Carving Nature at Its Joints

Natural Kinds in Metaphysics and Science

edited by

Joseph Keim Campbell, Michael O'Rourke, and Matthew H. Slater Carving Nature at Its Joints

Topics in Contemporary Philosophy

Editors

Joseph Keim Campbell, Washington State University Michael O'Rourke, University of Idaho Matthew H. Slater, Bucknell University

Editorial Board Members

Kent Bach, San Francisco State University Michael Bratman, Stanford University Nancy Cartwright, London School of Economics Richard Feldman, University of Rochester John Martin Fischer, University of California, Riverside Nicholas F. Gier, University of Idaho Philip J. Ivanhoe, Boston University Michael McKinsey, Wayne State University John Perry, Stanford University Stephen Schiffer, New York University Harry Silverstein, Washington State University Brian Skyrms, University of California, Irvine Holly Smith, Rutgers University Judith Jarvis Thomson, Massachusetts Institute of Technology Peter van Inwagen, University of Notre Dame

Meaning and Truth (published with Seven Bridges Press), Joseph Keim Campbell, Michael O'Rourke, and David Shier, eds.

Freedom and Determinism, Joseph Keim Campbell, Michael O'Rourke, and David Shier, eds.

Law and Social Justice, Joseph Keim Campbell, Michael O'Rourke, and David Shier, eds.

Causation and Explanation, Joseph Keim Campbell, Michael O'Rourke, and Harry S. Silverstein, eds.

Knowledge and Skepticism, Joseph Keim Campbell, Michael O'Rourke, and Harry S. Silverstein, eds.

Time and Identity, Joseph Keim Campbell, Michael O'Rourke, and Harry S. Silverstein, eds.

Action, Ethics, and Responsibility, Joseph Keim Campbell, Michael O'Rourke, and Harry S. Silverstein, eds.

Carving Nature at Its Joints, Joseph Keim Campbell, Michael O'Rourke, and Matthew H. Slater, eds.

Carving Nature at Its Joints

Natural Kinds in Metaphysics and Science

edited by Joseph Keim Campbell, Michael O'Rourke, and Matthew H. Slater

A Bradford Book The MIT Press Cambridge, Massachusetts London, England © 2011 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

For information about special quantity discounts, please email special_sales @mitpress.mit.edu

This book was set in Stone Sans and Stone Serif by Toppan Best-set Premedia Limited.

Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Carving nature at its joints : natural kinds in metaphysics and science / edited by Joseph Keim Campbell, Michael O'Rourke, and Matthew H. Slater. p. cm.—(Topics in contemporary philosophy) "A Bradford book." Includes bibliographical references and index. ISBN 978-0-262-01593-6 (hardcover : alk. paper)—ISBN 978-0-262-51626-6 (pbk. : alk. paper) 1. Naturalism. 2. Science-Philosophy. 3. Philosophy of nature. 4. Metaphysics. I. Campbell, Joseph Keim, 1958- II. O'Rourke, Michael, 1963-. III. Slater, Matthew Н., 1977–. Q175.32.N38C37 2011 113—dc22 2011002068

10 9 8 7 6 5 4 3 2 1

Contents

Foreword by John Dupré vii Acknowledgments ix

1 Introduction: Lessons from the Scientific Butchery 1

Matthew H. Slater and Andrea Borghini

2 Induction, Samples, and Kinds 33

Peter Godfrey-Smith

3 It Takes More Than All Kinds to Make a World 53 Marc Lange

4 Lange and Laws, Kinds, and Counterfactuals 85 Alexander Bird

5 Are Fundamental Laws Necessary or Contingent? 97 Noa Latham

6 Para-Natural Kinds 113

Roy Sorensen

7 Boundaries, Conventions, and Realism 129

Achille C. Varzi

8 Natural Kinds and Biological Realisms 155

Michael Devitt

9 Three Ways of Resisting Essentialism about Natural Kinds 175 Bence Nanay

10 Arthritis and Nature's Joints **199** Neil E. Williams

11 Predicting Populations by Modeling Individuals 231

Bruce Glymour

12 Similarity and Species Concepts 253

Jason G. Rheins

13 Species Concepts and Natural Goodness 289

Judith K. Crane and Ronald Sandler

14 How to Think about the Free Will/Determinism Problem 313

Kadri Vihvelin

Contributors 341 Index 343

Foreword

The contemporary problem of natural kinds is related to a long tradition of philosophical reflection, dating from at least as far back as Plato's discussions of Forms. Natural kinds offer an answer to the question, what is it for an individual thing to belong to a certain *kind* of thing and thus to be of the same kind as other particular things? Enthusiasts for natural kinds say that the division of things into distinct kinds is a fact about nature and, as such, is a fact that is disclosed to us by science. Philosophical reflection, they argue, tells us what it is for a thing to be a member of a kind; but what kinds there actually are, and what specific properties make a thing (or quantity of stuff) a member of a particular kind, are questions for science to resolve. Often these enthusiasts also believe that science answers the second of these questions by discovering the essence of a particular kind; a stock example is the alleged discovery that the essence of water is to be composed of H₂O. It is also common for contemporary philosophers to analyze other central metaphysical concepts-for example, causality or laws of nature—by appealing to natural kinds.

But there are many skeptics. Ian Hacking reminded us that the contemporary tradition of thought about natural kinds arose only in the nineteenth century, and we should be cautious about generalizing the topic to other parts of philosophical history. Still more recently, Hacking has suggested that the concept may have outlived its usefulness. This reflects a quite widespread skepticism about the extent to which nature presents us with kinds at all. At its strongest, this skepticism takes the form of radical nominalism: all sorting of things into kinds is a free creation of the mind. A weaker skeptical worry is that even if nature provides kinds, she may not provide essences. Kinds might, for instance, be defined by clusters of properties, none of which taken singly is necessary or sufficient for membership in the kind. And finally there are pluralists who argue that nature has provided kinds, but has not done so in any unique way. So natural kinds may overlap and cross-classify things, while things, therefore, might belong to many different kinds.

Hacking's hope that natural-kind talk will subside expresses, of course, one opinion within this range of options, and is, I think, unlikely to be realized: what is at stake in deciding between these positions is a fundamental question about the nature of reality. This volume, then, presents the diverse reflections of a distinguished group of philosophers on a central issue in metaphysics; it will surely be a welcome addition to the literature.

John Dupré Department of Sociology and Philosophy University of Exeter, United Kingdom May 2010

Acknowledgments

Earlier versions of the essays in this volume were presented at the eleventh annual Inland Northwest Philosophy Conference (INPC), held March 15– 17, 2008, in Pullman, Washington and Moscow, Idaho. For their financial and administrative support of the conference, we thank the philosophy departments at Washington State University (David Shier, Chair) and the University of Idaho (Douglas Lind, Chair), the College of Liberal Arts at Washington State University (Erich Lear, Dean), the College of Letters, Arts, & Social Sciences at the University of Idaho (Katherine Aiken, Dean), the research offices at both universities, the Provost's Office of the University of Idaho, and the administrative managers of both philosophy departments, DeeDee Torgeson and Tamara Johnstone-Yellin. We are also grateful for a grant from the Idaho Humanities Council, a state-based affiliate of the National Endowment for the Humanities, to help fund the Public Forum.

All presenters at INPC 2008 were encouraged to submit their work for publication in this volume and, after a process of peer evaluation, only a few were selected. We regret that we had to turn down numerous quality essays due to space limitations. We are grateful to the following referees for their assistance with the selection process: André Ariew, Mark Bedau, Anjan Chakravartty, Larry Colter, Rachel Cooper, Stephen Downes, John Dupré, Dorit Ganson, Clark Glymour, Matthew Haber, Carl Hoefer, Genoveva Martí, Mohan Matthen, John Perry, Julian Reiss, Harry Silverstein, Matt Stichter, Michael Strevens, Patrick Suppes, and Scott Woodcock. We thank the members of the editorial board of the Topics in Contemporary Philosophy series, and would also like to extend our gratitude to Phil Laughlin and Gita Manaktala at the MIT Press for assistance with this volume and for supporting the series.

Finally, special thanks to Delphine Keim Campbell, Laura Lanwermeyer, and Rebecca O'Rourke for their understanding and support.

1 Introduction: Lessons from the Scientific Butchery

Matthew H. Slater and Andrea Borghini

1 Carving Nature at Its Joints

1.1 Tao and the Art of Knife Maintenance

Good chefs know the importance of maintaining sharp knives in the kitchen. What's their secret? A well-worn Taoist allegory offers some advice. The king asks about his butcher's impressive knife-work. "Ordinary butchers," he replies "hack their way through the animal. Thus their knife always needs sharpening. My father taught me the Taoist way. I merely lay the knife by the natural openings and let it find its own way through. Thus it never needs sharpening" (Kahn 1995, vii; see also Watson 2003, 46). Plato famously employed this "carving" metaphor as an analogy for the reality of Forms (*Phaedrus* 265e): like an animal, the world comes to us predivided. Ideally, our best theories will be those which "carve nature at its joints."

