that there are trees and dogs but no trogs. I will show how having *inaccurate* beliefs about these can be a reproductive disadvantage.

The argument for Composite Accuracy (CA) runs as follows:

(CA1) Mistaken perceptual beliefs about the biological arrangement of simples in one's environment are, *ceteris paribus*, a reproductive disadvantage.

(CA2) Sometimes mistaken perceptual beliefs about how one's environment is populated with midsize composite objects entail mistaken perceptual beliefs about the biological arrangement of simples in one's environment.

(CA3) So, sometimes mistaken perceptual beliefs about how one's environment is populated with midsize composite objects can be, *ceteris paribus*, a reproductive disadvantage.

In CA1, "the biological arrangement of simples" is the arrangement of simples in one's environment DNA-wise, cellwise, proteinwise, foodwise, woodwise, furwise, dogwise, treewise, matewise, and in general *K*-wise where *K* corresponds to a natural kind sortal referenced in the explanations of evolutionary biology. These include, but are not limited to,

organisms and their parts. The idea behind CA1 is that facts about the biological arrangement of simples in one's environment are important for any midsize organism to track, because they are likely to hold out reproductive benefits and hazards. When some simples are arranged treewise, they have a causal-historical relationship with ancestor groups of simples that were also so arranged. And this causal history means that simples arranged treewise move and grow in certain ways, bear certain nutrients in their simples arranged leafwise and fruitwise, and offer safety in their simples arranged branchwise. Simples arranged wolfwise (before we bred them into dogwise arrangement) shared a causal history that gave them vastly different properties, including fast movement and vicious simples arranged teethwise. It was crucial for our ancestors to perceptually distinguish these kinds of biological arrangements quickly and accurately. Indeed, the ability to track facts about the biological arrangement of simples in one's environment is central to the debunker's explanation of our object beliefs.

According to the debunker's explanation we also *represent* how simples are arranged on the midsize level, and these representations are separable at least in principle from our composite object representations. After all, we can agree with another species about at least some of the facts about the arrangement of simples but disagree about how the world is carved up into midsize objects. For instance, when a dog and a tree are around, a troglodyte represents that there is some stuff arranged dogwise that is furry and moves around in a certain way, and that there is some other stuff arranged trunkwise that is woody and generally stays put. We would agree with the troglodyte about this, while disagreeing with his belief that the stuff in the scattered region containing both the furry and the woody stuff composes a trog.

We and the troglodyte are *correct* that there is some stuff arranged woodwise and furwise out there. After all, the nihilist-friendly version of evolutionary biology backs us up here. There are simples arranged woodwise and furwise because of long chains of causal interactions in the earth's history; one might say these are *natural* ways for simples to be arranged, because they correspond to properties of natural kinds from evolutionary biology (wolves and their descendants have fur, trees have woody trunks). Moreover, it would have been reproductively advantageous for us to evolve to track these facts accurately; presumably we could tell a corresponding story about troglodytes.

It is also possible to be *wrong* about how simples are arranged—to misrepresent the furry patches as leafy, for instance, or not predict their movements correctly. Presumably many lineages went extinct because they evolved visual faculties that inaccurately represented the arrangement of simples. As with any trait subject to natural selection, a single reproductive adaptation can involve a cost in one area that is be outweighed by a benefit in another area. It is possible that a less accurate means of perception might be very energy-cheap and so still a net reproductive benefit, or that members of a species have such highly unusual reproductive needs that a certain kind of systematic inaccuracy about the biological arrangement of simples in their environment was adaptive, all things considered. But *ceteris paribus*, more accurate perception of the biological arrangement of simples is a reproductive advantage, less accurate perception is a reproductive disadvantage.

One might reasonably ask whether these are not really *separate* representations of how simples are arranged, but just aspects or interpretations of composite object representations. If they are aspects of composite object representations, then they are properties (or combinations of properties) presented in a visual representation that represents

(or misrepresents) some things collectively as a composite object. However, *furriness* might be veridically represented in a visual representation as of a dog, even if there are no dogs—just as it might be in a visual representation as of a trog, even if there are no trogs. But in either case, to someone who has either the dog or trog visual representation, it seems like there's some furry dog-shaped stuff—some stuff arranged furwise and dogwise (or partial-trog-wise)—in their field of vision.

As to our representations of the arrangement of simples being an *interpretation* of our visual composite object representations, it is true in at least one sense. It is certainly not the case that our visual tree representations present the properties of trees as being borne collectively by *partless* objects. But when you're confronted with a visual experience of a nearby tree falling in your direction, whether you interpret this representation as being of a composite object, or of a plurality of simples—you don't need an interpretation to know that there is some tree-shaped stuff moving toward you fast, and you'd better run for cover.

This brings us to the motivation behind CA2. The debunker paints the following picture of visual representation. Mistakes about how the world is divided into midsize objects are self-contained, and do not entail mistakes about how simples in the world are arranged. Therefore, two beings can have identical representations of the way simples are arranged but differ significantly in their representations of how the world is carved up into midsize objects. Something like this is supposed to be the case with humans and trogs. When looking at how the world is carved into midsize objects, moving from our visual representations to that of troglodytes is akin to taking the same detailed, accurate topographical map of the world complete with forests, plains, and rivers, and simply redrawing the political boundaries with a marker.

But this neat separation is implausible. Our representations of how things are arranged in our visual field is implicit in our visual midsize composite object representations. If I am having the experience of seeing a blue frog, I am having the experience of seeing something blue. Let's say I am having a visual tree experience and that this means I am having the experience of a solid, so-shaped, green and brown, unified composite object in front of me, then I am also having an experience of some thing—or some *stuff*—that is so-shaped and green and brown in front of me. That I am having a visual experience as of a tree entails that I am having a visual experience of some thing or some stuff arranged treewise, however one fills in the details about what it means perceptually for something to appear treewise.

Nihilists appeal to this sort of idea regularly. Indeed, the idea that people are to be understood as really meaning that they see some simples arranged treewise when they claim that they see a tree (or at least when they do so outside the ontology room), or that the proposition about the simples is a nearby truth that renders the proposition about the tree nearly as good as true, are proposed nihilist solutions to the so-called "Problem of Reasonableness," the idea being that it is a bad thing that nihilism gives the result that almost everyone has false ontological beliefs.[23] The underlying assumption is that anyone would agree, trivially, to the notion that there's some treewise stuff over there if they already believe there's a tree over there. The former is just a weaker version of the perceptual belief they're already expressing when they say, "I see a tree over there." Some nihilists even argue that when one claims that some things compose a *K*, one is also committed to the claim that

[23] See Chapter 1 §4 for more discussion of the "Problem of Reasonableness" and proposed eliminativist solutions. The term comes from Korman (2009).

these things are "arranged *K*-wise" in whatever full, technical sense of the term nihilists intend.[24] (For the record, I am arguing for a far weaker claim.)

Let us return to our central question: does it matter in terms of reproductive fitness whether our ancestors evolved to accurately track facts about midsize composite objects? I have shown that visual representations of composites like trees are also representations of some thing(s), some stuff, or perhaps some so-shaped region of the world being arranged treewise. And I have also shown that, even for the debunker's explanation, it matters whether an organism gets it right about how things in the natural world are arranged. It follows that, at least in cases dealing with biological kinds (kinds of *composite* for the conservative, kinds of *arrangement of simples* for the debunker), getting things wrong about composite object facts means getting things wrong about how things are arranged, and getting things wrong about how things are arranged is, *ceteris paribus,* a reproductive disadvantage.

Let us take another look at our example of the troglodytes. Remember that in this example there are actually trees and dogs but no trogs. More germane to the debunker's explanation, simples are arranged treewise and dogwise, because these are arrangements corresponding to natural biological kinds. But nothing is arranged *trogwise*, at least according to the explanations of evolutionary biology. There is no biological property of *trogwise* arrangement; rather, the arrangement *trogwise* crosscuts the biological arrangements *treewise* and *dogwise.* However, to visually represent stuff as being arranged trogwise is to represent the stuff arranged trogwise as having some fundamental, natural unity to it just the way we would represent treewise or dogwise simples as having some natural unity to them

[24] See Brenner (2015: 1310-12) for one such nihilist argument. See Bennett (2009: 60-63) for an opposing view.

that warrants our believing that they are so arranged, and that gives them the kind of important natural properties that feature into our biological explanations. In fact, troglodytes visually experience there being a *more* fundamental connection between dogwise and trunkwise matter than there is between trunkwise and *branchwise* matter. (They may track the branchwise matter under a different sortal, but they do not visually represent them as parts of *trees*.)

But troglodytes are wrong. Treewise and branchwise matter *are* naturally unified by a long causal history that features into biological explanations. Trunkwise and dogwise matter are not so unified. In fact, they have vastly different causal histories that explain the vast differences in their current properties—differences that are *pro tanto* reasons that it would be inherently difficult to visually track them as some kind of unified whole, and inherently confounding to understand them as parts of a single natural kind.

It is possible that, *despite* the inherent difficulties posed by the biological inaccuracy of these representations, troglodytes or creatures like them may evolve in very different environments than ours or have bodies with very different reproductive needs than ours, such that visually representing the midsize world as containing trogs is an overall good adaptive tradeoff. Nothing I have said here contradicts this possibility. However, there are biological reasons that, *ceteris paribus*, in this world where there are trees and dogs and no trogs, it is more reproductively advantageous to visually represent the midsize world as containing trees and dogs. Put another way: *ceteris paribus,* visually misrepresenting the world of midsize objects carries a reproductive *price* for an organism in competition with its rivals with more accurate visual object representations. In a broader explanation of that organism's evolution, it may be that this price is paid by reproductive advantages elsewhere. My point is that once

we have identified such a price for an organism, an explanation is needed for how that price was paid. All of this supports the idea that the accuracy of our visual midsize composite object representations really *is* relevant to the evolutionary explanation of why we have the object beliefs we do.

In this section, I have argued that getting it wrong about composite objects can have a reproductive cost inasmuch as it misleads us about the biological arrangements of simples in our environment. In the next section, I will state the conservative explanation of our object beliefs and answer Korman's question.

6.6 The Conservative Explanation in a Nutshell

The conservative gives a simple explanation of why beings with lives and environments like ours would wind up with the sortal concepts we do: because the world is full of trees and dogs, and not trogs. Our recent ancestors flourished because they tracked composite objects accurately—particularly objects like trees and dogs that are natural kinds featuring in biological explanations of human evolution.

The conservative's explanation depends on the assumption that simples arranged treewise are just simples that meet the conditions for composing a tree, and in a world where composites exist, they just *do* compose a tree. I have shown why, in the domain of natural biological kinds, inaccurate perceptual beliefs about midsize composites in one's environment entail inaccurate perceptual beliefs about the biological arrangement of simples in one's environment. And I have shown why, even in the debunker's account, this kind of inaccuracy is, *ceteris paribus,* a reproductive disadvantage.

Can the conservative account explain why, if we discovered that our ancestors were troglodytes, the inaccuracy of their visual trog representations caused them to be less reproductively fit than their rivals who represented things as trees and dogs? In the previous section, I explained why creatures like troglodytes face a *pro tanto* reproductive disadvantage because of their misrepresentations of natural biological kinds *trees* and *dogs*. Korman agrees that a good explanation could be that troglodytes' way of tracking trees and dogs is simply less cognitively efficient than tracking by the sortals *tree* and *dog*, but he notes that this by itself has nothing to do with inaccuracy about the object facts (2019: 344-45). I hope to have shown that such cognitive inefficiency is a consequence of making mistakes about how reproductively important matter is arranged or distributed in the world, and that this kind of mistake can follow from visually misrepresenting what kind of midsize composite objects there are. Put positively, the conservative about objects can explain that our more recent ancestors were fitter than troglodytes because they no longer had to pay the reproductive costs that come with inaccurate perceptual beliefs about the biological arrangement of simples in one's environment.

7. Conclusion: Benefits of the Conservative Explanation

If I have argued well, I have shown that a hybrid externalist view like the conservative view examined in Section 6 can give an explanation of why we believe in ordinary midsize composite objects that is *at least as good* as the nihilist debunker's. It does not vindicate *all* of our objects beliefs, nor does it rule out the possibility of evolved object-misrepresenters like troglodytes. It does, however, give a coherent explanation of the way the accuracy of our midsize composite object representations can feature into explanations of our object beliefs.

If this debate is a tie, perhaps practical concerns like epistemic conservatism or the simple ease of use of thinking and writing in terms of composite objects will win the day.

I will conclude by noting two additional benefits of the conservative view. First, one of the benefits of the debunker's explanation is that provides *some* explanation of what sortals we have in terms of our particular evolutionary situation. It's because of our needs as midsize creatures in a certain environment that we sort things into *trees* and *dogs,* but never *trogs.* The conservative explanation, too, is sensitive to the needs and evolutionary situation of *homo sapiens.* We need to accurately represent things like *trees* because they exist and cause things in the world, and we gain reproductive advantages if we get the facts about them right. But as midsize creatures, we do not *visually* represent things like DNA, microbes, or protein molecules because we did not evolve to behaviorally interact with those things directly, though they be natural kinds and part of biological explanations.

Second, the conservative explanation does not attribute too much to the visual faculties alone. In the debunker's explanation, composite object beliefs do not track any facts about the world but have a strictly attention-focusing and behavior-guiding role. This may be attributing too much to perception. It is possible that perception generally has the function of representing an organism's environment veridically, within the constraints of things like available variation and overall reproductive cost to the organism. Perhaps focus and attention, and behaving in a reproductively advantageous manner, are managed by evolved mechanisms that are *distinct* from visual representation. As evidence that other mechanisms are in charge of this, we treat certain midsize objects very different from others. Potential sexual partners, potential food, and potential dangers are given special emphasis by responses from various areas in the nervous system that aren't central to visual perception. Other kinds

of objects we represent in the same kind of visual detail, but regard indifferently. Perhaps we simply perceive all midsize objects with a certain level of accuracy, and the rest of our brain helps us focus behavior on the most important objects.

Chapter 3:

An Indispensability Argument for Composite Objects

1. Introduction

Indispensability arguments are common in the philosophy of mathematics, where they are used to justify beliefs in mathematical entities such as numbers and sets on the grounds that they are indispensable to our best scientific theories. The target of these arguments is philosophers with a naturalistic attitude that translates into scientific realism about entities. The thrust of the arguments can be illustrated by an analogy: if you allow theoretical entities like electrons into your ontology because of their indispensability for science, you should allow numbers and sets in, too.[1] Some defenders of composition in metaphysics have begun to appeal, at least implicitly, to the important role composites play in our best scientific theories, and proponents of sparser ontologies have begun to field responses.[2]

My goal in this chapter is a modest one: to put forth a definite standard for indispensability and see whether composites are dispensable when compared to the leading competitor that does without them. To this end, I will respond to a line of recent arguments beginning with Rosen and Dorr (2002), Sider (2013), and Brenner (2018) that nihilism is compatible with our best science and even offers some advantages like combined ontological

[1] See Mark Colyvan (2001, 2019) for excellent summaries of the dialectic around these arguments, and for some attempts to strengthen the argument. See Quine (1976, 1980a, 1980b, 1981a, 1981b) and Putnam (1979a, 1979b) for the original sources of what Colyvan (2019: §1) calls the "Quine-Putnam Indispensability Argument." See Hartry Field (1980) for a major objection to the indispensability of mathematical entities.

[2] Those arguing science can inform us about physical composition in some capacity include Schaffer (2007); Morganti (2009, 2013, Chap. 5); Calosi et al. (2011); Calosi (2014); Graziani and Calosi (2014); Calosi and Morganti (Forthcoming); Gillett (2013); Healey (2013), Healey and Uffink (2013), Tallant (2014), and Hofweber (2016). For those issuing a pro-nihilist reply, see Rosen and Dorr (2002), Sider (2013), and Brenner (2018).

and ideological parsimony and the resolution of some theoretical puzzles. I will argue that composite objects are indispensable for our best scientific theories, and that therefore there is no reason to prefer sparser ontologies for scientific reasons alone.

Here is a plan of the chapter. In Section 2, I will state and motivate an indispensability argument for composites. Section 3 addresses some worries about scientific realism that might prevent one from accepting the first premise. In Section 3, I define a nihilist position that accepts scientific realism, but challenges the second premise about the indispensability of composites; I consider Brenner's (2018) account of the putative advantages of nihilism to evolutionary biology; and finally, I compare conventional evolutionary theory with a nihilist-friendly version in terms of theoretical virtues, particularly ontological and ideological parsimony.

2. The Argument for Composite Indispensability

In this section, I will present an indispensability argument for composite objects, and sketch the dialectic around the argument. Here is the argument for Composite Indispensability (CI):

(CI1) We ought to have ontological commitment to all the entities that are indispensable to our best scientific theories.

(CI2) Some composite objects are indispensable to our best scientific theories.

(CI3) So, we ought to have ontological commitment to some composite objects.