While Plato employed this metaphor to convey his view about the reality of Forms, its most common contemporary use involves the success of science—particularly, its success in identifying distinct *kinds* of things. Scientists often report *discovering* new kinds of things—a new species of mammal or a novel kind of fundamental particle, for example—or uncovering more information about already familiar kinds. Moreover, we often notice considerable overlap in different approaches to classification. As Ernst Mayr put it:

No naturalist would question the reality of the species he may find in his garden, whether it is a catbird, chickadee, robin, or starling. And the same is true for trees or flowering plants. Species at a given locality are almost invariably separated from each other by a distinct gap. Nothing convinced me so fully of the reality of species as the observation . . . that the Stone Age natives in the mountains of New Guinea recognize as species exactly the same entities of nature as a western scientist. (1987, 146)

Such agreement is certainly suggestive. It suggests that taxonomies are *discoveries* rather than mere *inventions*. Couple this with their utility in scientific inference and explanation and we have compelling reason for accepting the objective, independent reality of many different *natural kinds* of things. The members of such kinds would be the meat between the joints along which good theories cut. The goal of this introductory essay is to survey some important contemporary trends and issues regarding natural kinds, filling in the picture with key historical episodes. We conclude with a synopsis of the essays contained in this volume.

1.2 Applying the Metaphor

Not everyone appreciates Plato's metaphor. Some dislike its bloody connotations: perhaps we should refocus on garment-deconstruction and speak instead of "cutting nature at its seams." Others find it difficult to make much sense of the metaphor itself: even if actual butchery, past or present, bears out the Taoist ideal of the knife that never needs sharpening, what sense can we give to "nature's joints"? While there is undoubtedly much agreement about how to classify nature, it is not always clear how to interpret this. As Rosenberg (1987) reminds us, even impressively widespread cross-cultural classificatory prejudice might reflect our shared way of seeing the world—a human prejudice—rather than the reality of the divisions themselves.

Moreover, while agreement is common, so is *disagreement*. For example, the dispute about the proper definition of biological species has persisted long enough to have acquired a name: the *species problem*. This leads many to suggest that there are various acceptable ways of carving up biological reality, none of which is privileged over the others. If this is so, do we lose reason for thinking there are natural kinds, at least at this level of granularity? Though the metaphysical status of species has been a key battleground over questions about natural kinds, many related questions are discussed below and in the following essays. In general, we might want an answer to what Ian Hacking has called a "gentle metaphysical question": "are there natural kinds—real or true kinds found in or made by nature?" (1990, 135).¹

Broadly speaking, philosophers have pursued two strategies for fleshing out an answer to this last question. First, we may ask after the *metaphysics* of natural kinds. What (to press Plato's metaphor further) is the "skeletal structure" of nature? Joints are gaps: what are they gaps *between*? At first blush, it would seem that natural kinds are defined by similarity (Quine 1969). Things that are perfect duplicates would seem to be paradigm cases of members of a pristine natural kind. But there are several problems with this line of thought. First, the criterion is too loose. Perfect similarity is not sufficient for making the similar objects a natural kind. Imagine a factory stamping out perfect copies of a widget: few would wish to say that these widgets *thereby* form a natural kind. Second, the criterion is too strict. Requiring perfect similarity among instances of a natural kind would leave us without many of the kinds to which we are pretheoretically committed. 'Metal' or 'tiger' each plausibly names a natural kind of thing, yet we do not expect all metals or tigers to be perfect duplicates of one another. What we need, it seems, is a sense in which things can be similar enough to one another in a scientifically relevant way.

This leads us to a second strategy for identifying natural kinds: look toward their use. As we shall see, this strategy can come in either pure or mixed varieties. Let's start with the mixed (we'll purify in the next section), letting the *purposes* to which we put natural kinds inform our approach to their metaphysics. Consider Hempel's observation that

[t]he vocabulary of science has two basic functions: first, to permit an adequate description of the things and events that are the objects of scientific investigation; second, to permit the establishment of general laws or theories by means of which particular events may be explained and predicted and thus scientifically understood; for to understand a phenomenon scientifically is to show that it occurs in accordance with general laws or theoretical principles. (1965, 139)

In addition to aiding conceptualization and communication, grouping particular things on the basis of shared properties, regularities, dispositions, natural laws, and so forth enables understanding and control. We seek generalizations about what properties things have in common—what they *do*, how they behave. Establishing "general laws" which apply not only to particular objects but to *kinds* of objects allows us to explain and predict. On this model, large swaths of "the vocabulary of science" will necessarily become bound up with general laws. Ernest Nagel noted this connection when he wrote:

The statement that something is water implicitly asserts that a number of properties (a certain state of aggregation, a certain color, a certain freezing and boiling point, certain affinities for entering into chemical reactions with other kinds of substances, etc.) are uniformly associated with each other. (1961, 31 n.32)

Thus, a more nuanced metaphysical picture of natural kinds emerges: kinds as the extensions of *nomic predicates*—predicates that would appear in statements of natural laws.

Though appealing for a number of reasons, the nomic-predicate approach has its difficulties. First, though there are several competing accounts of natural laws,² philosophers seem far from reaching consensus over which is correct. Second, many of these accounts do not apply to rather large swaths of science—even where we suspect that there may be natural kinds. But while few recognize the existence of laws concerning particular species (see Lange 1995, 2004; Mitchell 2000; Woodward 2001), many would like to regard them as natural kinds. Then again, many would *not*. Finding an adequate account of natural kinds is thus complicated by disagreement both over what natural kinds should ideally *do* for us—both in and out of science—and whether categories of things are in fact natural kinds. Before addressing this strategy and its complications in more detail, we shall mention one further confusion encouraged by the phase '*natural* kind'.

1.3 The "Naturalness" of Natural Kinds

Recall that Hacking's gentle question asked whether there were kinds "found in or made by nature." It is not entirely clear how this modifier should be interpreted; nor is it clear that the modifier is appropriate. Granted, it commands some plausibility. As LaPorte notes, adhering to something like it countenances paradigmatic kinds like *tiger, elm,* and *water.* "*Toothpaste, lawyer,* and *trash,* on the other hand, fail to qualify as natural kinds" (2004, 16). But further reflection reveals that "being found in nature" is implausible as either a necessary or sufficient condition for being a natural kind:

Not all human-made kinds fail to be natural kinds. Humans have produced minerals, such as quartz and diamond, in the lab. Humans have also produced elements. Technetium is a synthetically produced element that has not been found to occur naturally on Earth. And humans have created new species of plants by inducing polyploidy. Not only are not all natural kinds produced in nature, but not all kinds in nature are natural kinds: Consider *mud*, *dust*, or *shrub*. These are too close to toothpaste and trash kinds to count as natural. Natural kinds are not distinguished by being found in nature. (LaPorte 2004, 18)

To foreclose on a system's objectivity due to "contamination" by human activity in general would be rash, even if certain kinds of human activity tip us off about such obviously nonobjective cases. It seems to be something about the character of those classification systems more than our simple complicity in their formation. Whatever one thinks of the underlying ontology, *systems* of classification are undeniably human artifacts—we are certainly involved in their creation.

More likely, the 'natural' compliment refers to some collage of a kind's being a nonarbitrary, nonsubjective, relatively elite grouping of things that is important to science. However, as we shall explain in more detail in section 4.3, there may be reason to want to free natural kinds from the exclusive dominion of science. Perhaps there are *social kinds* or *ethical kinds* or *metaphysical kinds* that also, somehow, deserve to be called 'natural'. For now, though, let us continue to focus on natural kinds in science and turn to their role in inductive inference.

2 Natural Kinds and Inductive Inference

Quine reintroduced the concept of a natural kind into philosophical discussion as part of an agreeably unified treatment of two paradoxes of confirmation: Hempel's (1945) ravens paradox and Goodman's (1983) "New Riddle of Induction." The ravens paradox can be generated by two plausible claims about confirmation: first, that positive instance of a generalization lends some support to that generalization; and second, that something which confirms a statement also confirms anything that is logically equivalent to it. The first claim is sometimes called "the instantial model" of confirmation. For example, if I'm trying to confirm the hypothesis that all ravens are black, it helps to find an *instance* of that generalization: a black raven. So far so good. Now the statement that all ravens are black is equivalent to the statement that all non-black things are nonravens. The instantial model says that every instance of a non-black nonraven—a red fire truck, a blue suede shoe, and so on—confirms it. But since this generalization is equivalent to our all ravens are black hypothesis, these miscellaneous things apparently confirm it too, opening the door for "indoor ornithology." That seems wrong.³

Goodman's "New Riddle" also infects that plausible instantial model of confirmation. Suppose we define a predicate 'grue' as applying to anything that is either green and observed before now or blue and unobserved. Assuming all observed emeralds have been green, they've all *also* been "grue" and thus on the instantial model support the conclusion that all emeralds are grue. Assuming that some emeralds are as yet unobserved, this entails the conclusion that some emeralds are *blue*.

Quine's solution in both cases was to call upon natural kinds as the extensions of "projectible predicates" to restrict the instantial model. Certain predicates—'raven' and 'emerald' among them—are posited to be distinguished in science by being confirmable by their instances. While 'raven' might name a natural kind, its complement—'non-raven'—does

not. Likewise, 'green' might name a natural kind of color, whereas 'grue' does not. Rather than seeking some metaphysical foundation for projectibility and letting *that* define natural kinds (what we are calling the "mixed approach" above), the present strategy puts all of the emphasis on projectibility and has that direct our approach to the metaphysics of natural kinds.

Quine's move seems productive. There does seem to be something suspiciously "unnatural" and miscellaneous about the grue things and the non-ravens that might interfere with their operating straightforwardly with our confirmatory practices. But as we saw above, it is difficult to say precisely what the compliment 'natural' amounts to. Without an answer to this question, we merely replace one difficult problem with another: identifying which predicates are *projectible*. Hacking puts this point nicely: "'Projectibility' becomes the name of an as yet unanalyzed feature of predicates, namely that they are and can be used inductively. Then the new riddle of induction achieves a succinct formulation, 'Which predicates are projectible?'" (1995, 202). But this just prompts the question again: what is it to be a natural kind? On the other hand, construing natural kinds simply as the extensions of projectible predicates leaves the problem of induction untouched. It looks as though we must choose which bird to pelt with our stone.

Quine toys with the former route, construing natural kinds in a manner Goodman painstakingly avoided: in terms of overall similarity.⁴ Ravens are relevantly similar to each other; non-ravens are not. Though in general cautious about kinds and the allied notion of comparative similarity, he believed the latter notion to be ready to hand in chemistry:

Comparative similarity of the sort that matters for chemistry can be stated outright in chemical terms, that is, in terms of chemical composition. Molecules will be said to match if they contain atoms of the same elements in the same topological combinations. . . . At any rate a lusty chemical similarity concept is assured. (Quine 1969, 135)

Quine saw the objectivity of chemical kinds as secured by their common chemical structure. This is, presumably what makes emeralds, but not nonemeralds, projectible. The italicized qualifier—"of the sort that matters for chemistry"—is important here. Presumably, what matters for chemistry is what matters *for chemists*: the particular reactivity of various chemical stuffs. And clearly, the topological structure of chemical substances' basic components is here of considerable importance. As we shall see below, Quine's thought found fertile ground with Kripke (1980) and Putnam (1975), who revitalized a form of *essentialism* about natural kinds that can be traced back to Aristotle. Free from Quine's antiessentialist scruples, they developed a modern version of the Lockean distinction between real and nominal essences. Natural kinds, they claim, are indeed individuated by hidden real essences. Unlike Locke, however, they were quite sanguine about our ability to *discover* such essences. For them, this was the bedrock upon which objective taxonomies could be built. In the next section, we trace some of this story.

3 The Question of Essentialism

Let us speak for a moment just about the qualitative features of objectswhat philosophers typically call their *properties*. Properties can be possessed in different ways. Ordinarily, that some object has a property P is an "accidental" matter—not in the sense of being regrettable or a fluke, but in that it might well not have had that property. For example, while Roger Federer is in fact a tennis player, he might not have been: he could have pursued a different career (and still have been the same person). Federer is also rational. But it is far less clear that he could have lacked this quality (while remaining the same person). If this is right, we say that the quality of rationality is essential to Federer, whereas that of being a tennis player is merely accidental. In general, the essential properties E of an object are those that determine what that object is. In other words, E includes those properties upon which the understanding of the object rests. It also includes some of the properties on which its existence depends (there may be others, which are non-essential, and on which its existence also depends.) In the Western tradition, the concept of an essential property dates back to Aristotle; it enjoyed much fortune in medieval and early modern philosophy, and is still somewhat in vogue.⁵

3.1 Aristotle on Essences

Setting his tennis prowess aside, Federer is still a unique individual—there is literally no one else who is he.⁶ On the other hand, he is many things that other people are as well. For example, he is a professional tennis player: one of the many who compete in tournaments. He is also *a* person: one of the many who inhabit the globe. So we have *one* individual—Federer—who is at the same time *many* things: he is one but he is also many. And thus we have "the problem of the one and the many." To solve this problem is tantamount to giving an explanation of kind-membership (or at least of possessing a property).

Plato tried to make sense of kind-membership by positing a relation of "taking part" or "participation" in a kind or a property (what Plato called "Forms"). For example, Mary and Hannah are both human as they participate in the *form of humanity*—an abstract, ideal, nonconcrete entity. This is how Plato proposed to understand the "jointedness" of nature: nature's joints are defined by the Forms. Yet Plato himself presented formidable objections to this project in the *Parmenides*—some of which seemed more compelling than the view itself. For this reason, perhaps, Aristotle set out to provide a different metaphysics. But as Karl Popper once put it, while Aristotle denied "Plato's peculiar belief that the essence of sensible things can be found in other and more real things . . . [Aristotle] agreed with him in determining the task of pure knowledge as the discovery of the hidden nature or Form or essence of things" (Popper 1950, 34). A pillar of the novel metaphysics was Essentialism, upon which Aristotle elaborates most famously in the *Categories*, the *Metaphysics*, and the *Posterior Analytics*.

In *Categories* 2 and 3, Aristotle draws some distinctions which provide the logical foundation for postulating the existence of essences. First of all, he claims that there are two kinds of predications: *to say of* and *to be in*. If B can be *said of* A, then B's definition can be predicated of A. On the other hand, if B cannot be said of A but it *is in* A, then B's definition cannot be predicated of A. For example, we can *say of* Rubi that he is a dog because whatever defines being a dog also defines Rubi. On the other hand, whiteness *is in* Rubi but cannot be said of Rubi, as he is not defined by whiteness, though of course, something else—for instance, snow—may be defined by whiteness. Although it appears that the focus of the *Categories* is to furnish guidelines for classificatory purposes, the distinction between "saying of" and "being in" is already a hint of the essentialist attitude more explicitly advocated in other works.

From here, Aristotle distinguishes four kinds of entities:

Of things themselves some are predicable of [i.e., said of] a subject, and are never present in a subject . . . Some things, again, are present in a subject, but are never predicable of [said of] a subject . . . Other things, again, are both predicable of [said of] a subject and present in a subject . . . There is, lastly, a class of things which are neither present in a subject nor predicable of [said of] a subject, such as the individual man or the individual horse. But, to speak more generally, that which is individual and has the character of a unit is never predicable of a subject. (*Categories* 2)

Following the standard scholastic interpretation of the "ontological square," we can devise: (i) *primary substances*, such as Rubi, that can neither

be *said of* nor *be in* other entities; (ii) *secondary substances*, such as dogness, that can be *said of* some other entities but that cannot *be in* other entities; (iii) *universal accidents*, such as whiteness, that can both *be said* and *be in* other entities; (iv) *individual accidents*, such as Rubi's whiteness, that can *be in* other entities but cannot be *said of* other entities.

From this analysis of predication Aristotle draws the conclusion that individuals (what he refers to as "primary substances") are the ultimate constituents of reality because they cannot be predicated, in any way, of other entities. You can say:

(1) Socrates is wise

but you cannot meaningfully say:

(2) Wisdom is Socrates

because Socrates is a kind of entity (i) that cannot be predicated, in any way, of other entities. Essences belong to (ii), while accidents may belong to (iii) or (iv). The distinctions drawn here, however, were meant mostly for classificatory purposes. How did Aristotle justify the postulation of essences in metaphysical terms?

To answer this question we should look into the *Metaphysics*, one of Aristotle's more mature works, especially books VII and XII, where the distinction between form and matter emerges more starkly. Here too he portrays the essence of an individual as that which defines it and without which it could not exist: "For the essence is precisely what something is . . . Therefore, there is an essence only of those things whose formula is a definition" (*Metaphysics* VII, pt. 4). But a new piece is added to the view: essences are now related to forms, "and so Plato was not far wrong when he said that there are as many Forms as there are kinds of natural objects" (*Metaphysics* XII, pt. 3). Yet Aristotle holds that Plato was wrong in claiming that forms by themselves are enough: "and so to reduce all things thus to Forms and to eliminate the matter is useless labour; for some things surely are a particular form in a particular matter, or particular things in a particular state" (*Metaphysics* VII, pt. 11). Thus, Aristotle sketches a theory of essences and individuals that will survive until present times.

In the *Posterior Analytics* Aristotle refines his theory of essences in the context of providing a secure path to knowledge. He puts forward a model of scientific explanation known as the Connecting Term Model according to which the fact A explains the fact C in virtue of another fact—B—which connects A to B and B to C. Why does eating sugar (A) necessarily make you gain weight (C)? Because eating sugar (A) necessarily increases your

bodily fat (*B*) and increasing your bodily fat (*B*) necessarily makes you gain weight (*C*). Aristotle's view stresses the necessity of the tie between the *explanandum* and the *explanans*, thus bringing what he regarded as decisive evidence in favor of essentialism: "Demonstrative knowledge must rest on necessary basic truths. . . . Now, attributes attaching essentially to their subjects attach necessarily to them. . . . It follows from this that premisses of the demonstrative syllogism must be connexions essential in the sense explained: for all attributes must inhere essentially or else be accidental, and accidental attributes are not necessary to their subjects" (*Posterior Analytics* I.6).

Aristotle's model for scientific explanation had a great impact on the future understanding of scientific method and constituted a knockdown argument against those who took a skeptical attitude toward essentialism: in a way, it proved that if scientific findings increase to any extent our knowledge, then they must do so by means of necessary connections; and said connections require essential attributes if they can be deemed necessary at all.