CI1 is motivated by a commitment to scientific realism. Here I will take scientific realism to be the view that the statements of our best scientific theories are to be taken

literally as sources of knowledge about a mind-independent world.[3] It follows from this broad commitment to scientific realism that we should be ontologically committed to the entities mentioned or predicted by our best scientific theories—if not to *all* of these entities, at least to the *indispensable* ones.[4]

Much rides on the definition of "indispensability" here. Mark Colyvan (2001) sets up the issue in the following useful way. He notes two alternatives to the above characterization of dispensability that are unacceptable. First, one could claim that *every* entity is dispensable. According to Colyvan, the results of Craig's Theorem entail that for *any* theoretical entity, an empirically equivalent version of the theory in which it formerly featured can be given with that entity eliminated. These are rearrangements of our theoretical vocabulary that do not answer: why eliminate *this* particular entity? Alternately, we might say that *no* entity is dispensable to a particular theory, because removing it would result in a *different* theory.[5] It is easy to show why neither of these is an adequate characterization of dispensability by imagining some theory *T* and adding the junk statement "*e* exists," where *e* is an entity that does no theoretical work whatsoever in *T* and is mentioned nowhere else in *T*. The resulting

[3] See Chakravartty (2017) for a helpful discussion of how the metaphysical, semantic, and epistemic senses of scientific realist commitment interrelate. Since the question here is about what entities we should believe in, I have given an epistemic formulation.

[4] In indispensability arguments for mathematical entities, *confirmational holism* was once thought to provide important motivation for CI1 hand in hand with *naturalism*. Here I take *naturalism* to be the idea that our account of reality should be consistent with and illuminated by our best science. Among other things, this means that science partially determines what kinds of things we believe in. However, this does not tell us which parts of our best scientific theories we should believe in. *Confirmational holism* is the idea that scientific theories are confirmed or confuted as wholes, rather than separately in terms of individual assumptions or hypotheses. According to confirmational holism, when a theory is empirically confirmed every element of that theory is confirmed. For physical theories, this would seem to include a great many mathematical entities. However, criticisms from Penelope Maddy (1992, 1995) and Elliott Sober (1993) have led many proponents to reject or at least de-emphasize this commitment. See Mark Colyvan (2019: §3; 2001: Chs. 1.2 and 2.5) on the role of confirmational holism.

[5] I take Hofweber (2016: 196-200) to be arguing something of this latter kind. His focus is that the content of our theory, and its predictions, are in terms of ordinary objects, not simples arranged object-wise.

theory is *T'.* We want a characterization of dispensability that explains why *e* in particular is dispensable, and neither of the above characterization does so (76-77).

This leads Colyvan to the following definition of *dispensability*, which I will be adopting in the present chapter:

> "An entity is *dispensable* to a theory iff the following two conditions hold:
>
> (1) There exists a modification of the theory in question resulting in a second theory with exactly the same observational consequences as the first, in which the entity in question is neither mentioned nor predicted.
>
> (2) The second theory must be preferable to the first" (Colyvan 77).

The motivation for (1) is fairly intuitive. The theory in question is one of our best scientific theories; its own observational consequences are key to its being highly successful. By offering an alternative theory without the entity in question, the dispenser is not trying to improve on those observational consequences, but to *preserve* them. The dispenser is trying to provide an entity-free theory that is *empirically equivalent*, where this means the dispenser's theory has the exact same observational consequences. So, if any modification of the theory that neither mentions nor predicts the entity in question is, as a result, not empirically equivalent to the original theory, then the entity does important theoretical and predictive work and is thus indispensable on these grounds.

The motivation for (2) is as follows. Once the dispenser has established that the science can be done without the entity in question—that the entity-free theory has the same observational consequences as the original—then the question of dispensability becomes the question: which of the two competing theories *is* our best theory? Which of them is the most

attractive? Once we have accepted scientific realism, that is the theory we should believe and the entities that cannot be given up without making that theory a *worse* theory are the ones that are indispensable to it.

How we determine which is the best theory turns on how we interpret "preferable" in (2). It is meant as a catch-all for how our theory rates according to theoretical virtues like explanatory power, parsimony, boldness/fruitfulness, conservativeness, modesty, and formal elegance.[6] The choice of virtues and their relative weights are bound to be controversial, but ultimately some criteria are needed for choosing between empirically equivalent theories.

In the above example of the empirically equivalent theories *T* and *T',* we can simply appeal to a theoretical virtue like ontological parsimony to rank the two theories. For instance, we can say that *e* does not feature in our *best* theory, *T,* but only into the less parsimonious theory *T'.* Removing *e* from *T'* makes it a more attractive theory, *T.* Therefore, *e* is dispensable to our best scientific theories. So, scientific realism does not compel us to believe in *e.*

The reasoning behind CI2 is that any composite-free version of one of our best scientific theories will fail at (1) or (2), or both. I will assume for the present argument that nihilism is the strongest competitor to conservative ontologies that are committed to ordinary midsize composite objects as well as to the theoretical composite objects mentioned or predicted by our best scientific theories. I will assume, for instance, that the problems faced by nihilism will extend in some degree to other sparse ontologies that let in only certain

[6] See Colyvan (2001: Ch. 4.3) and Quine and Ullian (1978: Ch. 6) for accounts of theoretical virtues.

classes of composites, or to even sparser ontologies than nihilism, and that if nihilist arguments fail here, so will these others.[7]

More specifically, I will argue that *even if* the nihilist's composite-free version of one of our best scientific theories meets (1), having the same observational consequences as the original, the nihilist-friendly theory will *not* meet (2): it will not be preferable to the original. To this end, I will assume that a nihilist-friendly dispensing theory is capable of meeting (1). The nihilist-friendly dispensing theory must be achieved by achieved by paraphrasing any statements of the theory mentioning or predicting a composite object *F* into statements mentioning or predicting simples arranged *F*-wise. I will assume that such a nihilist-friendly theory has (or can have with suitable modifications) the resources to express what is needed to have the same observational consequences as the original theory (though I will note in places any extra ideological cost this may incur).

Once we have accepted that the science can be done without composite objects, the debate is over whether the resulting nihilist-friendly theory is more attractive than the composite-laden theory. Here the burden is on the dispenser to show that it *is.* In Section 3, I will make the case for the indispensability of composites negatively, by refuting nihilist claims that a composite-free version of one of our best scientific theories will be preferable to the original. Taking the example of evolutionary biology, I will show that, all things considered, and in light of the relevant theoretical virtues, it is at best a tie between the

[7] See, e.g., Merricks (2001) or van Inwagen (1990) for eliminativist views that let in only conscious beings or only living things, respectively. See Ladyman and Ross (2007) for an extended argument in favor of ontic structural realism, which denies the existence of even mereological simples.

composite-free and composite-laden theory. In my conclusion, I will consider some tiebreakers that may well apply to the case of scientific theories.

3. Challenges to the Indispensability of Composite Objects

3.1 Defining a Nihilist Position Regarding CI1 and CI2

The nihilist is, of course, not forced to accept CI1. Metaphysical motivations may be more than sufficient for the nihilist to motivate her position and having an ontology that fits well with our best science might be a nice bonus, but hardly decisive. However, the fact that Rosen and Dorr (2002), Brenner (2018), and Sider (2013) have given serious counterarguments against the idea that nihilism is incompatible with our best scientific theories is evidence that even if they do not accept CI1, they accept the challenge to show that nihilism and scientific theories are far more compatible than their opponents think.

In what follows, I will assume that my audience, including my nihilist opponents, accept the kind of scientific realism that entails CI1. I will examine a nihilist argument against CI2 that invokes a standard nihilist paraphrase strategy. The claim against CI2 is that all references to composites can be paraphrased away without sacrificing empirical equivalence and with net gains in theoretical virtue.[8] More precisely, the claim is the following. Let *E* stand for the set of propositions making up evolutionary biology. The nihilist dispensability claim then would be that the theory resulting from nihilist paraphrase

[8] In Chapter 1, I gave reasons to believe that a nihilist paraphrase of evolutionary biology would be problematic, in large part because the nihilist needed the statements of the higher-level theory of evolutionary biology to be nearly-as-good-as-true in order to run a debunking argument against ordinary objects. Here the dialectic is slightly different, and the nihilist is free to endorse a strong anti-realism about claims referencing or predicting composites.

of *E* will neither mention nor predict composite entities; it will be empirically equivalent to *E* (fulfilling (1)); and it will be preferable to *E* (fulfilling (2)).

3.2 A Nihilist Dispensability Argument for Composite Objects in Evolutionary Biology

Brenner (2018: §4) offers a sketch of a dispensability argument for composites in evolutionary biology. He argues that there are no reasons to believe evolution requires quantification over composites, and that there are even theoretical benefits in the form of problems that are eliminated once nihilism is adopted. In this section, I will recap Brenner's argument.

Brenner begins by noting that some philosophers of science have treated not just organisms but large groups such as entire species as mereological sums on the grounds that natural selection must act on spatiotemporal entities that are cohesive wholes (e.g., Hull 1980). However, according to Brenner the nihilist can account for group selection without committing to mereological sums; she can say that natural selection selects not for groups whose group-level traits confer reproductive fitness, but for traits collectively possessed by simples ((arranged organism-wise) arranged group-wise). But the nihilist can accommodate the debate over levels of selection by asking which traits are selected by natural selection: those collectively possessed by simples arranged gene-wise, or simples arranged organism-wise, or by simples arranged organism-wise arranged group-wise, etc. (672-73).

Brenner notes that the paraphrase apparatus for the nihilist will be complex and will need to be spelled out in greater detail. He singles out the need for a same-organism relation to capture statements about genes, organisms, or groups over time:

> "Simples at t1 and some other simples at t2 will be associated with the same (illusory) organism in virtue of there being particular causal and spatio-temporal relations between the former and the latter simples, of the sort which are associated with the life of a single organism" (674a).

This strategy can be expanded to encompass same-genus, same-species, and other relevant properties borne by pluralities of simples even over vast stretches of time.

Brenner offers as a test case a strategy for reformulating the criteria for natural selection as put forth by Richard Lewontin (1970), according to which the following must be satisfied:

1. "Different individuals in a population have different morphologies, physiologies, and behaviors (phenotypic variation).
2. Different phenotypes have different rates of survival and reproduction in different environments (differential fitness).
3. There is a correlation between parents and offspring in the contribution of each to future generations (fitness is heritable)" (Lewontin 1970, p.1)

For Brenner, as long as the terms "individual," "survival," "reproduction," "parent," and "offspring" are understood in a suitably ontologically neutral manner, then nihilism is compatible with natural selection. For instance, "individual" must be taken to mean "simples arranged individual organism-wise." Here I take Brenner to mean not that this is what people *really* mean by these terms, but that there are at least paraphrases available under which the conditions for natural selection are intelligible under a nihilist ontology (674).

Finally, according to Brenner there are several beneficial consequences of a nihilist-friendly view of evolutionary theory. First, reproduction is completely redefined such that

nothing new is ever created: simples are just rearranged. This solves the problems of Reproduction vs. Growth, because if no new individuals are ever produced, there is no time when production of new biological material is the production of a new individual. Likewise, it solves the problem of the Reproduction of Collective Entities, because there is never a question of whether a collection of individual organisms produced a set of new individual offspring or whether the reproductive group as a whole has produced new offspring in the form of a new group. Second, it solves problems of ontic vagueness. For instance, it is a vague matter whether an organism is included under a species, because species grade into one another continuously via a long chain of reproductive ancestry. However, the problem is widespread: among others, the concepts of *organism* and *gene* are also vague. A nihilist-friendly conception of evolutionary biology has a clear answer: there are no species, organisms, or genes. There are only simples collectively arranged in various ways, and simples themselves are not susceptible to the same vagueness worries (675-76).[9]

To sum up Brenner's position, if we don't take it for granted that composites exist, we can recognize that simples arranged organismwise (or specieswise, genewise, etc.) are spatiotemporally located and in general can collectively bear the properties required by natural selection. Moreover, the elimination of composites from evolutionary biology can solve some enduring theoretical problems. In the next two sections, I will show that the nihilist-friendly version of evolutionary biology is not preferable on the grounds of the virtues involved in scientific theory choice.

[9] Though see Katherine Hawley (2004) for an argument that mereological simples do not escape vagueness problems.

3.3 Nihilist-Friendly Evolutionary Biology and Theoretical Virtues

In this section, I will evaluate the nihilist-friendly version of evolutionary biology in comparison with *E* according to some theoretical virtues. I will begin with the areas where it is fairly obvious that the two theories are equally virtuous and proceed to those that are more complicated. I will assume here that there are no insurmountable technical problems with the *F*-wise paraphrases in use by the nihilist.[10]

First, because the nihilist-friendly theory makes a claim about some simples arranged *F*-wise everywhere *E* makes a claim about an *F*, the two theories have the same observational consequences. That is, *E* and the nihilist-friendly theory are *empirically equivalent.* Because *E* is a highly successful theory at predicting a great variety of phenomena in the natural world, so is the nihilist-friendly theory. Neither has the advantage here. Because the two theories are empirically equivalent, certain other virtues such as *modesty*, the property of not making any more empirical claims than is necessary, and *refutability*, the property of making empirical predictions that could be shown false through testing, will also be possessed by both theories in equal measure. Finally, it follows from empirical equivalence that both theories equally possess *boldness* or *fruitfulness*—the extent to which their empirical predictions are able to successfully guide future research programs.[11,12]

Next, I will consider how the two theories compare in *explanatory power*. This can be tricky to characterize, but I will describe it simply as broadness of applicability (sometimes separately referred to as the distinct virtue of *generality*) along with the quality of not leaving

[10] Though see Uzquiano (2004) for potential worries.
[11] This is setting aside potential practical problems in working with the nihilist-friendly version of the theory and its formidably long nested predicates. See the section on parsimony below.
[12] See Colyvan (2001: Ch. 4.3) and Quine and Ullian (1978: Ch. 6) for more on theoretical virtues.

large areas or aspects of its subject matter a mystery. The two theories are equally general because any proposition of *E* that applies to all *Fs* will correspond to a proposition of the nihilist-friendly theory that applies to all pluralities of simples arranged *F*-wise.[13]

The two theories may, however, differ somewhat in how much of their subject matter is left a mystery, because the differences in their ontologies may be important. Recall that Brenner (2018) claims that the nihilist-friendly theory resolves some outstanding problems in *E* that involve ontic vagueness. It is a vague matter, for instance, as to whether an organism belongs to species *G* or to one of its immediate descendant species, species *H.* Starting with organisms that are definitely *G*s and working our way down the line of descent to those who are definitely *H*s, no matter which parent-offspring pairs we choose along the way we will never find a clear case of a parent who is a *G* giving birth to a child who is an *H.* Species are vague, and admit of borderline cases. Nihilism solves the problem by denying that there are species at all. To the extent that this is an outstanding theoretical problem, or to the extent that we want to rid our theory of vague entities, this is an explanatory advantage of nihilism.

However, there are two problems for the nihilist. The first is that some kinds of vagueness are actually interesting scientific results, and we don't want to paper over them with systematic changes to our ontology that are unmotivated by the science. It was a *distinctive* result of Darwin's theory of evolution by natural selection that species graded into each other and had no clear boundaries. This was something that its main theoretical rival, creationism, would never have predicted. Similarly, it may be fruitful to keep trying to solve

[13] This is in contrast to the theory E_{Lite} in Chapter 1, which consisted of highly particularized laws that lacked generality by completely avoiding reference to the putative composite entities of *E.* Here the nihilist has chosen to recoup generality by appealing to "arrangement *F*-wise predicates," and will be subject to some of the liabilities I outlined in Chapter 1 §8. I will elaborate on these further below.

puzzles like determining the boundaries between reproduction and growth. These are core biological concepts, and it might help us greatly to clarify their relationship. Thus, it isn't clear that the putative advantage that nihilism offers here really *is* an advantage.

This brings us to the second problem: the nihilist does not really achieve this advantage, because it does not remove vagueness for evolutionary biology, but merely pushes that vagueness into its predicates.[14] Given the predicate "arranged species *G*-wise" and the predicate corresponding to some descendant species, "arranged species *H*-wise," consider the following scenario: if we were following the line of descent of groupings of simples ((arranged organismwise) arranged species *G*-wise) to those ((arranged organismwise) arranged species *H*-wise), we would be hard pressed to identify any point along that line of descent when it would be appropriate to classify some simples arranged organismwise as ((arranged organismwise) arranged *G*-wise) and when we should start saying that they are ((arranged organismwise) arranged *H*-wise). The arranged species *G*-wise and arranged species *H*-wise predicates are vague.

Thus, any theoretical problems introduced by ontic vagueness in genes, organisms, species, or the like will carry over into vagueness of genewise, organismwise, and specieswise predicates. Another example: the nihilist-friendly theory will not always be able to say, whenever some simples are arranged organismwise (and are part of some same-organismwise succession of simples) and are growing in number, at exactly what point some of the simples stop bearing the *current* same-organismwise relation to the other simples in the succession and when they begin to bear a *different* same-organism relation, and only to

[14] See Karen Bennett (2009: 65-71) on the way a whole host of metaphysical problems re-arise for the nihilist in terms of "arranged *F*-wise" predicates.

subsequent groupings of simples (this, of course, is how the problem of growth versus reproduction persists for the nihilist).

The nihilist might object that vagueness in one's predicates is less objectionable than ontic vagueness.[15] Presumably, then, it is better to trade the latter for the former. So, we have at least moved some important problems from evolutionary biology into an area where they cause less damage. The problem is that, even if it ends up really being the case that we should get rid of ontic vagueness wherever possible, it is unclear how this is a *theoretical* advantage for nihilist-friendly evolutionary biology and not merely something that puts the theory in line with the demands of metaphysics.

Predicates like "arranged specieswise" or "arranged parentwise" will fulfill exactly the same predictive and explanatory role as "species" or "parent," so it is no surprise that problems with the latter will persist in the former. Concerns about the vagueness involved in these particular theoretical concepts—and the puzzles they produce—belong to evolutionary biology, while concerns about ontic vagueness *as such* belong to metaphysics. Reduction of objectionable kinds of vagueness might be a (non-scientific) tiebreaker between two competing theories, all things considered, but it does not directly bear on the virtues involved in scientific theory choice.

In this section, I have shown that the nihilist-friendly theory that replaces *E* is empirically equivalent to *E,* and so shares equally in many theoretical virtues that depend on such equivalence. Despite removing ontic vagueness from *E,* the nihilist-friendly theory retains vagueness in its predicates, and so ultimately solves no theoretical puzzles;

[15] See Brenner (2015: 1308-9) for one argument to this effect. Brenner does not address the issue of interpreting the vagueness in question as semantic in nature, and I will not address it here.

accordingly, it appears to have equal explanatory power. With such a tight competition between the two theories, I will examine the virtue of parsimony in case it is a tiebreaker. In the next section, I will show that, ultimately, the nihilist does not have an overall advantage in parsimony over the conservative if we take both ontological and ideological parsimony into account.

3.4 Nihilist Appeals to Ontological and Ideological Parsimony

In this section, I will examine the claims that nihilism results in a more parsimonious ontology and ideology than conservatism about objects, and this will answer our question as to whether a nihilist-friendly scientific theory would be more parsimonious than its composite-bearing rival. Nihilist claims to *ontological* parsimony originate in the fact that nihilism simply has fewer kinds of things in its ontology: there are only mereological simples, not composites. Most everything we believe in, from planets to protons, does not exist because these putative things are supposed to be made up of all kinds of parts, but really there are nothing but simples arranged proton-wise and planet-wise. In terms of *ideological* parsimony, the nihilist view also offers a clear and unified answer to van Inwagen's Special Composition Question (SCQ): when do some things make up a further thing? The nihilist answers: they never do. Thus, there is no need to provide conditions for composition of some things into other things as varied as planets or protons.

Unless he accepts brute composition, and with it the unpalatable consequence that nothing explains the fact that composition occurs in some situations and not others, the conservative about objects has to come up with an answer to the SCQ that gives necessary

and sufficient conditions for composition in all kinds of different situations.[16] An answer that gives the right results would have to include living things like otters, artifacts like dams, and other objects like rocks. This would probably be a very disjunctive set of conditions. More importantly, the number and kinds of objects yielded by the conservative's answer to the SCQ would be radically higher than those of the nihilist: a loss of ontological parsimony.

This difference in ontological parsimony must be taken in context, of course. On the one hand, the nihilist must start with the disadvantage that hers is a deeply counterintuitive view. On the other, nihilism is a unified all-purpose solution to several metaphysical puzzles that are generally considered difficult to answer from the standpoint of a conservative ontology.[17]

In this section, I will focus on the issue of whether nihilism's commitments in terms of ontology *and* ideology are favorable for nihilism on net balance because it results in greater parsimony. Whereas a theory's *ontology* is the set of things whose existence is assumed or entailed by it, *ideology* is the set of properties or relations whose existence is assumed or entailed by that theory.

4.4.1 The Challenge from the Special Arrangement Question:

Recently, Karen Bennett (2009: 62-66) has challenged nihilism's claims to greater overall parsimony by noting that any decrease of ontological commitments is purchased at the price of an *increase* in ideological commitments. In effect, if conservatives about objects are guilty of multiplying objects beyond necessity, nihilists are guilty of multiplying properties beyond

[16] But see Markosian (1998) for an argument in favor of brute composition.
[17] See Rettler (2018) for a good summary of nihilism's puzzle solving strategies and for an opposing view, but see Holly Kantin (2020) for a reply to Rettler.

necessity. This is because for every object *F* in a conservative's ontology, the nihilist will have to posit the property of simples being arranged *F*-wise. Because the nihilist is concerned with recapturing explanations about, say, multicellular organisms, these will have to be extraordinarily long, complicated properties that preserve the relevant compositional structure:

> "((((((there are simples arranged atomwise) arranged moleculewise) arranged organellewise) arranged cellwise) arranged organwise) arranged ...)" (Bennett 60).

For Bennett, in addition to being saddled with these strange, unparsimonious properties, the nihilist is also on the hook for something that directly parallels the SCQ. For every arrangement *F*-wise predicate, the nihilist must answer the question, "When is it the case that simples are arranged *F*-wise?" (66). Just as a conservative must provide an answer as to when some quarks compose a proton, when a proton composes an atom, when an atom composes a molecule, etc., so a nihilist must say when some simples arranged quarkwise are arranged protonwise, when some simples arranged protonwise are also arranged atomwise, and when some simples arranged quarkwise arranged protonwise arranged atomwise are also arranged moleculewise, etc. Jonathan Tallant (2014) has further developed this objection and calls the question the nihilist must answer regarding *F*-wise arrangement the Special Arrangement Question (SAQ). Importantly, the *ideological* cost of answering the SAQ is equivalent to the *ontological* cost of answering the SCQ, since for every composite object sortal *F* there will need to be a way for some simples to be arranged *F*-wise. The cost of commitment to the long, complicated *F*-wise predicates themselves is borne by the nihilist alone, and it is inversely proportional to the nihilist's ontological austerity.

4.4.2 Do Nihilists have to answer the SAQ?

Brenner (2015) has replied to this charge at some length. He argues that the nihilist does not have to answer the SAQ; however, if needed, the nihilist could do this without undermining the motivations for nihilism. And if the nihilist has to answer the SAQ, then so does the conservative, and the latter is no better off.

Brenner's argument for this first point is twofold. First, there is no reason why the nihilist *should* have to give a general account of arrangement *F*-wise. Probably, the conditions for this will vary by whatever specific sortal *F* we are talking about, and any general account is apt to be somewhat disjunctive. Nonetheless, the nihilist should be able to say something informative about the conditions for simples being arranged *F*-wise for some particular *F*.

Second, Brenner thinks the nihilist may not need to answer the SAQ because *nobody* needs to answer it. Brenner cites Ted Sider's (2013) view that nobody needs to use "arranged *F*-wise" predicates in their *fundamental* descriptions of the world, whether or not composition is true (perhaps because they can be replaced by fundamental non-arrangement predicates). Sider holds that the fundamental level is the only one wherein advantages in parsimony are relevant for truth, and that here nihilist will have the advantage over the conservative, who needs a fundamental parthood relation (Brenner 2015: 1298-99; Sider 2013: 239-41).

With regards to Brenner's first point, I just want to point out how similar the nihilist's position regarding the SAQ is to the conservative's regarding the SCQ. I think Brenner is right to say we shouldn't expect there to be any single, unified way of answering the SAQ because of the great diversity of (putative) composite sortal object kinds. This is for the same

reason that the conservative's answer to the SCQ is likely to be highly disjunctive. Any general story about the conditions for arrangement *F*-wise is likely to be just as disjunctive as some corresponding story about composition would be. This might mean that both questions are ultimately unreasonable. However, the motives for being able to give a general account of arrangement *F*-wise are the same as those for giving a general account of composition into *F*s. Composition into *F*s does a great deal of work in the conservative's ontology. If the composition relation cries out for explanation, then so do the "arranged *F*-wise" predicates in the nihilist's ideology. Even if everyone were converted to nihilism tomorrow, such that the average person understood the world around her in terms of pluralities of simples arranged humanwise, carwise, computerwise, cellwise, planetwise, etc.—she would still find these properties of being arranged *F*-wise to be among the most salient, important, interesting, and complex properties that simples have. Virtually the entirety of human knowledge would be understood in terms of *F*-wise arrangements, so there would be pressure to give not just a physical account of what these have in common, but to give a *metaphysical* account as well. We would expect the best such accounts to be as unified as possible. This is true even if, as Sider claims, arrangement *F*-wise predicates are non-fundamental.

I now turn to Brenner's second point, that it may be (as Sider suggests) that nobody will have to appeal to "arranged *F*-wise" predicates on a fundamental level, but that the conservative will have to include a fundamental parthood relation—ultimately giving the nihilist the advantage in terms of parsimony. The issue of fundamentality is a large one. Here I will only sketch a few reasons to think the nihilist's dialectical situation is not *prima facie* better than the conservative's.

To begin here, note that a conservative could make a similar move and declare that composites exist, but that neither the composites nor the parthood relation are fundamental.[18] Thus, by the standards of *fundamental* ontology the conservative's position is as parsimonious as the nihilist's. In fact, let us assume for comparison that our conservative is attempting to match a nihilist's fundamental ontology and ideology as closely as possible: he only commits to mereological simples and to as minimal a set of properties and relations as possible on the fundamental level. Let us assume for the moment that neither the nihilist nor the conservative needs any *F*s or things arranged *F*-wise, where *F* is a composite object, in their fundamental ontology or ideology.

Now, does the conservative need a primitive parthood relation in his fundamental ideology? This turns on whether we believe the fundamental level must furnish the basis for the non-fundamental parts of one's ontology and ideology, including the building blocks for complex non-fundamental predicates and entities. If it does, then the conservative needs a parthood relation in his fundamental ideology to ground the composition of mereological simples into even the most rudimentary composite objects (say, protons).

By the same token, the nihilist will need a primitive "arranged *F*-wise" predicate in order to ground the move from primitive fundamental predicates that apply directly to individual simples to the more complex predicates that pluralities of simples collectively bear at other, non-fundamental levels. This assumes that the simples at the fundamental level bear predicates in terms of, for example, spatiotemporal position, mass, spin, etc. These predicates by themselves don't say anything about whether the simples are arranged *F*-wise, or what

[18] For an extended discussion of this kind of ontological move on the part of a nihilist (a position he calls "Deep Nihilism"), as well as for the idea that the conservative can follow suit by developing "Deep Conservatism," see Korman (2015: Ch.6 §3.1-3.2).

would constitute *F*-wise arrangement. To define predicates that are in terms of counterfactuals about some nonexistent entity *F,* the nihilist needs a set of fundamental counterfactuals about conditions for composition into *Fs* and identity over time between *Fs.* These counterfactuals would need to be as complex and numerous as facts about the parthood relation in the conservative's ideology. (Depending on whether facts about composition are necessary truths, these may have to be counterpossible facts in the nihilist's fundamental ideology.)[19] This appeal to facts about nonexistent entities at the fundamental level of the nihilist's ideology is, to say the least, extremely odd, and it begs for explanation in a way that facts about parthood in the conservative's ideology do not.

Sider (2013: 239-41) does not spell out exactly why he thinks there will be no *F*-wise arrangement predicates in the fundamental level, but Brenner (2015: 1298-99) speculates that it may be because *F*-wise predicates could be replaced by more fundamental *non-F*-wise predicates. I am assuming that, for the purposes of recapturing conventional discourse about *Fs,* the nihilist ideology must ultimately accommodate *F*-wise predicates. Certainly, the nihilist could just eliminate *F*-wise predicates from the fundamental level—perhaps just relying on fundamental physical properties like mass, charge, and spin—and invoke a non-fundamental *F*-wise arrangement relation further up in the hierarchy of properties. However, to introduce *F*-wise arrangements at the level of, e.g., molecules rather than protons seems arbitrary and *ad hoc*. And it would mean the nihilist's ideology would no longer meet the criterion that the fundamental level contains the basis for one's non-fundamental ontology and ideology.

[19] See Ch.1 pp. 30-31 for my view of how this would play out in the nihilist debunker's dialectical situation, on the assumption that facts about composition are in fact necessary truths. Brenner (2021: 20) is neutral on this latter question.

Finally, there are reasons to think that protonwise, neutronwise, and mesonwise arrangement may be fundamental if *existential* dependence is a criterion for what is included in our fundamental ontology and ideology. Even if quarks are mereologically independent in the sense of being mereologically simple—that is, they are not dependent upon any parts making them up, because they have no parts—they may be existentially dependent on being in arrangements with other quarks, because individual quarks do not (and apparently cannot) exist outside of being arranged with other quarks either protonwise, neutronwise, or mesonwise. This means that even if our fundamental physics reduces matter into irreducible, fundamental particles, our fundamental physical descriptions of the world still must contain *F*-wise predicates because some kinds of fundamental particles, if they exist at all, must be arranged *F*-wise. If true, this would also mean that the minimal conservative view just discussed would need a primitive parthood relation (see Tahko 2018: §1.1).

4.4.3 Conservatives and the SAQ

Brenner merely raises these two objections as reasons we should not *assume* that anyone needs to answer the SAQ. He then argues that there is no reason the conservative will have an easier time answering it than the nihilist. On this latter point, he gives the example of Crawford Elder's (2011) analysis of what it means for something to be arranged *dog*-wise:

> "[M]icroparticles are dogwise arranged just in case (i) they are among a plurality of microparticles that between them are such as to cause the folk to judge that a dog is present, and (ii) they lie within the region occupied by a dog" (Elder 2011: 124).

Clause (ii) ties arrangement dogwise to the existence of dogs, so here Elder is clearly using a resource only available to conservatives. We can generalize Elder's analysis to *F*-wise arrangements in general. But Brenner sees three problems with the analysis. First, it doesn't

cover things arranged *F*-wise that don't compose anything, either in the case of complex arrangements that don't correspond to sortals for composite objects or in the case where something needs the artist. Second, it doesn't say anything informative about what makes the object fall under the sortal *F*. Finally, this is a poor conceptual analysis because it isn't necessary that the microparticles should be such that they cause folk to believe an *F* is present (because, for instance, there may be no folk); and the concept of there being things arranged *F*-wise does not include the concept of those things lying within a region occupied by an *F* or of those things composing an *F*.

I generally agree that Elder's analysis does not work, for reasons I will not elaborate on here.[20] However, I want to push back against some of the notions of "arrangement" Brenner is using here and clear up a potential confusion. The motive for the "arrangement *F*-wise" locution being in the nihilist's toolbox at all is that it allows her to recapture the claims of conservatives by specifying how some things are arranged *precisely as if* to compose an *F* even though they don't (because there are no *F*s). The extent to which it fails to do this is the extent to which it does not serve the nihilist's purpose. "Arrangement *F*-wise" is a philosophical term of art with a specific function. The challenge of the SAQ is the same as that of the SCQ: to give an account of this set of properties or relations that covers all and only the things that are arranged *F*-wise, and to do so in a way that is both unified and informative.

If we ask the notion of "arrangement" to do more work than this, we are moving the goal post for both the nihilist *and* the conservative. Though Brenner objects that Elder's

[20] In broadest strokes, I believe the ties to folk belief and lack of informativeness represent substantial problems.

account does not allow for cases where some things are arranged F-wise but do *not* compose an *F*, Brenner's examples suggest he is asking for an expanded sense of the term, under which something arranged *F*-wise does not meet the *conditions* for composing an *F*. For example, if some things are arranged statuewise but do not compose a statue (because of the absence of the right origin including an artist's intentions), this is a case where things would have only *some* of the properties necessary for a conservative to say they compose an *F*. In this case, the "arranged *F*-wise" locution does *not* recapture the conservative's claims about the matter. The conservative might say, "It is shaped just like a statue, but it isn't one." If saying some things are "arranged *F*-wise" means that they meet *all* the conditions for composing an *F* (if *F*s existed), then those wanting to speak only in terms of arrangements need some *other* locution to recapture what a conservative would say about such an *almost*-statue.

English speakers already have a loose, colloquial sense of "arranged" or "shaped" that either the nihilist or conservative is free to use where things meet the conditions for putative composition of an *F* even more loosely. Isn't that fluffy white stuff in the sky bear-shaped? Yes, it is. Aren't those things collectively running around the yard arranged dogwise? Sure, they are. In a sense, my terrier's assortment of things arranged toy-lambwise is also (approximately) arranged lambwise. But if the nihilist wishes to wed *this* use of the term to the technical, nihilist-friendly locution (or to hold the conservative to this standard) it will come at the additional cost of some predicates for distinguishing the *almost F*-wise and the *roughly F*-wise from the *truly F*-wise, and at the cost of making the account disjunctive.

In the same way, it is unclear why the conservative would be on the hook for giving an account of "arrangement *F*-wise" that would cover complex arrangement predicates that

don't correspond to composite object sortals but are merely specified by some description. This is beyond the purview of the SAQ because the very question presupposes that in any "arranged *F*-wise" predicate, *F is* a composite object sortal. As with the loose, colloquial sense of "arranged," the conditions for this descriptive sense of "arranged" will be disjoint from any conditions for being arranged such as to compose an *F* (if there were *F*s) and they apply to both the nihilist's account and the conservatives.