3.2 Locke on Essences

The tremendous success of Aristotle's metaphysics down the centuries secured the prominence of essences, granting them a chief role in the explanation of kind-membership. Along this path, philosophers' understanding of essences (and philosophical appreciation of their virtues and vices) changed dramatically. We don't have the space here to even survey these changes apart from a modern doctrine of essence whose import is still felt: that of John Locke.

During the early modern period epistemological issues undermined much of the scholastic philosophical tradition—and the Aristotelian doctrine of essences was no exception. Despite its previous success, it was newly on the brink. Even while granting that essential properties play a key metaphysical and conceptual role in delineating nature's joints, the means through which we come to gather information *about* these essences seem obscure. After all, it's by his accidental properties (elegant appearance, calm demeanor, tennis prowess) that Federer is known as an individual. Likewise, it seems that different natural kinds are regularly, though imperfectly, associated with their merely accidental properties. Gold, for example, is ordinarily identified by certain superficial properties: it's the stuff that's a shiny yellow ductile metal for which people will pay dearly. Are any of these properties *essential* to gold? Just how, in general, should we tell the difference between accidental, superficial properties and those which are *essential* to their bearers? How do we distinguish between what merely happens to be so and what *must* be so?

Locke took these questions seriously and advanced a novel proposal. First of all, he defined a quality of a subject as "the power to produce any idea in our mind" (*Essay* II.8.8). Next, he distinguished between *primary* and *secondary qualities*: the former being "utterly inseparable from the body, in what state soever it be" (II.8.9), the latter being the powers of the objects "to produce various sensations in us by their primary qualities" (II.8.10). Intuitively, the superficial properties of gold are its secondary qualities—the way it looks to us when we first encounter it in everyday experience. The primary qualities, on the other hand, are those which remain hidden to our senses but which specialized reasoning might reveal. Locke listed "solidity, extension, figure, and mobility" (II.8.10); we might now list a certain atomic structure, a typical charge or specific weight, and so on. Locke also considered a third category of qualities, bare powers powers of objects to modify other nonmental objects, such as the power of a key to open a lock. But we shall focus on the first two categories.

Locke's division among qualities of objects took its impetus from epistemic considerations. By definition, a quality is that which produces an idea in the mind. A champion of empiricism, he believed that if you cannot reliably come to know something through experience, you cannot say that it exists. This allowed him a fresh start also with respect to scientific essentialism, the stronghold of Aristotelian essentialism. Locke distinguished between a substance's nominal essence-"The measure and boundary of each sort or species, whereby it is constituted that particular sort, and distinguished from others, is that we call its essence, which is nothing but that abstract idea to which the name is annexed" (Essay, III.6.2)—and its real essence—"that real constitution of anything, which is the foundation of all those properties that are combined in, and are constantly found to co-exist with the nominal essence" (III.6.6). He then argued that our *ideas* of substances associate only with their nominal essences: "take but away the abstract ideas by which we sort individuals, and rank them under common names, and then the thought of anything essential to any of them instantly vanishes" (III.6.4). Precisely for this reason, he himself seemed ambivalent about our ability to fully grasp real essences. After all, our *ideas* of substances associate only with their nominal essences since we lack "microscopical eyes" to see real essences. Thus with Locke a new form of essentialism came into the picture, one which sees essences as abstract ideas that are applied to individuals-that is, a view which sees essences as *sorts* of things. As we shall see, we can identify a

parallel distinction between a sortal understanding of essences and an Aristotelian one in contemporary philosophers' treatments of these matters.

Almost a century after Locke published his *Essay*, David Hume's empiricism began to determine the philosophical fate of essentialism in England (and much of continental Europe) over the next two centuries. The cultural environment in which early twentieth-century philosophers of science wrote, steeped in post-Humean empiricism, had little truck with such seemingly occult notions as essence. The only sort of necessity worth having was a purely linguistic matter. It is relative to this trajectory that we can appreciate how dramatic was the revival of essentialism in the second half of the twentieth century. Two quite distinct branches can be identified in this revival: on the one hand we have the sortal tradition (sec. 3.3), and on the other the Kripkean–Putnamian one (sec. 3.4).

3.3 A Metaphysical Rebirth of Essentialism

The intuition that nature can be carved up into different *sorts* of things and that each thing is something of some *sort* lies at the basis of a wide-spread, metaphysically motivated revival of essences. Still revered by many, this *sortal* tradition, which flourished primarily in England, engrained a Lockean approach to essential properties and the close analysis of natural language.⁷ But the underlying doctrine is less homogeneous than it might first appear. Indeed, even if many defended a theory of *sortals*, few agreed on the meaning of that term. Following Feldman (1973), we can distinguish three necessary requirements that a predicate *P* has to satisfy to be a *sortal*:⁸

i. A predicate *P* is a sortal only if *P* singles out an individual.

ii. A predicate P is a sortal only if P is the partial or whole essence of the individual it singles out.

iii. A predicate P is a sortal only if, when P applies to an individual x, P cannot belong to any proper part y of x.

Arguably, (i), (ii), and (iii) serve different metaphysical purposes, yet there is no agreement between sortal theorists as to which of them a *sortal* should satisfy.⁹ At any rate, we may leave this issue to one side, as the sortal tradition had a considerably smaller impact on the debate over natural kinds than did the tradition initiated by Saul Kripke and Hilary Putnam, to which we now turn.

3.4 A Scientific Rebirth of Essentialism

In the 1970s, Kripke (1972, 1980) and Putnam (1975) independently defended the existence of essences—via rather different considerations. At

that time, Kripke was trying to offer a theory of reference which would account for, among other things, the way in which natural-kind terms function. His theory revamped the idea that the identity of an individual is necessary, that it is fixed in every possible scenario. Essences offered a handy explanation of this: the identity of an individual is fixed because it has some essential properties. As we have seen, however, at this point we face the epistemic challenges that confronted Locke.

Here lies Kripke's main innovation. He conjectured that essential properties are directly linked to our linguistic practices (such as naming) and our scientific concepts (such as genetic identity). Whereas previous theories of reference had it that names referred to individuals by way of descriptions, Kripke argued instead that a name reaches its bearer *directly* and continues to refer even if the properties we *in fact* use to identify it are missing. The name 'Federer' does not merely refer to that calm, elegant, person of Swiss origin who has won a certain number of tennis tournaments, but to *that guy*. The idea is that there is something *essential* about Federer since the first time we called him that name—perhaps something about his genetic makeup or origins (having the parents he did). Kripke moved to extend this plausible idea about proper names to natural-kind terms. When we first referred to 'water', say, we refer not to whatever satisfies certain characteristic properties (being clear, potable, liquid at standard temperature and pressure, and so on), but to that stuff. And when scientists discovered that that stuff was H₂O, they discovered the essence of water. Kripke produced an elegant proof that all identities were *necessary* identities.

Putnam's considerations on essences also proceeded from semantic considerations. Specifically, they grew out of the attempt to furnish a broader theory of meaning. In a deeply influential paper, "The Meaning of 'Meaning'" (1975), he distinguishes between two types of content: *narrow* and *wide*. Narrow content reflects the psychological state of an individual in isolation, whereas wide content includes content which is not part of that individual's thoughts but is nevertheless entailed by them. The existence of wide content suggests the existence of essential features of reality. For if the meaning of what we say about certain natural kinds (water, for example) is fixed in part by the *essence* of that kind, then we have good reason for accepting the existence of essences. Suppose we talk about this glass of water: its identity is not just fixed by the perceptual experience that you are having or what qualities you generally associate with water, but also by the very *essence* that the stuff we call "water." What is that essence? Well, one very plausible answer is that it is the properties which *explain* the co-occurrence of those superficial, "nominal" properties, whose essence is presumably (partially) captured by the molecular formula H₂O.

After Kripke and Putnam's contributions, the discussion of essential properties within the philosophy of science got a fresh start. They were able to bring back this notion in a way that was *prima facie* immune from the suspicion surrounding much of the ancient, Scholastic, and modern usages of it. Whether this is so is still a matter of much debate.

Let us consider one further twist in the story of essentialism about natural kinds. Plausibly, the role they play in scientific endeavors turns on their association with lawlike behaviors: we see the names of natural kinds habitually turn up in statements about natural laws. One might deny that this is coincidental and simply claim that kinds are law-involving, or *nomic*, predicates. But then the questions become: *What is it to be a nomic predicate? What are laws in general and what explains their apparent generality and necessity?*

Scientific essentialism attempts to answer this second question. The label first appeared in "The Philosophical Limits of Scientific Essentialism" (1987), by George Bealer. In that article, Bealer criticized Kripke-style essentialists, according to whom essential properties can be discovered a posteriori. Despite Bealer's aims, a number of influential authors embraced scientific essentialism and refined its metaphysical underpinnings (Bigelow, Ellis and Lierse 1992; Ellis and Lierse 1994; Ellis 2001; Bird 2007). In its present form, scientific essentialism is a hardcore metaphysical view, according to which kinds exhibit lawlike behaviors as manifestations of the *dispositions* which define them. Dispositions, roughly speaking, are abilities to act in one way or another given certain circumstances. On this view, laws of nature are immanent to the entities possessing certain dispositions. Although Kripke and Putnam never ventured into these sorts of metaphysical speculations, the gist of scientific essentialism owes a great deal to their revival of essentialism and to Kripke's suggestion that the essences of natural kinds may be discovered a posteriori.