My point is that these additional senses of "arranged *F*-wise" are irrelevant to the SAQ, and neither the nihilist's nor the conservative's answer needs to account for them. To drive the point home for the nihilist, let us look at a popular counterfactual analysis of the "arrangement *F*-wise" locution offered by some nihilists—one that Brenner defends at some length (2015: 1301-6). Here is Brenner's rough characterization of the analysis: "*x*s are arranged *F*-wise iff they are arranged in the manner in which they would be arranged if they composed an *F*" (1302). If we were to hold this analysis to the standard of needing to cover these additional senses of "arranged," it would be inadequate. This would not cover instances where some things would *almost* qualify, whether certain conditions for composing an *F* were only met approximately or not at all (including cases that would be expressed in the loose, colloquial sense of "arranged"). It would also not cover *F*s that were not sortals for composite objects, for there are no conditions under which things so arranged would compose something (unless, perhaps, we are looking at the nearest possible worlds wherein they *do* compose an *F*).[21]

[21] Brenner objects to Elder's account of "arrangement F-wise" because it depends on the microparticles composing an F. It's unclear whether the counterfactual reference to the *conditions for composing an F* are still parasitic on the concept of composition or F-ness in an illicit way. (If so, it would be a distinct objection from Tallant's GLOBALIZATION objection (2014: 1515)). I will not develop the objection in this chapter.

But does the conservative even need to answer the SAQ? Brenner (2015) argues that contra Bennett (2009), the conservative *does* need to answer the SAQ, and that it is an additional ideological cost over and above what is needed to answer the SCQ. Here I will sketch the disagreement and then defend Bennett's position. In what follows, I will assume that the nihilist could marshal sufficient resources to answer the SAQ.[22] What I will contest is the idea that she can do so without compromising the main motivations for nihilism (particularly ontological and ideological parsimony).

According to Bennett (2009: 60-64), the nihilist's ontology *plus* ideology will be at least as crowded and unparsimonious as those of conservatives, because "arranged *F*-wise" predicates will have to be compounded through each level of (putative) composition. Ontologically, the nihilist is only on the hook for mereological simples. However, ideologically, if she wants to recapture claims about putative dogs, she needs to have predicates for simples arranged protonwise, and (simples arranged protonwise arranged atomwise), and (((simples arranged protonwise) arranged atomwise) arranged cellwise…) and so on, to account for arrangement tissuewise, organwise, organ-systemwise, and on up to dogwise.

However, in Bennett's view the conservative's commitments are as follows. He can say that it is simples that instantiate atomwise arrangement, and cells that instantiate tissuewise arrangement, and perhaps whole organ-systems that instantiate dogwise arrangement, but simples do not directly instantiate *cellwise* or *tissuewise* or *dogwise* arrangement. Thus, he is not committed to predicates like "((((arranged atomwise) arranged

[22] See Brenner (2015: §3) for rebuttals to arguments from Tallant (2014), Elder (2011), and Unger-style sorites arguments purporting to show that accounts of *F*-wise arrangement are doomed to fail.

moleculewise) arranged cellwise) arranged tissuewise)" or the tortuously long nested property of dogwise arrangement, as it would need to be applied to simples (2009: 64).

Brenner's response is that clearly someone who believes that xs compose *F*s *also* believes in *x*s arranged *F*-wise. So, a conservative would have to endorse all the same compounded "arrangement *F*-wise" predicates that a nihilist would, and so he would have a bloated ontology *and* ideology. That is, the conservative would have to commit to whatever is needed to answer the SAQ *in addition to* the SCQ. And, contra Bennett, the conservative can't just say that it is, for example, only the molecules that are arranged dogwise, while the atoms are merely arranged moleculewise. By being so arranged, the atoms will thereby be ((arranged moleculewise) arranged dogwise). And this is just intuitive: the atoms seem shaped like a dog! (2015: 1311).

The challenge to the conservative is, then: Don't you have to say simples are arranged dog-wise? And doesn't this commit you to complicated arranged dogwise predicates you are saying only the nihilist needs? I will argue that if the conservative has a solution to the SCQ in hand, he already has the resources to answer the SAQ, even though it isn't necessary for him to do so. The solution to the SCQ need not be neat and clean. We will just assume for the present argument that the conservative has conditions for composition into *F*s in hand, whatever these conditions may be.

Aren't those simples arranged dogwise? They sure seem shaped like a dog. Does the conservative have to agree here? The conservative can take this as a question that is asked in the loose, colloquial sense of "arranged" or "shaped" we designated earlier. In this sense, he can freely answer "yes" without incurring any additional costs in terms of ideology. Whatever stuff is down there at the microscopic level, it is arranged dog-wise in the same

sense that whatever microscopic stuff is right now filling up my terrier's stuffed brontosaurus toy is arranged dinosaurwise.

However, assuming the question is meant more precisely, perhaps if it is posed by a metaphysician inside the ontology room, the conservative is free to agree or disagree. But in neither case does he incur the costs associated with the complicated "arranged dogwise" predicates that the nihilist must use. To be clear, the nihilist needs these predicates because in her ontology, simples are the *only* bearers of properties, and these properties must capture the conditions for composition into an *F*, if *F*s existed. However, the conservative has other options.

The conservative does not need a special predicate specifying the set of conditions for simples to be arranged dogwise because he already captures these conditions piecemeal, in terms of a hierarchy of composition. They are not conditions on the way simples are arranged because the conditions that apply to atoms, for instance, do not apply to simples. He doesn't have to answer the question of how simples can be arranged dogwise because he has a story about how they *compose a dog*. They do so by meeting the conditions for composing atoms that compose molecules that compose tissues that compose organs, etc., that compose a dog.

So, another option for the conservative is to answer yes, the simples are arranged dogwise, if by that you mean "arranged such that they compose a dog." Understanding "arranged *F*-wise" to be just meeting the conditions for composing an *F* (if there were *F*s), the conservative will agree that simples that compose an atom are arranged atomwise. But composition is transitive. If some things compose an A that is among some things that compose a B that are among some things that compose a C, then those things partly compose

a C.[23] So, those simples compose a dog by being arranged atomwise, where the atom they compose is part of a molecule that is part of a cell that is part of a patch of tissue that is part of an organ, etc. … that is part of a dog. And in being arranged atomwise for an atom in this particular situation, the simples are *arranged such that they (partially) compose a dog.*

Another option for the conservative is to allow that simples can satisfy a complex, nested predicate like ((arranged atomwise) arranged moleculewise)—but only *indirectly*, via composing an atom. In this way, for simples to be arranged dogwise, they would satisfy some complex dogwise arrangement predicate only by indirectly satisfying a series of predicates that are directly satisfied by composites at each level of the hierarchy. This reply seems consistent with Bennett's argument (2009: 64). It might be argued that allowing for special sense of indirect satisfaction of predicates is a burdensome addition to the conservative's ideology, and that this—along with the inherent costs of formulating *F*-wise arrangement predicates in the first place—is sufficient reason for rejecting this strategy.[24] Fortunately, I do not think the conservative needs to accommodate these predicates at all.

Here is my reasoning. First, I don't need a predicate in my ideology for every non-occurrent belief that I would ever assent to, especially if that belief were simply a conjunction of lots of separate things that happen to be true right now. Perhaps I believe that my friend is on her way back home from the grocery store toting a bag of groceries, and that she is riding a unicycle, happens to be facing Tokyo, and is whistling "La Marseillaise." Would I assent to the statement that my friend is slowly moving her groceries toward Tokyo via unicycle while whistling "La Marseillaise?" Yes. Does this mean I need a predicate for

[23] See Korman and Carmichael (2016: §1.1) on composition and transitivity.
[24] See Uzquiano (2004) for a summary of the technical worries for the "arranged *F*-wise" locution.

"slowly moving one's groceries toward Tokyo via unicycle while whistling "La Marseillaise?"" in my ideology to deal with just this situation? No. However, as Brenner (2021: 5) notes: just because we eliminate something from our ideology, doesn't mean we must eliminate it from our *vocabulary*. I can agree that in some indirect sense, the simples arranged atomwise are also arranged moleculewise, without inventing a special predicate to capture the relation. Second, the conservative also has a perfectly good story about how those simples are arranged, one that perfectly captures the conditions for those simples composing a dog.

Brenner offers two additional arguments as to why, when a conservative believes that xs compose an *F*, he is already committed to the idea that these xs are arranged *F*-wise. First, the fact that some xs are arranged *F*-wise is *the reason* they compose an *F* (1311). Second, nihilist is not introducing a new predicate that the conservative doesn't already use, but saying *I don't accept your one claim (about composition), but I accept the other (about arrangement F-wise)* (1312).

Regarding the first claim, the conservative should agree with Brenner only in very localized cases of composition that are able to be satisfied by simples. For the conservative, the fact that some simples are arranged atomwise *is* the reason they compose an atom—because to satisfy the arrangement atomwise predicate is just to satisfy the conditions for composition into an atom.[25] However, when considering some xs that could not collectively satisfy the conditions for composing and *F* unless they first compose a G, the xs (strictly speaking) cannot be arranged *F*-wise; only a G can be arranged *F*-wise. These xs can

[25] Strictly speaking, we should allow for the possibility that simples, if they exist, are diverse. For instance, there may be at least as many kinds of simples as there are kinds of fundamental particles in the Standard Model. Among these, *electrons* can be arranged atomwise, but *quarks* must first compose a proton.

compose an *F* only by composing a G that is arranged *F*-wise. So, outside of very low-level situations involving predicates that are able to be satisfied by simples, it is *false* that some xs being arranged *F*-wise is the reason they compose an *F*. For the believer in composition, there is no way that some simples can *be* arranged dogwise without first composing an atom. Thus, the conservative's story of arrangement dogwise, kidneywise, or moleculewise is *different* from the story told by nihilists, and so then are the conservative's ontological and ideological commitments.

Regarding the second claim, the conservative can answer: no, I don't make the claim that the simples compose the dog *and* the entirely separate claim that simples are arranged dogwise—if by this you mean the claim that the simples satisfy complex *F*-wise arrangement predicates and that this is my reason for supposing they compose a dog. I can just say: the simples compose a dog, and they do so by meeting the conditions for composing a dog (again, this means by composing an atom that with some other things composes a molecule that with some other things composes... a dog—the conservative's story I outlined above.)

4.4.4 Predicates for *F*-wise Individuation and Persistence Conditions

Finally, I want to conclude by noting that the nihilist will need not just "arranged *F*-wise" predicates, but relational predicates capturing individuation and persistence conditions in propositions about *F*s that rely on distinguishing or counting *F*s, or that attribute properties to *F*s over time.[26] Statements of evolutionary biology will have to distinguish, say, different

[26] In Chapter 1 §8, I argued that any nihilist who wants to recapture the content of scientific statements cannot simply use "arranged F-wise" predicates alone, because no one plurality of simples stays arranged F-wise long enough to do the things we attribute to *F*s over time. The nihilist will ultimately have to reify "arrangements F-wise" that have the persistence conditions of *F*s, thus letting composite objects in through the back door, or embrace a "schmidentity" relation that will connect different pluralities of simples arranged *F*-wise in a way that recaptures scientific propositions about same-*F* relations over time (where *F* is some composite object sortal). Brenner explicitly embraces this latter strategy (2018: 674).

baby chicks in the same nest. To do so, the nihilist-friendly theory of evolution will need individuation conditions for groupings of simples arranged *F*-wise, such that they are able to recapture statements about some *particular, countable F*. Otherwise, what the conservative regards as a nest of eight chicks might just be a giant mass of simples arranged chickwise. What is needed is a relational predicate that obtains among some group of simples arranged *F*-wise at a particular time *t*, such that they can be distinguished from other *F*-wise pluralities in the area at *t*. This would allow practitioners of the nihilist-friendly theory to capture propositions that ascribe properties like yellow color to, say, each plurality arranged particular-chick-wise, and to capture propositions that rely on distinguishing and counting the chicks. Natural selection measures success by the numbers: the number of offspring per parent, the number of genes for a particular trait in the gene pool, and so on. Call this the "arranged *particular-F*-wise" relation.

To capture statements of evolutionary biology about the same chick over time, the nihilist-friendly theory needs to tie together groupings of simples bound together by the "arranged *particular-F*-wise" relation at various times from $t_1 \ldots t_n$ with a predicate expressing an "arranged *same-F*-wise" relation. Brenner acknowledges that this kind of same-organism relation must be in the nihilist's paraphrase repertoire (2018: 674).

Both the "arranged *particular-F*-wise" relation and the "arranged *same-F*-wise" relation are strange in their dependence on the individuation and persistence conditions for non-existent objects, and unparsimonious because they depend on a large set of counterfactuals (or counterpossibles) that seem to all be required in the fundamental ideology because they cannot be reduced to more fundamental facts about such nonexistent objects.

This is in addition to the parsimony concerns I mentioned above, dealing with the standard "arranged *F*-wise" paraphrase strategy.

In this section, I have given reasons that, contra Brenner (2015), the nihilist really does seem to be committed to answering the SAQ and to lose a great deal of ideological parsimony in comparison to the conservative. In fact, there seems to be a perfect balance of ontological *plus* ideological parsimony between nihilism and conservatism. Everywhere the nihilist eliminates a composite object kind, it gains an arrangement predicate. If the basic ideological machinery for composition into *F*s should be included in the fundamental level of the conservative's ontology, so the basic machinery for the move from simples to simples collectively arranged *F*-wise should be in the nihilist's fundamental ontology.

4. Conclusion: Tiebreakers

If I have argued successfully, I have shown that the advantages of nihilism, whatever they may be, must belong to metaphysics: its ability to resolve long-standing problems like those involving ontic vagueness, the problem of the many, or the problem of material constitution. A change from conventional evolutionary biology to a nihilist-friendly theory seems to confer no benefits that would justify the change—no benefits, that is, in terms of theoretical virtues.

Nonetheless, if the nihilist-friendly scientific theory fares equally well in competition with the original theory with regard to all the virtues involved in scientific theory choice, consistency with metaphysics may very well be a tiebreaker. Since the debates between sparse, conservative, and permissive ontologies goes on, it is unclear where such an advantage would lie. Only time and continued debate will tell us the answer.

I have said little about the practical considerations of moving from one theory to the other. However, in the absence of either scientific *or* metaphysical reasons for theory change, these may be decisive. The nihilist theory of evolution would be a *new* theory, and it will be judged at least in part by how it compares to the old one. The principle of *conservatism* counsels us that, all things being equal, the fewer previous beliefs your theory rejects, the more plausible it is.[27] Conservatism speaks against the nihilist-friendly theory, because the latter systematically falsifies the great majority of the claims of *E*. The motivation for the principle of conservatism is twofold: to maintain a steady, secure connection to one's base of empirical evidence as one moves from old theories to new ones; and to avoid the costs involved with moving to a new theory. The nihilist can easily push back on this first motive, as the nihilist-friendly theory falsifies the claims of *E*, but it does so in a way that sacrifices none of the new theory's connection to empirical evidence. After all, it is empirically equivalent to *E*.

As for the second motive, that of avoiding the costs of moving to the new, nihilist-friendly theory, I will close with two thoughts. The nihilist theory will be difficult to use, and difficult to understand. First, the theory uses all the ponderous, nested arrangement *F*-wise predicates previously mentioned, which will be difficult to adjust to and use for both theorists and practitioners. Of course, after a while, speakers can adjust and develop shorthand and figures of speech to efficiently communicate even the most complex and difficult ideas. But this brings us to the final point. The nihilist-friendly theory of evolution asks students, theorists, and practitioners to accept that strictly speaking there are no organisms or species, and that survival and reproduction must be reinterpreted in completely novel terms. If these

[27] Loosely adapted from Quine and Ullian (1978: 67).

cognitive gymnastics were motivated by a new theoretical paradigm that promised to genuinely resolve long-standing issues and illuminate difficult aspects of the subject, perhaps making testable new predictions, then there would be no question of their being justified. However, according to the analysis in this chapter, the nihilist is holding forth costs without overriding benefits.[28]

[28] See Foley (1983) and Feldman (2003: 143-44) for some objections to epistemic conservatism.

Chapter 4:
Interfaces and Objects

1. Introduction

In this chapter, I will raise objections to the Interface Theory of Perception (hereafter ITP) and some of its skeptical implications for human perceptual beliefs. The view was introduced by Donald Hoffman (2009) and developed primarily by Hoffman, Manish Singh, and Chetan Prakash.[1] (Hereafter, for simplicity, I will attribute the theory to Hoffman, who has developed the most comprehensive philosophical arguments supporting, and supported by, ITP.) ITP is the thesis that the perceptions of any evolved organism, including those of human beings, constitute a species-specific user interface that guide the perceiver to perform fitness-enhancing actions in a way that almost certainly does not represent the world or its features veridically.

Hoffman uses evolutionary reasoning to draw sweeping skeptical conclusions about human perceptual beliefs in a manner that parallels some recent evolutionary debunking arguments about ordinary object beliefs in the literature.[2] By comparison, Hoffman's argument is unusual in that it appeals not just to reasoning and examples from evolutionary biology, but to experiments utilizing evolutionary game theory and genetic algorithms; in addition, it has a more broadly skeptical conclusion: it also threatens our beliefs about spacetime and causality.

[1] See Hoffman, et. al. (2015) for a full statement of the view. See also Prakash et. al. (2020) for the most recent game theoretical results, and Hoffman (2019) for the most recent comprehensive statement of the view and its implications, presented for the general reader.

[2] For an overview of object debunking arguments, see Korman (2014, 2019a, and 2019b). Some examples can be found in Merricks (2001: 72–76) and Benovsky (2015: §2).

Recently, philosophers have criticized various aspects of ITP's game-theoretical analysis of perception.[3] This chapter will focus on articulating and criticizing the broader evolutionary argument supporting ITP, and the debunking argument reasoning from ITP to perceptual skepticism. I will argue that even if evolutionary biology straightforwardly grounds Hoffman's experiments and entails ITP, the argument justifying ITP faces a self-defeat problem that is not circumvented by an appeal to Universal Darwinism.

Here is a plan of the chapter: In section two, I present ITP, the argument supporting it, and its skeptical implications. In section three, I argue that the argument for perceptual skepticism via ITP is self-undermining and present Hoffman's response appealing to Universal Darwinism. I then set up a dilemma for Hoffman: either evolutionary biology is an instance of Universal Darwinism, or it is not. In either case, the argument for perceptual skepticism via ITP fails.

2. ITP and Perceptual Skepticism

The basic idea behind ITP is that our perceptions are related to the world in roughly the same way that the elements of a graphical user interface on a computer (for example, the icons on a Microsoft Windows desktop) are related to the complex physical hardware inside that computer. They do not veridically represent the underlying hardware situation (in fact, they don't purport to represent it at all), but they do allow us to interact productively with it. Likewise, our perceptions do not veridically represent the world, but they allow us to interact with it in a way that enhances our overall reproductive fitness.

[3] For examples, see Cohen (2015), Hummel (2015), and Martínez (2019).

This is in contrast to the conventional view that our perceptions, however limited, evolved to be generally veridical because it was adaptive for our ancestors to see the world as it is. Hoffman holds that the conventional view cannot be true because perceptual strategies tuned to fitness without the additional constraint of veridicality tend to dominate perceptual strategies that report at least some aspects of the world veridically. Hoffman and others have attempted to show this by applying evolutionary game theory to test various mathematical models of perception against each other. This is Hoffman's main evidence for ITP. To illustrate his experimental results in a way that's more intuitive, Hoffman also points out numerous examples of evolutionary satisficing in the perceptual systems of other species and reasons about human perception by analogy.

In this section, I will state and motivate the argument from ITP to perceptual skepticism. Along the way, I will give examples from evolutionary biology, sketch some of the results from Hoffman's experiments in evolutionary game theory, and develop Hoffman's central interface metaphor. I will break ITP*'s* argument for perceptual skepticism into two main parts:

1) The argument that evolutionary biology entails ITP
2) The argument from ITP to perceptual skepticism

I will address Part I in Section 2.1 and Part II in Section 2.2.

2.1 The Argument that Evolutionary Biology entails ITP

To support ITP, Hoffman appeals to reasoning and empirical evidence from evolutionary biology, according to which natural selection is the main force driving the development of species and their traits. Perception is a trait that evolves by natural selection, and as such it is

subject to the constraints of overall reproductive fitness in the organisms possessing it. If, during the evolution of a species, the development or improvement of some trait X in members of that species would sufficiently compromise their overall reproductive fitness, genes for developing or improving X in that species' gene pool will eventually lose out to rival alleles and be driven to extinction. This limitation of the development or improvement of the trait X in the name of overall reproductive fitness is what's known as evolutionary *satisficing*. Roughly, it's the reason certain traits don't function better than they do: these traits were "good enough" at their task to secure the reproductive success of organisms bearing them, and genes for improving them either never occurred in the gene pool or were outcompeted by rival alleles that were more conducive to overall reproductive fitness.

One of Hoffman's central examples of evolutionary satisficing in perception is that of the Australian male jewel beetle *Julodimorpha bakewelli.*[4] It flies around in search of attractive females to mate with, and its behavioral rule of thumb is to seek out anything shiny, dimpled, and brown. Unfortunately, people now litter the beetles' habitats with sundry discarded beer bottles that satisfy just this description: they are superlatively shiny, dimpled, and brown. In fact, they appear so sexually attractive to male beetles that they act as a supernormal stimulus; the male beetles will prefer them to *actual* female jewel beetles and will try to mate with the bottles until they are eaten by the ants that have learned to live and hunt near the discarded bottles. According to Hoffman, when a male beetle attempts to mate with a bottle, it's a symptom of a perceptual error—specifically, an error of perceptual categories. In response to the presence of the bottle, the beetle forms a representation of a

[4] See Gwynne and Rentz (1983), Hawkeswood (2005), and Robertson et. al. (2013) for more information on *Julodimorpha bakewelli* and other animals that have fallen into evolutionary traps involving evolved perceptual limits.

superlatively sexy female beetle; but this perception is inaccurate, owing to evolved limitations in the beetle's visual system.[5]

Hoffman stresses that evolutionary satisficing is an adaptive compromise that supports overall reproductive fitness for the organism within its particular niche, taking into account such factors as its existing structures, its environment, the cost of using the perceptual organ, and the cost of perceptual error. In the beetles' evolutionary past, it may have paid off to have a perceptual category for "attractive female" that was cheap and easy, allowing male beetles to reliably use a few easy-to-detect markers like shininess, dimpled-ness, and brownness to detect them. This would work perfectly well in the environment of the modern jewel beetle's ancestors, where there were no beer bottles and the only shiny, dimpled, brown objects were females.

Put negatively, we may assume that if traits for more accurate perceptual system had ever been available (i.e., one that would have prevented the error with the bottles), these traits must have lost out to the traits for the perceptual system the beetles currently have. After all, perceptual systems that give more information about their objects tend to be more costly to build, maintain, and use. It's possible that the cost of error was so low in the ancestral environment that the cheapest-to-build system that identified female beetles with even a relatively low frequency of success would always win out.

I want to pause briefly here and note what I think we should and should not take from Hoffman's empirical examples here. His example of the jewel beetle does not necessarily

[5] It is controversial just how complex are the representations we should ascribe to jewel beetles. Perhaps they just represent shiny, dimpled, brown things as something-to-be-mated-with (their physiological responses make this latter part incontrovertible). I think it's at least plausible that in their ancestral environment, the best candidate for *representatum* here would be something like "attractive mate."

show anything about the beetle's ancestral environment, or how reliable the senses were in that environment. It is a good example of how the limitations of an evolved perceptual system can result in systematically mistaken perceptions in a radically *changed* environment. However, Hoffman is not arguing that *our* environment has changed dramatically. I think Hoffman just means to establish here that the evolved perceptual systems of organisms have limits, established at least in part by what was needed to achieve reproductive fitness in their ancestral environments, and that an organism might fall short of some *particular* standard of perceptual veridicality because of these limitations. For instance, the jewel beetles make mistakes that are easy to identify by *our* standards of perceptual veridicality, of which their perceptual systems fall short. Under normal visual conditions, we would never confuse a beer bottle and a beetle.

Ultimately, Hoffman's examples of jewel beetles and the like are easy to understand examples of satisficing meant to prime the reader for his more radical conclusions. Hoffman considers the real, rigorous arguments for ITP to involve mathematical models, game theory, and genetic algorithms. I will turn to these shortly, as they are the primary support for a premise I will call *Fitness Beats Truth.* First, it will help us to cast the broad claims of evolutionary biology relied upon by ITP as a specific thesis:

1. *Evolutionary Biology*: All species of living things on earth (including our own species, Homo sapiens), have evolved primarily by the process of natural selection, and the traits of organisms from all species have primarily been molded according to the requirements of overall reproductive fitness in their ancestral environments.

Since *Evolutionary Biology* is a claim about all life on earth and the way natural selection molds the traits of various species, it entails specific claims about its instances: particular evolved species and their particular evolved traits. Here we are most interested in

human beings and their evolved trait of perception, which we will capture in the following thesis:

> 2. *Human Perception*: Our own trait of perception has been molded according to the requirements of overall reproductive fitness in our ancestral environments.

This much is uncontroversial among evolutionary biologists and perceptual psychologists. However, Hoffman has a radical view of the way the requirements of reproductive fitness have molded human perception.

In order to explain and motivate the third premise, *Fitness Beats Truth,* it will help us to first contrast Hoffman's view with what he takes to be the conventional view of perceptual veridicality, and then to define a few key terms. The conventional view is that our perceptual systems are adaptive (at least sometimes, and at least in part) *because* they represent the world veridically.[6] That is, a perceptual system tuned to fitness is likely to deliver us at least some truth about the world as well. According to this view, evolution will tend to produce perceptual systems that deliver veridical perceptions because these will keep us safe, allow us to find mates, and in general situate us to produce more offspring. If we identify things like mates or food more accurately than our reproductive rivals, we are likely to out-compete them for these resources. And if we are better than our reproductive rivals at distinguishing the tiger from the tall grass in which it's hiding, we are less likely than they are to become its dinner.

Hoffman points to what he sees as abundant exceptions to the conventional view. The jewel beetles are far from an isolated case; there are similar examples throughout the animal

[6] See Hoffman (2009: §1.2) and Hoffman et. al. (2015: 1480-81) for a characterization of the conventional view.

kingdom.[7] It is sometimes suggested that we should not attribute veridical perceptions to things with very rudimentary perceptual systems (e.g., flies and frogs) but we should attribute them to humans, at least under normal conditions.[8] However, Hoffman believes that perceptual systems in general are subject to a deeper and more fundamental kind of satisficing than is normally appreciated, and that ultimately human perceptual systems are no more likely to deliver us veridical representations of the world than those of flies, frogs, or jewel beetles. In fact, he contends, the selective pressures on our ancestors' perceptual systems almost certainly involved tradeoffs that prevented them from evolving even the most minimal kind of veridicality such as representing the world's real structure. Moreover, for Hoffman the conventional view suggests the evolution of our perceptual capacities was sharply discontinuous with that of other evolved life, and this is a strike against the conventional view, because it runs afoul of one of Darwin's central insights: the unity and continuity of life on earth.

To clarify Hoffman's position and motivate *Fitness Beats Truth,* I will now summarize some of his terminology and methods from evolutionary game theory. Evolutionary game theory is a kind of mathematical modeling based on game theory in classical economics, which has been used to test hypotheses about various evolved strategies

[7] In addition to male Western Australian jewel beetles, Hoffman lists "dragonflies that mistake gravestones for water, gull chicks that prefer red disks on cardboard to their real mothers, frogs that die of starvation when surrounded by mounds of unmoving edible flies, and birds that prefer brightly speckled rocks or the eggs of cowbirds to their own eggs" (Hoffman et. al. 2015: 1490).

[8] For example, David Marr, in his theory of vision, seems to suggest this kind of distinction between human perceptions, which really represent features of the external world, and more primitive perceptual mechanisms in, e.g., frogs and flies (1982: 340). It may be that these creatures veridically represent things (like "a surface I can land on" or "a black speck I can eat") but simply have very few representational types, and nothing to match our general representations of 3-dimensional space or 4-dimensional spacetime. I thank Kevin Falvey for this insight.

in organisms including those for resource gathering, perception, and altruistic behavior.[9] Hoffman and his team use it to test various kinds of perceptual strategies in competition with each other, to see whether individual strategies will outcompete rival strategies and drive them to extinction, be driven to extinction themselves, or stably coexist with rival strategies in certain ratios in the population.

In his experiments, Hoffman identifies several kinds of perceptual strategies, all of which he terms *interface strategies.* A perceptual interface strategy defines a mathematical relation between a set of perceptual experiences and a set of world states, both of which are mathematically defined in terms of measurable spaces. The strategies can be roughly divided into a group of *realist strategies* of various kinds, and what he calls a *strict interface strategy*. The realist strategies range from those in which perceptual experiences must capture all states of the world with complete accuracy, such that the function from world states to perceptual experiences is *isomorphic*, to those where perceptual experiences need not match world states at all, but the function from world states to perceptual experiences must preserve all world *structures*. This latter, structure-preserving realist strategy is called a *critical realist strategy.* By contrast, a *strict interface strategy* is one that simply has no veridicality constraints: perceptual states need not match the world, nor does the function from world states to perceptual experiences need to preserve any of the world's structures. In his experiments, Hoffman focuses on the competition between strict interface strategies and critical realist strategies, because the latter make the weakest (and so easiest to satisfy) kind

[9] See Barnard and Sibly (1981) on resource gaining strategies of producer-vs.-scrounger, Harper and Pfennig (2007) on perceptual strategies involving Batesian mimicry, and Axelrod and Hamilton (1981) on strategies relating to altruistic behavior.

of realist claim. If critical realist strategies fare poorly, then *a fortiori* the stronger kinds of realist strategies will also lose out to strict interface strategies.

We are now in a position to state the *Fitness Beats Truth* thesis:

3. *Fitness Beats Truth*: In the evolution of perception, it is highly probable that the requirements of overall reproductive fitness will cause strict interface strategies to drive rival realist interface strategies to extinction, *or* that those same requirements will cause realist interface strategies to not evolve in the first place.

The idea behind *Fitness Beats Truth* is that "fitness-tuned" strict interface strategies will outperform "truth-tuned" realist interface strategies over time, gradually driving the latter to extinction.[10] Hoffman's reasoning is that this is because strategies that represent resource quantities veridically are almost certainly *less* efficient at reporting fitness payoffs. Fitness payoffs are unlikely to correlate with resource quantities directly. The possibilities for functions representing fitness payoffs are vast, and only a tiny subset of these match those that track real-world resource quantities. And fitness payoffs are likely to diverge from resource quantities for biological reasons. Fitness payoffs are determined partially by the organism's needs: there is a *right* quantity of fitness-enhancing resources that would be best to acquire, in terms of fitness payoffs—a right concentration of oxygen to breathe, a right amount of calories to consume, a right amount of heat to have in one's immediate environment. It would be less beneficial to act to acquire a quantity of resources that was either too little or too much for the organism to use in the pursuit of survival and reproduction.

[10] This is a generalization from Hoffman's experimental results, not to be confused with the more specific "'Fitness Beats Truth' (FBT) Theorem," from Prakash et. al. (2020: 17), which I briefly describe below. See also note 10 below.

In what follows, I will show how Hoffman takes his experiments to support *Fitness Beats Truth.* I will first describe the simple perceptual interface games reported in Hoffman et. al. (2015), as these present the interface strategies in their simplest form. I will then briefly describe two variations on these basic interface games: the more complex Bayesian visual strategies tested in Prakash et. al. (2020), and the experiments involving genetic algorithms and the evolution of perceptual strategies presented in Mark (2013).

Hoffman (2015) sets up a series of simple interface games wherein players representing strict interface and critical realist strategies are forced to compete for resources distributed over territories. The resources in a territory vary from 0-100, and one might assume that more is better, and the players should simply estimate the real amount of resources and try to beat the other player to the territories with the most resources. This would only be true if more resources always meant a higher fitness payoff. However, the *payoff function* for a set of resources is defined by a Gaussian function (a bell curve) whose value peaks at 50 resources and approaches zero as the number of resources approaches either 0 or 100. The perceptual resources of the game are four colors: red, yellow, green, and blue. The critical realist strategy uses the four colors to represent the range of resource values, in order (red 0-24, yellow 25-50, green 51-75, blue 76-100). The strict interface strategy uses the four colors to represent different value ranges, not for the value of resources in the territory, but for the value of the fitness function (red where resource values approach 0 or 100 at the base of the curve, yellow and green for the next higher ranges on either side of the curve, and blue for the large area at the top of the curve where resource values approach 50 from either side).

The critical realist strategy's perceptual function is *monotonic*, because it preserves the order of the linear function describing resource values from 0-100 with an ordered series of colors (red, yellow, green, blue). This makes critical realist perceptions stand in a *homomorphic* relation to states of the world, because they preserve some of its structure (the rough quantity of resources in a territory). By contrast, the strict interface strategy's perceptual function tracks the fitness payoff function, which is *non-monotonic* with the resource quantity. They payoff function goes up until resources hit 50, then goes all the way back down as they approach 100. The colors representing the payoff function as the resource quantity climbs from 0-100 proceed in this order: red, yellow, green, blue, green, yellow, red. This procession of colors does *not* preserve the order of increasing resource quantity, and so the perceived colors cannot be said to veridically represent the structure of the resource quantity.

Hoffman believes that it's very unlikely that fitness functions will vary monotonically with the truth about resource quantities and other real-world measures, for the very reasons I mentioned earlier when motivating *Fitness Beats Truth*. First, monotonic functions—those that vary in an order-preserving way with function measuring a real quantity—are a tiny subset of possible payoff functions. Secondly, biological organisms need homeostasis; there is a right amount of some resource that's fitness-enhancing for us, which means we should seek the right amount, and not too much or too little. This means where resource measures are linear, payoff functions are likely to be Gaussian, as in the above example. Hoffman and his team estimate that fitness functions and payoff functions vary with "unbiased probability one" (Hoffman et. al. 2015: 1486-90; Prakash et. al. 2020; 8-9).

Now, how do the rival strategies perform in the resource game? In general, the strict interface strategy has the advantage because its perceptions are a homomorphic representation of fitness payoffs. All the strict interface player has to do is choose territories that appear as blue wherever they are available, because blue represents the highest level of fitness payoff. On the other hand, while the critical realist player can rule out red or blue as having the lowest fitness payoffs, she has no reason to prefer yellow to green; these are redundant categories in terms of fitness, and neither of them guarantees the highest possible payoff (the player can expect about 62.5 on average for either green or yellow). According to Hoffman, because it is an efficient communication channel for information about payoffs (but not "truth" about resource quantity), the strict interface strategy will dominate the critical realist strategy (and, *a fortiori,* more restrictive realist strategies) in this resource competition and drive it to extinction. The situation only gets worse as the complexity increases, or as the cost of information and computation increases. Introducing dispersion (i.e., noise) into the perceptual channels also disadvantages realist strategies (2015: 1486).