By construing laws as manifestations of the essential dispositional natures of different natural kinds of things, scientific essentialists effectively solve two problems about laws of nature and their relation to natural kinds. First, the vague intuition that natural kinds were somehow implicated in natural laws becomes precise and understandable. Second, by making the laws expressions of the *essential* nature of different kinds, scientific essentialists dispense with one of the most difficult problems in giving an account of natural laws: making sense of their apparently "intermediate" strength of necessity.¹⁰ Essentialists thus hold that not only are

the laws somehow more robust than accidental generalizations, but that they *had* to be just the way they are.

4 Applications

4.1 Physico-Chemical Kinds

Chemical kinds have long been a favorite example of essentialists. For as both Quine and Putnam noted, it seems quite plausible that the sort of similarity that would matter for this domain would be molecular structure: the arrangement of certain kinds of atoms. Putnam claims that the essence of water—what it is to *be* water—is to have the molecular structure denoted by 'H₂O'. The superficial properties we *associate* with water—for example, its being a good solvent for certain types of compounds—are explained by its structure. More specifically, the structure *plus* the character of its constituent atoms gives rise to these properties.

How then should we understand what divides atoms into different kinds? An analogous story seems likely: the arrangements of subatomic particles (viz., protons, neutrons, and electrons) explains why oxygen covets electrons and why hydrogen is comparatively willing to give them up. But then we need a story about the character of these subatomic constituents. What explains why protons have the charge and mass that they do? According to the Standard Model of particle physics, the answer lies in its composition of quarks and *their* dispositions. Thus, we have a recursive picture of the identity of physico-chemical kinds. The identity of a kind at a certain level of compositional complexity is fixed by arrangements of things at a lower level of complexity. One might wonder at this point whether it is, so to say, "turtles all the way down" or whether complexity bottoms out. Contemporary physics seems to support the latter view. It treats certain kinds (such as quarks and electrons) as *fundamental* in that they apparently lack structure. They are part of the bottom level of physical complexity and thus kinds whose essence can no longer be understood structurally. On the other hand, the very use of the word 'atom' (meaning "something that is partless") for one of these intermediate levels suggests that we ought be cautious about identifying a particular level as fundamental!

While the foregoing sketch may look quite plausible and unproblematic, there are deep and persistent issues involved. We have not discussed *how* reference to physical or chemical kinds is achieved. Is it, as Kripke and Putnam suggest, a *direct* matter? Reference aside, we might also wonder whether the proffered essences are plausible. Take any glass of water: it is filled with many things that are not composed of H_2O . In addition to various isotopic forms of water (various "heavy waters," for instance), there are doubtless other impurities (e.g., minerals, trace elements, dissolved gasses, even microorganisms). The same could be said for the sample initially "baptized" as water. What makes it the case that this initial dubbing fixed on the H_2O sameness relation?¹¹

4.2 Biological Kinds

Such worries notwithstanding, the essentialist view of natural kinds has seemed compelling enough to extend to higher levels of organization. Hopes initially turned toward extending kindred notions of structure to the biological realm: perhaps tigers have a certain genetic structure which alone makes them tigers. We cannot "define" tigers as, say, fierce striped *feline quadrupeds* because some tigers lack these qualifications (Kripke 1980, 119–120). Just as water behaves differently in different conditions, tigers get maimed or adapt certain behavioral patterns in different environments. Tigers are not easily genetically maimed, though, and their genetic structure is causally upstream from their stripes and fierceness. Insofar as genetic structure remains stable-serving as an explanation for our habitual association of a certain nominal essence with tigers-it seems an admirable candidate for the office of "real essence of tiger." As Robert Wilson characterizes this view: "species essence is not constituted by [observable] morphological properties themselves, but by the genetic properties—such as having particular sequences of DNA in the genome-that are causally responsible for the morphological properties" (1999, 190).¹²

But again, while initially tempting, this view faces several objections. First, even if we are impressed by the structural account of physicochemical kinds, we should bear in mind that "genetic structure" and "molecular structure" do not play the same causal role. An organism's genetic structure does not determine its superficial properties in nearly as direct a way as molecular structure does the superficial properties of homogeneous chemical kinds (see Lewontin 2000 for a nice discussion of this point). Second, the fact of evolution and the considerable diversity of species raises the question of whether there even *is* a genetic essence that all and only the members of a particular species share (see Devitt 2008; Okasha 2002; Walsh 2006; Wilson 1999). And third, many philosophers of biology (e.g., Dupré 1981; Kitcher 1984; Mishler and Donohue 1982) have concluded that we ought be *pluralists* about biological classification (at least at the rank of species). How might this affect a conviction that species divisions carve nature at its joints?

There is much to say in response to these worries (and there are others besides), and the philosophical community remains largely divided. Some suggest that we can reconfigure our understanding of natural-kinds essentialism in light of the majority view in systematics to accommodate "historical essences" (Griffiths 1999; LaPorte 2004; Okasha 2002). What makes a tiger the kind of thing it is not some intrinsic genetic property, but a historical property concerning its origin-its location on the tree of life, say. Others take the common practice of treating species historically as suggesting a radically different metaphysical approach to species. Rather than treat species as kinds, perhaps we should understand them as individuals-spatiotemporally extended objects, "hunks of the genealogical nexus"—perhaps as a way of resisting pluralism about species or rendering it a purely pragmatic issue (Ghiselin 1974; Hull 1978). Others may be content to simply abandon the attempt to extend Plato's metaphor of natural joints to the biological realm. Hacking's "gentle metaphysical question" is general: it can receive a positive answer without natural kinds being particularly common in science. One could conceivably be pushed all the way back to construing only the fundamental physical particles as natural kinds.

Yet this smacks of parochialism. As Dupré remarks, biology "is surely the science that addresses much of what is of greatest concern to us biological beings, and if it cannot serve as a paradigm for science, then science is a far less interesting undertaking than is generally supposed" (1993, 1). Whether or not one agrees with Dupré's assessment, it seems plausible that many biological categories do play an inferential and explanatory role commonly associated with natural kinds. This puts pressure on the traditional essentialist view of natural kinds.

A number of philosophers have been pursuing a suggestion of Richard Boyd's (1991, 1999): that there may be a class of phenomena accurately described as "homeostatic property clusters."¹³ This apparently non-essentialist understanding of natural kinds appears better able to make sense of biological diversity. Roughly speaking, Boyd eschews essential properties which "hold together" and explain the co-occurrence of the various superficial properties associated with a kind, suggesting instead that a cluster of properties might secure *its own* stability, constituting a sort of homeostatic mechanism. Insofar as such homeostatic property cluster (HPC) kinds accommodate our inductive and explanatory practices, we are within our rights to regard them as *real* (see Neil Williams' essay below for further discussion).

4.3 All Kinds of Kinds

Thus far, we have restricted our scope to scientific kinds—and a rather limited swath there. Might our concepts also carve nature at *other* joints? In addition to natural kinds of *things*—particles, organisms, and so on might there be natural kinds of events, processes, forces, laws, states of affairs, and so on? Even within the biological sciences, we see quite a diversity of classificatory concepts being employed. Biologists (both implicitly and explicitly) draw upon a rich stock of biological categories (e.g., "predator," "decomposer," "muscle tissue," "afferent neuron," "neurodegenerative disease") in deepening our knowledge of the organic world.

And what about kinds outside of the natural sciences—for example, from the social sciences and beyond? There seems to be no a priori reason to exclude these farther-flung applications. For even nonscientific kinds often seem to come with particular dispositional behaviors which are entrenched in different kinds of relations. In the social sciences, we might wonder whether there are genuinely different *kinds* of people, societies, economic systems, and so on. There is currently a vigorous debate in the philosophy of science concerning the status of racial divisions: do race terms name natural kinds of people (Andreasen 1998; Kitcher 1999; Zack 2002; Pigliucci and Kaplan 2003; Hacking 2005; Glasgow 2009)?

Psychology offers a particularly rich set of examples. Griffiths (1997) has explored the question about whether emotions and other psychological states might be natural kinds. Boyd (1999, 155) even flirts with the notion that the categories "feudal economy" and "capitalist economy" might name natural kinds, finding no difficulties in principle with construing unabashedly human creations as nevertheless *natural* in the relevant sense.

What about other conventional-seeming categories? At some point, we may wish to distinguish between *natural* and *social* kinds. Consider a citizen *versus* an illegal alien, a sole proprietorship *versus* a limited liability company, a not-for-profit organization *versus* a for-profit business. These are examples of classifications that can play key roles in a society, and their roles are governed not by *natural laws* but by laws in the more familiar and mundane sense—each is associated with different rights, duties, and privileges. But there are examples that might be less clearly identified as "natural" or "social." Consider, for example, the kinds that you find in front of you every day on supermarket shelves or on your plate—*food kinds*. A chicken can be *free range*, an egg *certified organic*. Although it may be disputed that vernacular expressions are able to pick out natural kinds (Dupré 1993, 26ff.), nonetheless they pick out kinds that are important for practical purposes.