Later experiments (Prakash et. al. 2020) expand on Hoffman's original resource competition games by building a modified version of Bayesian probabilistic estimation into the perceptual models. In these experimental games, each player observes the available territories, estimates the optimal one to choose, and receives the payoffs; then the other player follows suit, with the players taking turns over many rounds. As in the original games, each territory is assigned a resource quantity, and the payoff function is Gaussian. Using a standard Bayesian perceptual model, the "Truth" strategy attempts to estimate the world state that has the highest probability of being the "true" one for each sensory state for a given territory and then chooses among these states the territory with the highest fitness payoff.

The "Fitness-Only" strategy makes no attempt to estimate the true state of the world corresponding to a sensory state, but directly estimates the fitness payoff associated with a given choice. (For each given sensory state, this calculation depends on a posterior probability distribution of possible world states and fitness values corresponding to each world state.) The results of competition between these two strategies are similar to those in the original tests: the "Fitness-Only" strategy drives "Truth" to extinction with high probability. This result is captured in a formal theorem named "The 'Fitness Beats Truth' (FBT) Theorem" (Prakash et. al. 2020).[11]

In a further set of experiments, a fellow proponent of ITP, J.T. Mark (2013), uses genetic algorithms to investigate whether veridical strategies will evolve in the first place. These experiments are based on algorithms introduced by M. Mitchell (1998) where a series of robots are forced to evolve a foraging strategy, over successive generations, to successfully gather soda cans on a map without knowing the map's layout. In Mitchell's experiments, the robots of the first generation are wildly incompetent, while the final generations use sophisticated, highly successful foraging strategies. Mark modifies the algorithm to study the coevolution of foraging and perceptual strategies and gives each square on the map a variable number of cans from 1-10 and assigns a Gaussian fitness function (such that the peak fitness payoff is 10 for squares with 5 cans, and tapers to 0 for squares with either 0 or 10 cans). The robots randomly represent squares with each number with either red or green, but these representations are genetic and can evolve over time. After 500 generations, the robots are much more skilled foragers. Hoffman notes that the

[11] "Over all possible fitness functions and a priori measures, the probability that the Fitness-only perceptual strategy strictly dominates the Truth strategy is at least $(X - 3)/(X - 1)$, where X is the size of the perceptual space. As this size increases, this probability becomes arbitrarily close to 1: in the limit, Fitness-only will generically strictly dominate Truth, so driving the latter to extinction" (Prakash et. al. 2020: 17).

perceptual strategies of these latter robots are fitness-tuned rather than truth-tuned. Here a veridical strategy would use a monotonic (order-preserving) function to relate the two colors and the 11 quantities, perhaps representing 0-5 with red and 6-10 with green. However, the highly evolved foragers use one of two strategies: they either represent 0,1,9, and 10 with red and all the middle quantities with green, or they represent 0,1,9, and 10 with green and all the rest with red. For Hoffman, these experiments show that veridical perceptions are unlikely even to evolve; they show up in the gene pool only by random chance, and no evolutionary pressures seem to favor them (Hoffman et. al. 2015: 1487-88).

Hoffman concludes from the experimental data that our perceptions of the world are almost certainly not veridical, because they are almost certainly of the "strict interface" variety (2015: 1489). These games were designed to test the realist strategies making the weakest claims, e.g., the critical realist strategy only claims to preserve world structures but gives no guarantee that our individual perceptions will match states of the world. *A fortiori*, realist strategies that make stronger claims about veridicality should fare even more poorly against strict interface strategies. The strength of strict interface strategies is that they lack the constraint of representing the world veridically and are free to track fitness payoffs with maximum flexibility and precision.

Hoffman's case for *Fitness Beats Truth* is built upon experimental evidence from evolutionary game theory, which draws its principles and content from evolutionary biology, among them natural selection, heredity, genetic variation, reproductive fitness, competition for scarce resources, and specific traits of living organisms like perception. If the modeling is successful, it illuminates aspects of the science in which it functions as a model. *Fitness Beats Truth* purports to do so by using evolutionary principles to model and make predictions

about specific subject matter from the theory: it makes predictions about how perception must evolve. Indeed, Hoffman characterizes his experiments as applying reasoning from evolutionary biology to cognitive traits like perception (Hoffman et. al. 2015). His point is clear: if we accept that evolution is true, we should accept *Fitness Beats Truth*.

Let us capture this whole discussion in premise (4), the claim that if *Evolutionary Biology* is true then *Human Perception* and *Fitness Beats Truth* are true:

4. If *Evolutionary Biology* is true, then *Human Perception* and *Fitness Beats Truth* are true.

Now it's a simple matter to show that *Human Perception* and *Fitness Beats Truth* entail ITP. If our own trait of perception has been molded according to the requirements of overall reproductive fitness in our ancestral environments, and these requirements will cause strict interface strategies to drive rival realist interface strategies to extinction with high probability, then it is almost certain that *our* perceptions form a strict interface with the world. This is the last premise in the argument:

5. If *Human Perception* and *Fitness Beats Truth* are true, then ITP is true.

From here, (3), (4), and (5) entail ITP:

6. ITP*:* So, it is almost certain that our perceptions constitute a strict interface (i.e., a *non-realist* interface) with the world. (That is, it is almost certain that our perceptual systems evolved to achieve fitness payoffs without representing the world veridically.)

For reference, here is the whole argument for ITP:

1. *Evolutionary Biology*: All species of living things on earth (including our own species, Homo sapiens), have evolved primarily by the process of natural selection, and the traits of organisms from all species have primarily been molded

according to the requirements of overall reproductive fitness in their ancestral environments.

2. *Human Perception*: So, our own trait of perception has been molded according to the requirements of overall reproductive fitness in our ancestral environments.
3. *Fitness Beats Truth*: In the evolution of perception, it is highly probable that the requirements of overall reproductive fitness will cause strict interface strategies to drive rival realist interface strategies to extinction, *or* that those same requirements will cause realist interface strategies to not evolve in the first place.
4. If *Evolutionary Biology* is true, then *Human Perception* and *Fitness Beats Truth* are true.
5. If *Human Perception* and *Fitness Beats Truth* are true, then ITP is true.
6. ITP*:* So, it is almost certain that our perceptions constitute a strict interface (i.e., a *non-realist* interface) with the world. That is, it is almost certain that our perceptual systems evolved to achieve fitness payoffs without representing the world veridically.

We can see here that (4) and (5) entail that *Evolutionary Biology* $\rightarrow$ ITP. In the next section, I will give the debunking argument from ITP to perceptual skepticism.

2.2 The Argument from ITP to Perceptual Skepticism

In this section I will present the argument from ITP to Perceptual Skepticism, and further develop Hoffman's interface metaphor to make the results more intuitive. Here is the argument:

6. ITP: It is almost certain that our perceptions constitute a strict interface (i.e., a *non-realist* interface) with the world. That is, it is almost certain that our perceptual systems evolved to achieve fitness payoffs without representing the world veridically.

7. If ITP is true, then it would be a massive coincidence at best if our perceptual representations were veridical, and we have no reason to believe such a coincidence has occurred in our case.

8. *Perceptual Skepticism*: So, we should not think that our perceptual representations are veridical.

Premise (7) captures the implication of ITP that the interface strategies that perform the best in terms of achieving fitness payoffs are not the kind that are constrained to do so by means of representing the world veridically. Rather, if they deliver veridical representations, it is by sheer coincidence, because veridicality does not seem to be *relevant* to achieving fitness payoffs. Moreover, we have no reason to believe that there *has* been such a coincidence in our case. The reasoning behind *Fitness Beats Truth* and ITP applies to perception across the board. If it is sound reasoning and applies to other organisms, we would need a principled reason for believing it doesn't apply to our particular case. Moreover, the evolution of our perceptual systems is continuous with other organisms, and there is no empirical evidence that the basic functioning of our senses is fundamentally different from those of our ancestors.[12] Even our significantly larger brains, for instance, just pack in more neurons than those of our ancestors; they are an example of a difference in degree, not in kind.

It follows that we should not believe the deliverances of our perceptions. According to Hoffman, this means that we should not believe that the objects and properties we seem to see in the world are really there; the same goes for the apparent structural and causal relations among such objects in spacetime, and even for spacetime itself. These, too, are just elements of our perceptual interface, which is tuned to fitness rather than truth.

At this point, one might reasonably ask: how can a perceptual system guide us successfully in the world if all it gives us are non-veridical perceptions? Here we return to Hoffman's central metaphor. Our perceptions are related to the world as the icons on a Windows desktop are related to the underlying hardware. My folder containing all my

[12] See, for example, Jon Kaas (2010: 1127-30) for a concise history of the evolution of the visual system.

information for an upcoming course is yellow and rectangular. However, what I'm doing when I click on the coursework folder and navigate the subfolders is control electrical currents through a series of transistors on multiple circuit boards, making rapid changes to the physical states of flash memory chips. Nothing in this lower-level situation is yellow or rectangular, and when I move the icon to a more convenient spot on my desktop, I haven't moved the underlying hardware's location. The higher-level situation with icons is not a veridical representation of the underlying hardware situation or its properties. Of course, it does not purport to represent the underlying hardware situation accurately (and certainly not as yellow or rectangular), and we don't take it as such. Hoffman's point is that within our own perceptual interfaces, we *only* have access to the icons and desktop, not to the underlying hardware. So, we should not believe that our own representations of objects, spacetime, and causality are literal representations of what's out there in the world.

Nonetheless, in the Windows desktop situation it is most practical for me to use the icons and not bother with the lower-level situation. I can use it to pull up my information and show my students slides and use my notes to teach. Moreover, I would never carelessly drag the icon containing my course information to the Recycle Bin icon and empty it. I take the icons seriously, but not literally. That is, I do all of this knowing that the underlying hardware situation is very different, and that there is no actual *folder* on my Windows desktop (nor is there any actual desktop). These icons usefully hide the complexity of the underlying hardware and show me only what I need to accomplish important tasks. Trying to accomplish my teaching tasks by directly engaging with the lower-level hardware would be nearly impossible. Even a brilliant computer engineer with full knowledge of the circuitry

would find it prohibitively difficult to edit a PowerPoint slideshow by tweaking individual transistors.

Let's say I have a visual experience as of there being a red tomato in front of me. I can seem to reach out and touch the tomato, feel its weight in my hand, and thump it on the counter to hear the dull *thud* of its body. I can seem to smell and taste it. According to Hoffman, there is no real, mind-independent object with the properties of a tomato out in the world where it seems to be in space and time. Rather, the tomato representation is an icon of my species-specific perceptual interface. Instead of taking it to be an object in spacetime, I should see it as a packet of fitness information. The icon of the tomato informs me of potential fitness-enhancing actions I can take within the interface, such as *eating the tomato* (which I'm not to take as *literally* eating a tomato). According to Hoffman, my brain is not perceiving a tomato but *constructing* a tomato icon, much as my computer's CPU is constantly constructing the images that make up my Windows desktop. When, within the interface, I look away from the tomato icon, my brain stops constructing it, much as my computer stops generating icons of my teaching slides when I navigate away from them to check my email (Hoffman 2019: 17-21).

For Hoffman, the idea of causality is just another useful fiction of our interface. The relationship between the icons on your desktop is real, but when I drag a file to the recycle bin and empty the bin, I haven't caused an actual folder to go into an actual bin. Rather, I have effected some useful change in stored information that involves neither physical files and folders nor bins. Likewise, when within my own perceptual interface, I manipulate apparent objects in spacetime, I should not regard the causal relationships between objects in my interface as corresponding to *actual* causal relationships in the world. Rather, I should

regard spacetime as a communications channel and physical objects as messages about fitness. The notion of causality simply helps me predict the consequences of my actions so I can better utilize these messages (Hoffman 2019: 123-4).

In this section, I have given the argument from *Evolutionary Biology* to ITP, and from ITP to *Perceptual Skepticism*. ITP is a radical and counterintuitive view—a point of which Hoffman and his colleagues are well aware. In the next section, I will raise a self-defeat problem for ITP, and show that Hoffman's response is inadequate to save the theory.

3. Objection that ITP is Self-Undermining

ITP appeals to an empirical thesis to make its evolutionary case for skepticism: the *Evolutionary Biology* thesis presupposes that organisms exist and that their environments are made up of living and nonliving objects.[13] Indeed, Hoffman's empirical examples of jewel beetles, frogs, etc. involve organisms interacting with their respective object-filled environments. In addition, the game theoretical experiments model—and are explained and motivated by—the subject matter and general principles of the empirical science of evolutionary biology. But if ITP is true, then if we accept the plausible premise (7) we must accept *Perceptual Skepticism.* This means that observations made through the perceptual interfaces of human scientists are almost certainly false, raising the problem that ITP is self-undermining because it seems we no longer have any reason to believe the *Evolutionary Biology* thesis. If so, then the *Human Perception* thesis and *Fitness Beats Truth* are no longer

[13] In addition, if we are justified in believing that DNA molecules, cells, and other explanatorily crucial microscopica exist, we must be justified in believing we've observed them through microscopes, and so *a fortiori* we are justified in believing that there are microscopes.

supported, and this leaves ITP itself unjustified. Hoffman appears to be in the following unfortunate dialectical situation:

(D1) If *Evolutionary Biology*, then ITP.

(D2) If ITP and (7), then *Evolutionary Biology* is unjustified (so, ITP is unjustified).

(D3) So, either (7) is false, or ITP is unjustified.

(D4) So, *Perceptual Skepticism* is unjustified.

The argument for (D1) was given in Subsection 2.1 above, and I have just given reasons to support (D2). The idea behind (D3) is that if (D2) is true, ITP is unjustified. But also, if (D1) is false and *Evolutionary Biology* really doesn't entail ITP, then ITP is also left unjustified. Alternately, *Evolutionary Biology* might be justified and (D1) might be true, meaning ITP is also justified, but only if premise (7) is false. This last would only be possible if there really were a reason to think that in the case of our perceptions alone, we have reason to believe that tracking fitness payoffs is best accomplished by veridically representing the world. First, if we are granting all of the premises through ITP, then it really would be a *very* unlikely coincidence that our perceptions happen to be veridical. So, we'd need to have very good reasons to think we're an exception. Second, this would run counter to the motivation for Hoffman's whole argument, which holds that the results of his game theoretical experiments apply to all evolved perceivers and that it would be a strike against a theory if humans were discontinuous with the rest of evolved life. But finally, as captured in (D4), this would still mean that *Perceptual Skepticism* is not justified, and this is one of the most salient and important consequences of ITP.

Hoffman is aware of this potential problem and proposes a solution. He straightforwardly accepts the consequence of ITP that the evidential claims supporting

evolutionary biology—which, after all, appeal constantly to our perceptual object beliefs—are almost certainly false, leaving *Evolutionary Biology* unjustified. However, he argues that when it comes to supporting ITP, the heavy lifting is done not by the empirical science of evolutionary biology but by Universal Darwinism, the idea that *if* there are replicators *anywhere* with the needed kind of heredity, variation, and selective pressures, these replicators will evolve according to the process of natural selection (2019: 64-6). Because it does not presuppose a physical world of spacetime replete with objects and their causal interactions, Hoffman believes there is no tension between Universal Darwinism and *Perceptual Skepticism*. The latter does not undermine one's justification for the former. So, in response to the self-undermining charge above, Hoffman makes the following revised claim:

(D1*) If *Universal Darwinism*, then ITP.

In the rest of this section, I will show that (D1*) is false.

First, let me clarify what I'm taking Hoffman to mean by Universal Darwinism. Universal Darwinism is sometimes taken as a rather conservative expansion of evolutionary principles to include situations markedly outside the realm of known terrestrial life. For instance, it is reasonable to expect any new life we discover on earth or on other planets to have evolved primarily by natural selection, even if the replicators that are the medium for hereditary transmission are radically different from DNA and RNA. More controversially, it is sometimes argued that units of culture (for example, words, stories, catchy songs, or even replicating chunks of code within a computer simulation) fit the requirements for Darwinian evolution. On this stronger version of the view, Universal Darwinism is taken to be a kind of naturally occurring algorithm that operates in any population of replicators that meets the

requirements of natural selection: it is *substrate independent*.[14] Hoffman supports this latter, more radical version of Universal Darwinism. It will help us to cast this as a premise formulated as closely as possible to *Evolutionary Biology*:

> *Universal Darwinism:* Wherever conditions for evolution by natural selection exist—i.e., wherever the right sort of heredity, variation, and selective pressures exist among replicating beings in whatever medium—evolution by natural selection will occur among these beings, and the traits of beings from all species so evolved will primarily be molded according to the requirements of overall reproductive fitness in their ancestral environments.