Returning to the philosophical terrain, those key *formal relations*—such as identity, parthood, membership (in a set), spatiotemporal location—can be regarded as kinds within the metaphysical realm (Sider 2009). Might there even be natural kinds of *absences*? On the other hand, consider the role of kinds in ethics (Boyd 1988): moral realists may wish to say that wrong actions comprise a natural kind (or even a hierarchically nested series of natural kinds).

We humans love to draw lines around different portions of the world, so there should be no shortage of fascinating possibilities to consider when we ask whether we are, in so doing, carving nature at its joints.

5 The Essays

So much by way of introduction. Hopefully you are eager to read the fine essays you have before you.

As we saw above, one of the central roles philosophers have attributed to natural kinds is that they serve as the metaphysical basis for inductive inference. Only predicates in whose extensions stands a natural kind are "projectible"—a theme sounded in different ways by Quine and Goodman. Godfrey-Smith, in "Induction, Samples, and Kinds" (chap. 2) challenges this orthodoxy by suggesting that there are in fact two varieties of inductive inference that have been run together. In only one of these varieties does the "naturalness" of kinds play any significant role: at stake in these inferences are generally dependence relations linking properties. As such, the number of samples is, in principle, irrelevant to the strength of the inference. If we can establish the dependence relation by examining only one positive instance, we can get the generalization in all of its glory. But there is another strategy of inference in which the strength of the inference to a generalization depends on the quality of our sampling: in particular, that it is broad and random. Here apparently pathological cases, like Goodman's "grue," can be explained away in familiar terms as certain kinds of "observation selection effects." Godfrey-Smith argues that distinguishing these two inductive strategies can go a long way toward relieving some longstanding philosophical (perhaps innate!) confusions about induction.

Marc Lange turns his sights on the growing support for scientific essentialism in his essay, "It Takes More Than All Kinds to Make a World" (chap. 3). As we pointed out above, elementary physical particles appear to be admirable candidates for natural kinds, if anything is. Assuming something like the Standard Model is correct, they are intrinsic duplicates defined by a small collection of properties (such as charge and spin). But, Lange points out, if there is something to the modern physical practice of recognizing different "tiers" of natural laws—if, for example, there are symmetry principles that abstract away from particular laws like Coulomb's law—we need to make sense of certain "counterlegals," that is, counterfactuals involving breaks of laws. Scientific essentialists contend that the essence of charged particles such as electrons give rise to Coulomb's law. But how can the essentialist make sense of counterlegals such as 'Had Coulomb's law failed to be true, the fundamental dynamical laws would still have held'? What essence could possibly account for this subjunctive fact? This is the sense in which it takes more than all of the *actual* kinds in order to make a world complete with laws. The scientific essentialist would need far more.

Along the way, Lange elaborates a view on the relation between laws and subjunctives that he defended in *Natural Laws in Scientific Practice* (2000) and more recently in *Laws and Lawmakers* (2009), and discusses the vexed question of what makes some properties "natural," offering the very interesting suggestion that it might be that a property could be natural in one possible world and unnatural in another.

In "Lange and Laws, Kinds, and Counterfactuals" (chap. 4), Alexander Bird questions one of the key contentions in Lange's essay: that if there had been kinds of particles other than the actual kinds, the force laws (and laws connecting fundamental and derivative properties) would still have held. One reason for not accepting this, suggests Bird, is that we don't yet know what the fundamental laws are. Perhaps whatever these turn out to be are not as independent from the existence of certain kinds of particles as we are tempted to suppose. Moreover, in at least some cases, we find interesting connections between the existence of certain kinds and fundamental laws. For example, the non-existence of certain conceivable particles (e.g., Helium-2) seems to be governed by fundamental forces (e.g., the strong force). As Bird explains, a natural way of resisting his skepticism involves forbidding "backtracking" reasoning about counterfactuals. But this plausibly both undermines the inference from the claim about the independence of laws and fundamental kinds to other results claimed by Lange and leads to some odd consequences (e.g., that the first event-the Big Bang, possibly—would have a kind of physical necessity). More generally, Bird suggests that the idea of a hierarchy of laws formed by Lange's proposal about laws is not quite as secure or important as Lange thinks. This debate will no doubt continue.

As we mentioned above, one way of thinking about the dispute between those who see laws as necessary and those who believe them to be contingent involves investigating their connection with kinds or properties. Noa Latham pursues the thread in his paper, "Are Fundamental Laws Necessary or Contingent?" (chap. 5), arguing against the grain that there is in fact no significant distinction between necessitarian views of laws (espoused by the scientific essentialists) and contingentist views (those like Lange who deny that laws are metaphysically necessary). These views are best understood as notational variants of a single view. Latham's argument turns on claims about the metaphysics of property-individuation-for one, that it makes no sense to think about stripping away all of the nomological features of a property, leaving a sort of contingent shell. But from this extreme contingentist view about property-identity, there is much leeway-and possibly no fact of the matter-about how much we should pack into our concept of properties. There might still be reasons for locating oneself at one end of the spectrum (e.g., the necessitarians do not face the difficult problem of multiplying senses of necessity; contingentists have a more linguistically natural view), but Latham claims that these reasons fall short of the kind of metaphysical strength that their proponents have in mind.

Shifting gears somewhat, Roy Sorensen's essay, "Para-Natural Kinds" (chap. 6), flirts with rejecting the prevalent view that only *substances* can be natural kinds. What about absences (gaps in an electron shell, craters in the moon)? What about shadows? On reflection, even these "nothings" evince classificatory possibility. Sorensen calls them *para-natural kinds*: absences *defined* by natural kinds. It's not surprising that we might have been tempted to treat certain absences as natural kinds, for like reflections they take on many of the hallmark features—lawfulness, projectibility, and so on—possessed by the natural kinds which define them. Such features allay general worries about the "subjectivity" of absences. The absence of a chapter in this volume on what kind of doughnut Plato would prefer is a subjective absence salient only to those who might have expected one. In contrast, Sorensen contends that para-natural kinds are mind-independent.

The road to essential properties passes through the individuation of their bearers: if something has an essence, then it is *something*. In his essay, "Boundaries, Conventions, and Realism" (chap. 7), Achille Varzi questions the existence of boundaries between individuals and events of all sorts, thereby disputing the existence of essences' bearers. His argument moves from the distinction between *artificial* and *natural* boundaries (also labeled *fiat* and *bona fide*, respectively). When we uncover a *natural* boundary (one that is *not* merely fiat), we thereby have a reason to believe that we are in

the presence of a genuine individual (or event). On the other hand, when confronted with artificial boundaries, the suspicion of being in the presence of a genuinely artificial individual (or event) surfaces. In his essay, Varzi surmises that all boundaries are artificial, and he substantiates such a thesis by surveying a host of examples—from geography to geopolitics to biotechnology. From this it follows that every individual (or event) is, to some extent, artificial; but from this it does not also follow that anything goes. He concludes by reassuring us that artificial boundaries are, in the end, all that we need "to solve, in an arbitrary but efficient way, coordination problems" of all sorts, and that such a stance is compatible with rigorous metaphysics, such as those advanced by Putnam or Goodman.

But suppose that one were to resist Varzi's challenge in the name of some form of "realism" about natural kinds and essences; what does it take-Michael Devitt wonders in his essay "Natural Kinds and Biological Realisms" (chap. 8)—to be such a realist? Moving from the species problem as a case study, Devitt defines realism as that view according to which certain entities play a role that is causally significant because of the kind of thing they are (i.e., things that "cut nature at its joints"). This understanding of realism should, however, be kept distinct from two other notions: one according to which realism is committed to the mindindependent existence of certain entities; and another according to which realism is committed to the existence of universals. Devitt thus shows that we ought to keep separate issues about the realism of certain taxa (i.e., the groups of organisms themselves) from issues about the realism of *categories* (a second-level issue). This sets the stage for considering recent debates over Mark Ereshefsky's (1998) "pluralistic anti-realism" and the "pluralistic realism" of philosophers like Philip Kitcher (1984) and John Dupré (1993). Devitt argues that the clash between these views is merely apparent: at stake is not the mind-independent existence of species, but rather whether species categories have a sufficiently robust explanatory significance compared to other scientific kinds. Devitt's suggestion is that the plausibility of the pluralist position with respect to the species problem is evidence of their minor explanatory role. He concludes by arguing that higher taxa play an even more modest explanatory role and, thus, that the Linnaean hierarchy should be dispensed with.

We noted above the controversy about biological essentialism. In his essay, "Three Ways of Resisting Essentialism about Natural Kinds" (chap. 9), Bence Nanay argues that contemporary biological practice decisively legislates against it. He notes first that essentialism about biological kinds involves three central tenets: that all and only members of a certain kind possess a common essence, that such *real* essences give rise to the *nominal* essences of a kind, and that essences facilitate our inferential practices by causing the co-occurrence of the various superficial properties associated with the kind. The first tenet seems to commit the essentialist to the existence of property-*types*. Thus one could resist it by adopting nominalism about properties. This way of arguing, as Nanay remarks, needn't carry much weight—especially if it is motivated by controversial metaphysical rather than biological commitments. Instead, he argues that we should see Ernst Mayr's influential (now nearly ubiquitous) idea of the biological realm being best described by "population thinking" as pushing us toward nominalism about property-types. This move puts Nanay in position to block the second and third tenets of kind-essentialism as well: property-types play no causal role in evolution; they are statistical abstractions. As such, they cannot explain or facilitate anything—contra the second and third tenets.

Taking the prize for best title, Neil Williams's essay "Arthritis and Nature's Joints" (chap. 10) attempts to throw another log on essentialism's funeral pyre. Many diseases, he argues, seem poorly accommodated by essentialism. Rheumatoid arthritis, for example, is presently defined in an exclusively clinical way (as presenting with four of seven diagnostic features). Now while it might turn out that these symptoms possess a common cause, it seems a bit implausible to claim that if they are *not* we should be forced to relinquish our practice of construing arthritis as a single disease kind. Williams draws upon the resources of Boyd's homeostatic property cluster account of kinds in order to make sense of disease kinds. In many ways, diseases seem an ideal test-case for the HPC account. Williams essay thus contributes both to our understanding of disease classification and an apparently flexible approach to natural kinds.

Species *taxa* play a key role in predicting how populations evolve. The methods employed to carry out such predictions, however, are not free from theory-laden assumptions. In his essay, "Predicting Populations by Modeling Individuals" (chap. 11), Bruce Glymour addresses the so-called "dynamic" and "statistical" interpretations of evolutionary theory, showing that they mistakenly take their outcomes to model populations while they are in fact modeling individuals. Glymour argues that the central concept at stake in predicting populations is selection. This is measured by monitoring either selection differentials or selection gradients, where the former is understood as the difference in fitness among classes of individuals. When considering this method, the way 'fitness' is defined assumes a central role; the model of selection is, in this case, a population genetic

model. The latter is a more complex notion, tracing the probability that a certain trait has of causing modifications in phenotypic or genotypic traits—selection gradients are defined at the individual level and they do not depend on fitness. When adopting this method, the model of selection will be tailored to specific populations, monitoring the causes of survival and reproductive success for its individuals. Glymour argues that the method of following selection gradients has epistemic advantages over methods based on selection differentials, as the former can more easily account for differences at a higher level (populations) in terms of differences at the lower level (individuals).

Another essay in the volume regarding the species problem, Jason Rheins's "Similarity and Species Concepts" (chap. 12), focuses on the role the similarity relation plays in sorting out species. Rheins's argument starts with a characterization of the similarity relation: since it is always relative to a respect or parameter, similarity is a more ductile theoretical tool than sameness. Rheins then introduces the metaphysical distinction between *immoderate* and *moderate* realism. The first envisages that any universal trait is existentially independent of the existence of any individual. On this view, universals may be said to exist as unrepeatable entities, which are numerically one and the same. The other form of realism, by contrast, sees universals as existing immanently in individuals. A universal cannot exist independently of the existence of some individual which instantiates it. And when the same universal is found in more than one individual it is because we have a repetition of instances. After introducing realist versions for three of species concepts-biological, ecological, and evolutionary-Rheins argues that the similarity relation is more suitable than simple qualitative sameness in accommodating such views. Indeed, according to Rheins, the fact that species are divided by a similarity relation does not entail that they are not real. Species concepts based on similarity are consistent with moderate realism based on an objective type of qualitative similarity, whose specifics vary from case to case and provide us with a satisfactory explanatory and predictive power.

The effects of the way organisms are classified into species are felt not only in biological circles but—most remarkably—in ethics as well. In their essay, "Species Concepts and Natural Goodness" (chap. 13), Judith Crane and Ronald Sandler discuss Philippa Foot's account of natural goodness, according to which an organism's worth is based on the potential it has for flourishing in ways that are proper for members of its species. Endorsing a pluralist conception of species, Crane and Sandler explore how well Foot's account sits with our biological findings and their most direct

philosophical consequences. After introducing the various species concepts that have been advanced by biologists and philosophers of science, the authors argue that Foot's account rests on what they label the axiological species concept (ASC). Central to this is the idea of "life form"-clearly reminiscent of Aristotelian doctrines and often regarded as synonymous with "species"—which expresses those traits that are distinctive of the way in which members of a given species live. Although ASC is ultimately deemed a viable species concept, Crane and Sandler argue that its endorsement needs to be backed up by normative commitments that are foreign to biology, such as those coming from ethology, from the thesis that vice and virtue involve emotions and desires (beyond physiological phenomena), or from the conviction that ethical norms may apply across (very) different environments and cultures. Thus, a natural goodness approach cannot be justified only on the basis of biological findings, but rather calls for some meta-ethical and normative commitments that are independent of them.

The volume concludes with an essay by Kadri Vihvelin, "How to Think About the Free Will/Determinism Debate" (chap. 14), which considers a lurking issue in the natural kinds business. Suppose that we sharpen our conceptual cutlery so much that we attain an accurate knowledge of all the joints of reality and, hence, of the laws governing them. Regardless of whether such laws are probabilistic, we might then be in a position to predict, for any instant of the world, what the next future instant can be like, in a way which is independent of the agents' deliberations. This is a way of capturing the idea that nature might unfold *deterministically*. On the other hand, the way in which we represent (most of) our actions assumes that, for any of those actions, there is a metaphysical possibility of choosing whether or not to do it—in these cases we represent ourselves as free agents. But if determinism is true, this representation is false. A certain variety of natural kinds realism thus seems to clash with the idea that we are free agents. According to Vihvelin, the problem of free will versus determinism is indeed the problem of explaining whether this apparent conflict is genuine. In her essay, she first discards a number of misguided ways in which free will and determinism have been conceived. According to her, the problem of free will versus determinism stems from two obvious facts: first, that determinism prima facie denies that natural kinds realism and freedom of the will are compatible; and second, that indeterminism prima facie leaves room for the two being compatible. Vihvelin's proposal reconciles these facts. In her view, determinism is compatible with free will, as freedom does not rest on an agent's actually
doing something, but on her ability to so act. In other words, we are free any time we are counterfactually *able* to do otherwise, even if we do not exercise such an ability.

Acknowledgments

Some of the material in this introduction has been adapted from Slater 2005. The authors would like to thank Michael O'Rourke and three anonymous referees for their comments on previous drafts.

Notes

1. Hacking calls this question "gentle" to differentiate it from a sterner one: "can natural kinds be characterized by essential properties? The gentle question is about what there is, the stern one, about what must be" (1990, 135).

2. There are, among others, nomic-necessitation approaches commonly associated with the work of Dretske (1977), Tooley (1977), and Armstrong (1983), best-systems approaches associated with Ramsey (1978) and Lewis (1973), primitivist approaches (Carroll 1994; Maudlin 2007), subjunctive approaches (Lange 2000), essentialist approaches (Bird 2007; Ellis 2001), and eliminativist approaches (Cartwright 1980, 1999; Giere 1999; van Fraassen 1989).

3. Hempel was quick to point out that it merely *seemed* wrong: for the statement that all ravens are black is, in a sense, a statement not just about ravens, but about the entire universe.

4. He rejected the thought that natural kinds would serve any permanent role in scientific investigation for precisely this reason, being somewhat cautious of a theory-neutral notion of overall similarity—but that's another story.

5. It should be noted that "essence" acquires a very different meaning in other philosophical contexts, most notably: in Hegel's philosophy, where it stands for the deeper structure of reality, in contraposition with the superficial "phenomena"; in Husserl's phenomenology, in which essences are the content of eidetic intuitions; and in the Existentialist tradition, where it is bestowed a negative connotation in opposition to "existence." We shall, however, leave these uses of the term aside as they are not relevant to the philosophical context here under consideration.

6. Even an identical twin or a doppelgänger would not be him.

7. It includes among its advocates well-known philosophers such as David Wiggins, Michael Dummett, John Wallace, and Robert Ackermann.

8. Because they are unnecessary for present purposes, we will ignore here some of the distinctions between kinds of sortals, such as the distinction between "phase" and "proper" sortals. The first is predicated of a phase of an entity—for example,

"child" is predicated of a phase of a human being's life, namely their childhood. The latter is predicated of the entire life of an entity—for example, "person" is predicated of a human being for his entire life.

9. Thus, Wiggins (1979, 1986) seems to defend (i), Brody (1980) defends (ii), and John Wallace (1965), Robert Ackermann (1969), and Jonathan Lowe (1998) defend (iii)—a view that Wallace attributes to Frege.

10. Natural laws are not *logically* necessary: there is no contradiction or incoherence in imagining that, say, the law of universal gravitation is false. And yet, it seems clear that laws are somehow "more necessary" than mundane, accidental facts (e.g., that all the coins in my pocket are made of copper). See Lange 2009 for an accessible and insightful discussion into this issue.

11. Abbott (1997), LaPorte (1998), and Brown (1998) discuss the impurity problem. For critical discussion of Putnam's views of natural-kind term reference, see Zemach 1976; Mellor 1977; Devitt and Sterelny 1987; LaPorte 1996; and Stanford and Kitcher 2000.