Hoffman holds that ITP can draw support from *Universal Darwinism* and address the question of whether evolution favors true perceptions, without recourse to the empirical science of evolutionary biology. This is made possible by *Universal Darwinism*'s substrate-independence: it does not depend on or presuppose the existence of a physical world populated by objects causally interacting in spacetime. According to Hoffman, here is how the landscape of evolutionary theory looks after one accepts ITP:

> Still recognizable… are the landmarks of universal Darwinism: variation, selection, and heredity. But gone from objective reality are physical objects in spacetime, including those central to biology: DNA, RNA, chromosomes, organisms, and resources. This doesn't entail solipsism. *Something* is there in objective reality, and we humans experience its import for our fitness in terms of DNA, RNA, chromosomes, organisms, and resources. But the FBT Theorem tells us that, whatever that something is, it is almost surely not DNA, RNA, chromosomes, organisms, or resources. It tells us that there is good reason to believe that the things that we perceive, such as DNA and RNA, don't exist independent of our minds. The reason is that the structures of fitness payoffs, which shape what we perceive, differ from the structures of objective reality *with high probability* (2019: 65).

Here the "FBT Theorem" is the "Fitness Beats Truth Theorem" (see Prakash et. al. 2020: 17), which I've paraphrased in Section 2 as part of the more general *Fitness Beats Truth* thesis.[15]

[14] See, for instance, Dennett (1995: 48-60).
[15] See note 10 for a full statement of the theorem.

If we accept ITP, we should regard the science of evolutionary biology as literally false, and we should not believe that DNA, organisms, or other physical objects relied on by the theory exist. We shouldn't believe that the physical relationships between these objects exist, or that the patterns in their behavior tell us something about the world's real structure. We should take the physical sciences generally as incredibly well-tested, detailed, sophisticated information about our own fitness—not about the real world.

The abstract algorithm of *Universal Darwinism* does not declare that natural selection *does* occur or in which substrates the conditions for it exist. This would take away from the substrate-neutrality that makes *Universal Darwinism* immune from self-undermining. This raises a difficult question for proponents of ITP: why should you and I believe *Universal Darwinism* applies to *us*? This depends on ITP*'s* answer to a more fundamental question: is *Evolutionary Biology* an instance of *Universal Darwinism*, or not?

Herein lies a dilemma for ITP. Either evolutionary biology is an instance of *Universal Darwinism*, or it is not. If it is not, then ITP is unjustified because the *Human Perception* thesis has been undermined. If it is, then ITP is wrong in its most important prediction (that our species has evolved a *non-realist* interface that delivers non-veridical perceptions), and premise (7) must be false, leaving *Perceptual Skepticism* unjustified. I will spend the rest of this section elucidating the two horns of the dilemma.

3.1 First Horn of the Dilemma

Hoffman seems certain to take the first horn and deny that evolutionary biology is an instance of *Universal Darwinism*.[16] This is most consistent with ITP*'s* central motivations and results. We are to take *Evolutionary Biology* as a thesis about the contents of our particular user interface that provides fitness information, but not information about the world. But according to ITP, *Universal Darwinism* is true—*really* true. It's true of *whatever* is out there making up the real world, regardless of how it appears to anyone through their interface. A mere relation of interface icons cannot be an instance of *Universal Darwinism* any more than a cartoon of Wyle E. Coyote plummeting off of a cliff can be an instance of Newton's law of universal gravitation.[17] Taken as a literal claim about the world, it's false that Wyle E. Coyote plummets off of the cliff, and a false proposition cannot instantiate some true proposition of law.

However, to say that the science of evolutionary biology is *not* an instance of *Universal Darwinism* carries severe drawbacks. It cuts off *Universal Darwinism* from empirical biology's robust experimental support and explanatory resources. In fact, it utterly transforms the relationship between the two theories: what was once an elegantly

[16] Hoffman's own position is not entirely clear. For instance, when scientists like Richard Dawkins and philosophers like Daniel Dennett draw inferences from evolutionary biology to Universal Darwinism, they seem to be doing so on the assumption that the former is an instance of the latter—and Hoffman is eager to enlist their support in justifying such a move. See Dawkins (1983) and Dennett (1995) for these more conventional characterizations of the relation between evolutionary biology and Universal Darwinism. Moreover, Hoffman also invokes empirical evidence from evolutionary biology as if it is an instance of Universal Darwinism. Of course, in other places Hoffman explicitly states that the claims made by our empirical scientific theories are false. See, e.g., Hoffman (2019: 65) and Hoffman et. al. (2015: 1501). However, see (2015: 1489) for an explicit appeal to empirical evolutionary biology.

[17] By stipulating a "genuine" instance, I also mean to rule out as potential instances those observations that are true instances of a universal generalization simply because they make the antecedent of the conditional false, in the way that most observations to date have confirmed my hypothesis that all goblins wear little green hats. (Most observations have been of non-green-hat-wearers that were not goblins, confirming the logically equivalent hypothesis that all non-green-hat-wearers are not goblins.)

explanatory, copiously detailed, nuanced supporting relationship between the empirical science of evolutionary biology and the theory of *Universal Darwinism* is now a staggering coincidence between *Universal Darwinism* and the intricate but non-veridical world of one's interface icons.

This opens up two thorny questions for ITP. First, what is justifying *Universal Darwinism*? Second, and more importantly, if evolutionary biology is not an instance of *Universal Darwinism*, why should we think *Universal Darwinism* applies to us?

Let us start with the first question: what justifies *Universal Darwinism*? All of the empirical reasons for believing *Universal Darwinism* stand or fall with *Evolutionary Biology.* Even the fact that empirical biology makes incredibly accurate predictions about what will happen within our interface carries no justificatory weight if we have no theory of how our interface relates to the real world, and it's crucial to ITP*'s* radical claims that we don't have such a theory.

However, there are reasons to believe Hoffman thinks *Universal Darwinism* and ITP are *a priori* justified. This is implicit in his claim that *Universal Darwinism* survives self-undermining because UD is a substrate-independent algorithm that doesn't rest on empirical evidence (Hoffman 2019: 89-91). He also stresses that ITP merely shows us that our *perceptual* faculties are unreliable; it does not cast doubt on our faculties of mathematical and logical reasoning (Hoffman et. al. 2015: 1500). This suggests that Hoffman believes logic and math alone suffice to establish *Universal Darwinism* and ITP.[18] For these reasons, I

[18] However, Hoffman also seems to contradict (2) at key points. Recall that he leaves open the possibility that we *can* reason from the contents of our interfaces to those of the actual world (2019: 124) and that he takes the findings of ITP to contribute to empirical research in visual psychology (2009: §1.8). Moreover, he explicitly

will proceed on the assumption that *Universal Darwinism* is to be justified *a priori.*[19] Moreover, for the purposes of the present argument, I will assume that it *can* be justified *a priori.*

I now address the second question. Even if I believe that *Universal Darwinism* is true, if I don't believe it applies to me, I lose any reason to believe that I am an evolved being whose traits like perception have evolved by natural selection. The *Evolutionary Biology* thesis, along with my own observations of my physical body and environment, gives me reason to believe I'm a certain kind of physical organism with evolved traits, and thus that the *Human Perception* thesis is true. Without *Evolutionary Biology*, I have no reason to believe the *Human Perception* thesis. In fact, I have no reason to believe my conscious experiences arise from brain processes in an evolved, physical body.[20] I'm no longer in a position to believe that my apparent perceptions of the world are the deliverances of an evolved trait that would be subject to *Universal Darwinism.*

It's possible that my consciousness and perceptual faculties are non-physical.[21] But I don't know anything about *non*-physical processes and laws—if there be such—or how they apply to me, or if they indeed satisfy the conditions for *Universal Darwinism*. Nor does

states that ITP makes testable predictions, at least in principle, about whether or not there are physical objects when no one's looking (Ibid, Ch.6). I shall take him to mean that these are in-principle possible inferences and empirical predictions, but ones that are exceedingly remote. Anything stronger would undercut ITP*'s* radical conclusions.

[19] It's possible that we could have learned Universal Darwinism from a *false* scientific theory. When, as a child, one reads a story about fictional bears counting apples and learns that 2+2=4, one learns a truth even though the story is literally false. Perhaps Universal Darwinism is *a priori* justified, and evolutionary biology is just a set of experiences that enable us to learn it. This is Kant's qualification that we can learn *a priori* truths through enabling empirical experiences. See (Kant 1787 [1965: 43(B3)]).

[20] Indeed, Hoffman presses the point that we ultimately shouldn't believe our conscious processes arise from neural activity at all (2019: Ch.3).

[21] Ultimately, Hoffman speculates that we are non-physical conscious agents defined in terms of math and set theory. See Hoffman (2019: Ch.10 and Appendix).

Universal Darwinism by itself entail anything about the nature of consciousness. Even assuming I am a perceiver of *some* kind (and not just, for example, a being having a long and vivid dream), I have no reason to believe I'm the kind of *evolved* perceiver that would be covered by the predictions of *Fitness Beats Truth*.

To convince us to accept that Universal Darwinism applies to us, Hoffman might appeal to the possibility that we are some other kind of life besides what is described in evolutionary biology. For instance, *Universal Darwinism* applies in principle to "memes"—bits of remembered and transmitted mental content within, e.g., human culture. Languages and their elements, works of art, and whole scientific theories are said to evolve along at least quasi-Darwinian lines. Perhaps we are memes, or clusters of memes? Plausibly, memes would fit well into the ITP picture because abstract ideas like words and numbers have systems of internal relations and, at least on certain understandings, it is intuitive to say that like other abstracta they are causally inert and don't have a location in spacetime. But there are substantial problems with both motivating this view and reconciling it with ITP.

One problem is that we need reasons to accept this over a physicalist view, on which it is plausible that memes can *only* exist in a world with physical objects causally interacting in spacetime—a world very much like the one represented in our perceptions. In fact, the examples appealed to by universal Darwinists like Dawkins and Dennett are things that occur within the physical substratum of humans and their thoughts, social interactions, and information technology. On this understanding, memes are things that cannot replicate without organic brains or physical technologies like television, radio, and networked computers. Memes are implemented both neurally and digitally, but they are subjected to physical processes of change over time. They are real patterns of neural wirings in human

brains, or patterns of code on microchips. Natural selection is a physical process that includes selection by organisms who (wittingly or unwittingly) use the memes and pass them on via physical media.

Perhaps there is some substrate that is entirely non-physical in the sense the ITP requires—being outside of space, time, and physical objects as we understand them—and which would permit natural selection along Darwinian lines. But this would take a great deal of explaining from Hoffman. For one thing, even if Darwinian processes somehow work outside of space, it's unclear how *any* ordered, algorithmic process could work outside of time. One step must follow another, and in the correct order, to achieve the right kind of change. Also, the concept of *change* doesn't make sense outside of time. For another, on the conception that abstracta like numbers, sets, and algorithms exist outside of space and time, they are in some sense unchanging and causally inert. We use a Darwinian algorithm to explain why we observe the changes we observe in the physical world, but algorithms by themselves don't *do* anything, just as recipes that languish in the pages of an unopened cookbook don't do anything. Darwinian algorithms were formulated to explain processes that are physical and causal in nature—perhaps fundamentally so. That the algorithm of *Universal Darwinism* will hold within some wholly alien substrate outside of spacetime, physical objects, and causality, is not something we have any reason to accept without argument. Of course, even given a satisfactory notion of how *Universal Darwinism* works in such a world, we would still need a premise like *Human Perception* to connect us to the evolutionary picture.

3.2 Second Horn of the Dilemma

On the second horn, Hoffman can claim that *Evolutionary Biology* is an instance of *Universal Darwinism*. This would provide the missing link between *Universal Darwinism* and the *Human Perception* thesis. I would then have reason to believe that I'm an evolved being because I know that I'm exactly the kind of thing evolutionary biology explains: I'm a member of *Homo sapiens,* an evolved perceiver. Because *Evolutionary Biology* is an instance of *Universal Darwinism,* I know that I'm covered by its laws, and subject to the limitations of evolved perception as expressed by *Fitness Beats Truth*. Thus, I accept ITP.

However, the drawbacks of choosing the second horn are substantial. If we take *Evolutionary Biology* to be an *instance* of *Universal Darwinism,* this means we take *Evolutionary Biology* to be a true thesis: literally true of the real world, not merely interface-limited information about our own fitness. Since *Evolutionary Biology* is an empirical thesis grounded in empirical observation, and we are to believe it is true, this means our perceptions of organisms and their environments have been veridical and the causal relationships we've observed in nature have been real relationships that we've captured in our laws and explanations. This means a great many of our perceptions of the world are veridical: at the very least, they reveal much of its real structure.[22] But ITP predicts that this is almost certainly not the case. This by itself gives us strong reasons to doubt ITP.

However, ITP by itself does not entail *Perceptual Skepticism.* It is possible for ITP and *Evolutionary Biology* to be both true and justified, as long as premise (7) is false and *Perceptual Skepticism* is unjustified. Let's assume for the moment that we *do* accept both

[22] That is, *critical realism* must be true at the least. Put in Hoffman's terms, this as a situation where we have to make little or no correction when drawing inferences about the real world from what goes on in our interface.

ITP and premise (7), and thus accept *Perceptual Skepticism.* While it is possible for *Evolutionary Biology* and *Perceptual Skepticism* to both be true, it is not possible to rationally accept them both if our only grounds for accepting *Evolutionary Biology* are empirical observations, which are the very kind of thing undermined by *Perceptual Skepticism.* On the other hand, if we *do* have some justification for accepting *Evolutionary Biology*—one that isn't a completely *ad hoc* means of escaping this dilemma—we must have a reason for thinking that, despite ITP being true, human perception is the rare case where fitness payoffs *were* achieved by means of veridical perceptions. So, it seems that in that case, (7) is false, and we have no justification for *Perceptual Skepticism.*

Hoffman could try to stake out some middle ground and claim that *Evolutionary Biology* is something like a *quasi*-instance of *Universal Darwinism*. Perhaps it is an instance only inasmuch as it reflects a kind of real non-physical pattern consisting of, for example, sets, numbers, algorithms, and some basic logic. This would seem to keep the view clear of commitments to perceptual beliefs in physical objects, spacetime, and causality. However, if ITP and *Perceptual Skepticism* are true, then we are only to take patterns and structures in our interface as information about fitness, not about the state of the real world. To suggest that we take these as *direct* evidence about the real world—even the most abstract patterns having to do with math and logic, and even if they happen to match our *a priori* reasoning—would be an illicit appeal to critical realism, the notion that our perceptions reveal some of the world's real structure.

According to Hoffman, it is at least theoretically possible to take patterns and structures in our interface as a kind of *indirect* evidence about the real world, but for now this is merely an extremely remote possibility. We could approach this in one of two ways: we

could formulate and test a theory about our interface and use it to ask what we can or can't infer from the structures in the interface, or we could formulate and test theories of objective reality and how it appears in our interface. To rescue Hoffman's view from the second horn of the dilemma, we would need such a theory in hand, and one powerful enough to justify *Evolutionary Biology*, or to otherwise get us from *Universal Darwinism* to *Human Perception* or some nearly equivalent premise in order to justify ITP. However, having a theory that lets us reason *indirectly* from our perceptions to states of the world may force us to revise or reject *Perceptual Skepticism*, depending on just how far our perceptual beliefs are from the way the world really is. And we do *not* currently have such a theory in hand, and for Hoffman, until we do, we are to regard the elements of our interface as *not* corresponding to the structure of the world (2019: 124).

In this section, I have shown that *Universal Darwinism* does not entail ITP. This is because *Universal Darwinism* alone does not entail the *Human Perception* thesis, and thus we have no reason to believe the constraints on perception specified in *Fitness Beats Truth* apply to us. To justify *Human Perception, Universal Darwinism* must have recourse to *Evolutionary Biology* as an instance of the theory. But this would mean taking *Evolutionary Biology* to be true. However, this move is entirely unmotivated within ITP, and we cannot rationally accept both *Evolutionary Biology* and *Perceptual Skepticism.*

Conclusion

If my arguments have been successful, I have shown that the evolutionary argument for ITP and *Perceptual Skepticism* faces a serious self-undermining problem rooted in the relationship between *Evolutionary Biology* and perceptual debunking arguments, which is not resolvable by appeal to a substrate neutral *Universal Darwinism.* In a broader sense, I

hope to have contributed to the literature on evolutionary debunking arguments about perceptual beliefs in ordinary objects. My aim was to show why certain dialectical moves are forbidden to debunkers, and in doing so to help reframe the debate in a more productive way.

There is also a broader problem, which is outside the scope of this chapter, about the relationship between scientific models and the larger scientific theories they depend on and inform. Scientific models gain their strength and explanatory power by being embedded in broader empirical theories, and those in turn benefit from the rigor, focus, and enhanced visualization of the former. Because it was unnecessary for my argument, I did not press the case that *Fitness Beats Truth* itself may be unjustified after *Evolutionary Biology* is removed and replaced by *Universal Darwinism.* As I suggested above, Hoffman seems to sometimes assume his entire argument is *a priori* justified.[23] However, one could certainly ask how we are to interpret his models of perception, replicators, resources, as well as his appeals to typical biological properties such as homeostasis, once we have removed physical objects like organisms from the larger theory of evolution. These and other questions will have to be the subjects of future work.

[23] See Hoffman et. al. (2015: 1500).

References

Aizawa, Kenneth, (2008). Neuroscience and Multiple Realization: A Reply to Bechtel and Mundale. *Synthese*, *167*(3), 495–510.

Aizawa, Kenneth and Carl Gillett (2011). The Autonomy of Psychology in the Age of Neuroscience. In P.M. Illari, F. Russo, and J. Williamson (Eds.), *Causality in the Sciences* (203–23). Oxford University Press.

Axelrod, Robert, and Hamilton, William D. (1981). The Evolution of Cooperation. *Science*, 211: 1390–96.

Bagwell, Jeffrey N. (2021). Eliminativism and Evolutionary Debunking. *Ergo, 8*(17), 496-522. doi: https://doi.org/10.3998/ergo.1154

Barker, Jonathan, (2019). Debunking Arguments and Metaphysical Laws. *Philosophical Studies*, *177*(7), 1829–55.

Barnard, Christopher J., and Richard M. Sibly (1981). Producers and scroungers: A general model and its application to captive flocks of house sparrows. *Animal Behaviour*, *29*(2), 543–550.

Bennett, Karen (2009). Composition, Colocation, and Metaontology. In D. Chalmers, D. Manley, and R. Wasserman (Eds.), *Metametaphysics* (38–76). Clarendon.

Benovsky, Jiri (2015). From Experience to Metaphysics: On Experience-Based Intuitions and their Role in Metaphysics. *Noûs*, *49*(4), 684–97.

Boghossian, Paul A. (1997). What the Externalist Can Know A Priori. *Proceedings of the Aristotelian Society*, *97*, 161-175.

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

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

Brenner, Andrew (2021). Mereology and ideology. *Synthese*, 198, 7431–7448.

Brigandt, Ingo and Alan Love (2017). Reductionism in Biology. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosoph*y (Spring 2017 Ed.). Retrieved from https://plato.stanford.edu/archives/spr2017/entries/reduction-biology/

Brown, Jessica (1995). The Incompatibility of Anti-Individualism and Privileged Access. *Analysis* 55(3), 149-156.

Brueckner, Anthony (1992). What an Anti-Individualist Knows A Priori. *Analysis, 52*(2), 111-118.

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

Calosi, Claudio (2014). Quantum mechanics and priority monism. *Synthese, 191*, 915–928.

Calosi, Claudio, Vincenzo Fano, and Gino Tarozzi (2011). Quantum ontology and extensional mereology. *Foundations of Physics, 41*, 1740–1755.

Chakravartty, Anjan, "Scientific Realism", *The Stanford Encyclopedia of Philosophy* (Summer 2017 Ed.), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2017/entries/scientific-realism/>.

Cohen, Jonathan (2015). Perceptual Representation, Veridicality, and the Interface Theory of Perception. *Psychonomic Bulletin and Review*, 22(6): 1512-1518.

Colyvan, Mark (1999). Contrastive Empiricism and Indispensability. *Erkenntnis, 51***,** 323–332.

Colyvan, Mark (2001), *The Indispensability of Mathematics.* New York: Oxford University Press.

Colyvan, Mark, "Indispensability Arguments in the Philosophy of Mathematics", *The Stanford Encyclopedia of Philosophy* (Spring 2019 Ed.), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2019/entries/mathphil-indis/>.

Contessa, Gabriele (2014). One's A Crowd: Mereological Eliminativism Without Ordinary-Object Eliminativism. *Analytic Philosophy*, *55*(2), 199–221.

Daly, Chris and David Liggins (2010). In Defense of Error Theory. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *149*(2), 209-230.

Dawkins, Richard (1976). *The Selfish Gene.* Oxford: Oxford University Press.

Dawkins, Richard (1983). Universal Darwinism. In: *Evolution from Molecules to Man.* Bendall, D. S., Ed. Cambridge University Press. 403-425.

Dennett, Daniel (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life,* New York: Simon and Shuster.

Devitt, Michael and Kim Sterelny (1999). *Language and Reality: An Introduction to the Philosophy of Language,* 2nd Ed., Oxford: Blackwell Publishers.

Elder, Crawford (2011). *Familiar objects and their shadows*. Cambridge: Cambridge University Press.

Fairchild, Maegan and John Hawthorne (2018). "Against Conservatism in Metaphysics," *Royal Institute of Philosophy Supplement*, *82,* 45–75.

Feldman, Richard (2003). *Epistemology.* New Jersey: Prentice Hall.

Field, Hartry H. (1980), *Science Without Numbers: A Defence of Nominalism*, Oxford: Blackwell.

Foley, Richard (1983). Epistemic Conservatism. *Philosophical Studies*, *43*(2), 165-182

Fodor, Jerry A. (1974). Special Sciences (Or: The Disunity of Science as a Working Hypothesis). *Synthese*, *28*(2), 97–115.

Gillett, Carl (2003). The Metaphysics of Realization, Multiple Realization, and the Special Sciences. *Journal of Philosophy*, *100*(1), 591–603.

Gillett, Carl (2013). Constitution, and multiple constitution, in the sciences: Using the neuron to construct a starting framework. *Minds and Machines*, *23*, 309–337.

Graziani, Pierluigi, and Claudio Calosi, Eds. (2014). *Mereology and the sciences: Parts and wholes in the contemporary scientific context*. Berlin: Springer.

Gwynne, Darryl T., and David C. F. Rentz (1983). Beetles on the Bottle: Male Buprestids Mistake Stubbies for Females (Coleoptera). *Australian Journal of Entomology*, 22: 79-80. https://doi.org/10.1111/j.1440-6055.1983.tb01846.x

Harper, George R., and David W. Pfennig (2007). Mimicry on the edge: why do mimics vary in resemblance to their model in different parts of their geographical range? *Proceedings of the Royal Society B.*, *274*(1621): 1955–1961.

Hawkeswood, Trevor J. (2005). Review of the biology and host-plants of the Australian jewel beetle Julodimorpha bakewelli (White, 1859) (Coleoptera: Buprestidae). *Calodema*: 3: 3-5.

Hawley, Katherine (2004). Borderline Simple or Extremely Simple. *The Monist*, *87*(3), 385-404.

Healey, Richard (2013). Physical composition. *Studies in History and Philosophy of Modern Physics*, *44*, 48–62.

Healey, Richard and Jos Uffink (2013). Part and Whole in Physics: An Introduction. *Studies in History and Philosophy of Modern Physics, 44*, 20–21.

Hoffman, Donald D. (2009). The Interface Theory of Perception: Natural Selection Drives True Perception to Swift Extinction. In Dickinson, Sven, Michael Tarr, Ales Leonardis, and Bernt Schiele (Eds.), *Object Categorization: Computer and Human Vision Perspectives* (148-265). Cambridge University Press.

Hoffman, Donald D., Manish Singh, and Chetan Prakash (2015). The Interface Theory of Perception. *Psychonomic Bulletin and Review,* 22: 1480–506.

Hoffman, Donald D. (2019). *The Case Against Reality: Why Evolution Hid the Truth from Our Eyes.* New York: W. W. Norton and Co.

Hofweber, Thomas (2016). *Ontology and the Ambitions of Metaphysics*. New York: Oxford University Press.

Hull, David (1972). Reductionism in Genetics—Biology or Philosophy? *Philosophy of Science*, *39*(4), 491–99.

Hummel, J. E. (2015). Is Interface Theory a Special Case of Q-morphs or Full-blown Solipsism? *Psychonomic Bulletin and Review* (invited commentary). http:// www.psychonomic.org/ news/ 297868/ Is-Interface-Theory-a-Special-Case-of-Q-morphs-or-Full-blown-Solipsism.htm.

Johnston, Mark (1992a). Reasons and reductionism. *The Philosophical Review*, 101, 589-618. doi: 10.2307/ 2186058.

Johnston, Mark (1992b). How to speak of the colors. *Philosophical Studies*, 68, 221-263. doi:10.1007/ BF00694847.

Johnston, Mark (1993). Objectivity reconfigured: Pragmatism without verificationism. In J. Haldane and C. Wright (Eds.), *Realism and reason* (pp. 85-130). Oxford University.

Kaas, Jon (2010). Visual System: Evolution of. In E. Bruce Goldstein (Ed.), *The Encyclopedia of Perception* (1127-30). Sage Publications.

Kaiser, Marie I. (2011). The Limits of Reductionism in the Life Sciences. *History and Philosophy of the Life Sciences*, *33*(4), 453–76.

Kant, Immanuel (1787 [1965]). *Critique of Pure Reason*, Norman Kemp Smith (Trans.), New York: St. Martin's Press, 1965.

Kantin, Holly (2020). Why Compositional Nihilism Dissolves Puzzles. *Synthese, 197*, 4319–4340.

Kimbrough, Steven O. (1978). On the Reduction of Genetics to Molecular Biology. *Philosophy of Science*, *46*(3), 389–406.

Korman, Daniel Z. (2006). What Externalists Should Say About Dry Earth. *Journal of Philosophy*, *103*(10), 503-520.

Korman, Daniel Z. (2009). Eliminativism and the Challenge from Folk Belief. *Noûs*, *43*(2), 242–64.

Korman, Daniel Z. (2014). Debunking Perceptual Beliefs about Ordinary Objects. *Philosophers' Imprint, 14*(13), 1–21.

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

Korman, Daniel Z. (2019a). Debunking Arguments. *Philosophy Compass,* 2019; 14:e12638. https://doi.org/10.1111/phc3.12638

Korman, Daniel Z. (2019b). Debunking Arguments in Metaethics and Metaphysics. In A Goldman and Brian McLaughlin (Eds.), *Metaphysics and Cognitive Science* (1–34). Oxford University Press.

Korman, Daniel Z., and Carmichael (2016). Composition. Oxford Handbooks Online. Online Publication Date: Jun 2016. DOI: 10.1093/oxfordhb/9780199935314.013.9

Kovacs, David M. (2019). How to be an Uncompromising Revisionary Ontologist. *Synthese* (2019). Advance online publication. https://doi.org/10.1007/s11229-019-02196-8

Ladyman, James, Don Ross, James Collier, and David Spurrett, (2007). *Every thing must go: Metaphysics naturalized.* Oxford University Press on Demand.

Lewontin, Richard C. (1970). The units of selection. *Annual Review of Ecology and Systematics*, *1*, 1–18.

Liggins, David (2008). "Nihilism Without Self-Contradiction," *Royal Institute of Philosophy Supplement*, 62: 177-196.

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

Maddy, Penelope (1995). Naturalism and Ontology. *Philosophia Mathematica*, *3*(3): 248–270.

Mark, Justin T. (2013). *Evolutionary pressures on perception: when does natural selection favor truth?* Ph.D. Dissertation, University of California, Irvine.

Markosian (1998). Brutal Composition. *Philosophical Studies*, *92*, 211-249.

Marr, David (1982). *Vision.* San Francisco: Freeman Press.

Martínez, Manolo (2019). Usefulness Drives Representations to Truth: A Family of Counterexamples to Hoffman's Interface Theory of Perception, *Grazer Philosophische Studien* 96(3), 319–341. DOI: https://doi.org/10.1163/18756735-09603004

McLaughlin, Brian and Michael Tye (1998). Externalism, Twin Earth, and Self-Knowledge. In Barry C. Smith and Cynthia Macdonald (Eds.), *Knowing Our Own Minds: Essays on Self-Knowledge* (285-320). Oxford University Press.

Mendelovici, Angela (2010). *Mental Representation and Closely Conflated Topics*, PhD thesis, Princeton University.

Mendelovici, Angela (2013). "Reliable Misrepresentation and Tracking Theories of Mental Representation," *Philosophical Studies* 165: 421-443.

Merricks, Trenton (2001). *Objects and Persons.* Oxford University Press.

Merricks, Trenton (2003). Review: Replies. *Philosophy and Phenomenological Research*, *67*(3), 727–44.

Merricks, Trenton (2014). Comments on Van Inwagen's 'Inside and Outside the Ontology Room'. Unpublished comments from the California Metaphysics Conference at the University of Southern California.

Mckinsey, Michael (1991). Anti-Individualism and Privileged Access. *Analysis*, *51*(1), 9-16.

Mitchell, Melanie (1998). *An Introduction to Genetic Algorithms*. Cambridge: Bradford Books MIT Press.

Mogensen, Andreas L. (2015). Evolutionary Debunking Arguments and the Proximate/Ultimate Distinction. *Analysis*, *75*(2), 196–203.

Morganti, Matteo (2009). Inherent properties and statistics with individual particles in quantum mechanics. *Studies in History and Philosophy of Modern Physics, 40*, 223–231.

Morganti, Matteo (2013). *Combining science and metaphysics: Contemporary physics, conceptual revision and common sense*. London: Palgrave Macmillan.

Morganti, Matteo and Claudio Calosi (Forthcoming). Humean supervenience, composition as identity and quantum wholes. *Erkenntnis*.

Nagel, Ernest (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Harcourt, Brace and World.

O'Leary-Hawthorne, John and Michaelis Michael (1996). Compatibilist Semantics in Metaphysics: A Case Study. *Australasian Journal of Philosophy, 74*(1), 117–34.

Polger, Thomas W. and Lawrence A. Shapiro (2016). *The Multiple Realization Book.* Oxford University Press.

Prakash, Chetan., Kyle D. Stephens, Donald D. Hoffman, and Manish Singh (2020). Fitness Beats Truth in the Evolution of Perception. *Acta Biotheoretica*. https://doi.org/10.1007/s10441-020-09400-0

Putnam, Hilary (1975). The Meaning of 'Meaning'. In Keith Gunderson (Ed.), *.Language, Mind and Knowledge* (131-193). University of Minnesota Press.

Putnam, Hilary (1979a). What is Mathematical Truth. In *Mathematics Matter and Method: Philosophical Papers, Volume 1*, 2nd Ed. Cambridge: Cambridge University Press, pp. 60–78.

Putnam, Hilary (1979b). Philosophy of Logic. Reprinted in *Mathematics Matter and Method: Philosophical Papers, Volume 1*, 2nd Ed. (323-357). Cambridge University Press.

Quine, Willard Van Orman (1976). Carnap and Logical Truth. Reprinted in *The Ways of Paradox and Other Essays*, Revised Ed. (107-132). Harvard University Press.

Quine, Willard Van Orman (1980a). On What There Is. Reprinted in *From a Logical Point of View*, 2nd Ed. (1-19). Harvard University Press.

Quine, Willard Van Orman (1980b). Two Dogmas of Empiricism. Reprinted in *From a Logical Point of View*, 2nd Ed. (20-46). Harvard University Press.

Quine, Willard Van Orman (1981a). Things and Their Place in Theories. In *Theories and Things* (1-23). Harvard University Press.

Quine, Willard Van Orman (1981b). Success and Limits of Mathematization. In *Theories and Things* (148-155). Harvard University Press.

Quine, Willard Van Orman and Joseph S. Ullian (1978), *The Web of Belief*, 2nd Ed. McGraw-Hill.

Rettler, Bradley (2018) Mereological Nihilism and Puzzles About Material Objects. *Pacific Philosophical Quarterly 99*, 842–868.

Robertson, Bruce A., Jennifer S. Rehage, and Andrew Sih (2013). Ecological Novelty and the Emergence of Evolutionary Traps. *Trends in Ecology and Evolution* September 2013, *28(9)*: 552-60.

Rosen, Gideon and Cian Dorr (2002). Composition as a fiction. In R. Gale (Ed.), *The Blackwell guide to metaphysics* (151–174). Blackwell.

Schaffer, Jonathan (2007). From nihilism to monism. *Australasian Journal of Philosophy*, *85*, 175–191.

Schwartz, Stephen P. (1978). Putnam on Artifacts. *The Philosophical Review*, *87*(4), 566-574.

Shoemaker, Sydney (1994). Phenomenal Character. *Noûs*, *28*(1), 21-38.

Sider, Theodore (2013). Against parthood. In K. Bennett and D.W. Zimmerman (Eds.), *Oxford studies in metaphysics* (Vol. 8) (237-293). Oxford: Oxford University Press.

Sober, Elliott (1993), "Mathematics and Indispensability" in *The Philosophical Review*, Vol. 102, No. 1 (Jan. 1993), pp. 35-57

Sober, Elliott (1999). The Multiple Realizability Argument against Reductionism. *Philosophy of Science, 66*(4), 542–64.

Street, Sharon (2006). A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies, 127*(1), 109–66.

Tallant, Jonathan (2014). Against Mereological Nihilism. *Synthese, 191*, 1511–1527.

Tan, Peter (2019). Counterpossible Non-Vacuity in Scientific Practice. *Journal of Philosophy, 116*(1), 32–60.

Thomasson, Amie (2003). Realism and Human Kinds. *Philosophy and Phenomenological Research, 67*(3), 580-609.

Uzquiano, Gabriel (2004). Plurals and Simples. *The Monist*, 87(3), 429-451.

van Inwagen, Peter (1990). *Material Beings*. Cornell University Press.

van Inwagen, Peter (2014). Introduction: Inside and Outside the Ontology Room. In *Existence: Essays in Ontology* (1–14). Cambridge University Press.

Vavova, Katia (2015). Evolutionary Debunking of Moral Realism. *Philosophy Compass*, *10*(2), 104–16.

Williamson, Timothy (2007). *The Philosophy of Philosophy.* Blackwell.

Weiner, Jonathan (1994). *The Beak of the Finch.* Vintage Books.