12. Kitcher also provides an illustration of the pull of genetic essences. "Structural explanation" often involves investigation into the genetic basis of morphological features—for example, viral protein sheaths. "We learn that the features that originally interested us depend upon certain properties of the viral genome. At this point our inquiries are transformed. We now regard viruses as grouped not by the superficial patterns that first caught our attention, but by similarities in those properties of the genome to which we appeal in giving our explanations. . . . The achievement of an explanatory framework goes hand in hand with a scheme for delineating the 'real kinds' in nature" (1984, 321–322). Kitcher admits, of course, that this example "mixes science with science fiction"—as we shall see, the general strategy faces other serious problems.

13. See Kornblith 1993, Wilson 1999, and Chakravartty 2007.

References

Abbott, B. 1997. A note on the nature of water. Mind 106:313-319.

Ackermann, R. 1969. Sortal predicates and confirmation. *Philosophical Studies* 20:1–4.

Andreasen, R. O. 1998. A new perspective on the race debate. *British Journal for the Philosophy of Science* 49:199–225.

Aquinas, T. 1997. *De Ente et Essentia.* Trans. R. T. Miller. Internet Medieval Source Book. http://www.fordham.edu/halsall/Sbook.html.

Aquinas, T. 1998. *De Principia Naturae*. Trans. J. Bobik. Notre Dame: University of Notre Dame Press.

Armstrong, D. M. 1983. *What Is a Law of Nature?* Cambridge: Cambridge University Press.

Bealer, G. 1987. The philosophical limits of scientific essentialism. *Philosophical Perspectives* 1:289–365.

Bigelow, J., B. Ellis, and C. Lierse. 1992. The world as once of a kind: Natural necessity and laws of nature. *British Journal for the Philosophy of Science* 43:371–388.

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

Boyd, R. 1988. How to be a moral realist. In *Essays on Moral Realism*, ed. G. Sayre-McCord. Ithaca: Cornell University Press.

Boyd, R. 1991. Realism, anti-foundationalism, and the enthusiasm for natural kinds. *Philosophical Studies* 61:127–148.

Boyd, R. 1999. Homeostasis, species, and higher taxa. In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge, Mass.: MIT Press.

Brody, B. 1980. Identity and Essence. Princeton: Princeton University Press.

Brown, J. 1998. Natural kind terms and recognitional capacities. *Mind* 107 (426): 275–303.

Carroll, J. W. 1994. Laws of Nature. Cambridge: Cambridge University Press.

Cartwright, N. 1980. Do the laws of physics state the facts? *Pacific Philosophical Quarterly* 61:75–84.

Cartwright, N. 1999. The Dappled World. Cambridge: Cambridge University Press.

Chakravartty, A. 2007. *A Metaphysics for Scientific Realism*. Cambridge: Cambridge University Press.

Devitt, M. 2008. Resurrecting biological essentialism. *Philosophy of Science* 75 (3): 344–382.

Devitt, M., and K. Sterelny. 1987. Language and Reality. Cambridge, Mass.: MIT Press.

Dretske, F. 1977. Laws of nature. Philosophy of Science 44:248-268.

Dupré, J. 1981. Natural kinds and biological taxa. Philosophical Review 90:66-90.

Dupré, J. 1993. The Disorder of Things. Cambridge, Mass.: Harvard University Press.

Ellis, B. 2001. Scientific Essentialism. Cambridge: Cambridge University Press.

Ellis, B., and C. Lierse. 1994. Dispositional essentialism. *Australasian Journal of Philosophy* 72:27–45.

Ereshefsky, M. 1998. Species pluralism and anti-realism. *Philosophy of Science* 65:103–120.

Feldman, F. 1973. Sortal predicates. Noûs 7:262-282.

Fine, K. 1994. Essence and modality. Philosophical Perspectives 8:1-16.

Fine, K. 1995. Senses of essence. In *Modality, Morality and Belief*, ed. W. Sinnott-Armstrong. Cambridge: Cambridge University Press.

Ghiselin, M. 1974. A radical solution to the species problem. *Systematic Zoology* 23:536–544.

Giere, R. 1999. Science without Laws. Chicago: University of Chicago Press.

Glasgow, J. 2009. A Theory of Race. New York: Routledge.

Goodman, N. 1983. *Fact, Fiction, and Forecast,* 4th ed. Cambridge, Mass.: Harvard University Press.

Griffiths, P. 1997. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.

Griffiths, P. 1999. Squaring the circle: Natural kinds with historical essences. In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge, Mass.: MIT Press.

Hacking, I. 1990. Natural kinds. In *Perspectives on Quine*, ed. R. B. Barrett and R. F. Gibson. Cambridge: Blackwell.

Hacking, I. 1995. Entrenchment. In *Grue! The New Riddle of Induction*, ed. D. Stalker. LaSalle: Open Court.

Hacking, I. 2005. Why race still matters. Daedalus 134:102–116.

Hempel, C. 1945. Studies in the logic of confirmation. *Mind* 54:1–26, 97–121.

Hempel, C. 1965. *Fundamentals of Taxonomy*. In his *Aspects of Scientific Explanation*. New York: The Free Press.

Hull, D. 1978. A matter of individuality. Philosophy of Science 45:335–360.

Kahn, M. 1995. The Tao of Conversation. Oakland: New Harbinger.

Kitcher, P. 1984. Species. Philosophy of Science 51:308–333.

Kitcher, P. 1999. Race, ethnicity, biology, culture. In *Racism*, ed. L. Harris. Amherst: Humanity Books. Reprinted in his *In Mendal's Mirror*. Oxford: Oxford University Press, 2003; citations are to the 2003 version.

Kornblith, H. 1993. *Inductive Inference and its Natural Ground*. Cambridge, Mass.: MIT Press.

Kripke, S. 1972. Naming and necessity. In *Semantics of Natural Language*, ed.G. Harman and D. Davidson. Dordrecht: Reidel.

Kripke, S. 1980. Naming and Necessity. Cambridge, Mass.: Harvard University Press.

Lange, M. 1995. Are there natural laws concerning particular biological species? *Journal of Philosophy* 92 (8):430–451.

Lange, M. 2000. Natural Laws in Scientific Practice. New York: Oxford University Press.

Lange, M. 2004. The autonomy of functional biology: A reply to Rosenberg. *Biology* and *Philosophy* 19:93–109.

Lange, M. 2009. Laws and Lawmakers. New York: Oxford University Press.

LaPorte, J. 1996. Chemical kind term reference and the discovery of essence. *Noûs* 30 (1):112–132.

LaPorte, J. 1998. Living water. Mind 107:451-455.

LaPorte, J. 2004. Natural Kinds and Conceptual Change. Cambridge: Cambridge University Press.

Lewis, D. 1973. Counterfactuals. Cambridge, Mass.: Harvard University Press.

Lewis, D., and R. Langton. 1998. Defining 'intrinsic'. *Philosophy and Phenomenological Research* 58:333–345.

Lewontin, R. C. 2000. The Triple Helix. Cambridge, Mass.: Harvard University Press.

Lowe, J. E. 1998. *The Possibility of Metaphysics: Substance, Identity, and Time*. Oxford: Clarendon Press.

Maudlin, T. 2007. The Metaphysics within Physics. New York: Oxford University Press.

Mayr, E. 1987. The ontological status of species: Scientific progress and philosophical terminology. *Biology and Philosophy* 2:145–166.

Mellor, D. H. 1977. Natural kinds. *British Journal for the Philosophy of Science* 28:299–312.

Mishler, B. D., and M. J. Donohue. 1982. Species concepts: A case for pluralism. *Systematic Zoology* 31 (4):503–511.

Mitchell, S. 2000. Dimensions of scientific law. Philosophy of Science 67:242-265.

Nagel, E. 1961. *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace, and World.

Okasha, S. 2002. Darwinian metaphysics: Species and the question of essentialism. *Synthese* 131:191–213.

Pigliucci, M., and J. Kaplan. 2003. On the concept of biological race and its applicability to humans. *Philosophy of Science* 70:1161–1172.

Popper, K. 1950. *The Open Society and Its Enemies*, rev. ed. Princeton: Princeton University Press.

Putnam, H. 1975. The meaning of 'meaning'. In his *Mind, Language, and Reality: Philosophical Papers*. Cambridge: Cambridge University Press.

Quine, W. v. O. 1969. Natural kinds. In his *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Ramsey, F. 1978. Foundations. London: Routledge & Kegan Paul.

Rosenberg, A. 1987. Why does the nature of species matter? *Biology and Philosophy* 2:192–197.

Sider, T. 2009. Ontological realism. In *Metametaphysics*, ed. D. Chalmers, D. Manley, and R. Wasserman. New York: Oxford University Press.

Slater, M. H. 2005. Monism on the one hand, pluralism on the other. *Philosophy of Science* 72 (1):22–42.

Stanford, P. K., and P. Kitcher. 2000. Refining the causal theory of reference for natural kind terms. *Philosophical Studies* 97 (1):99–129.

Strawson, P. F. 1959. Individuals: An Essay in Descriptive Metaphysics. London: Methuen.

Tooley, M. 1977. The nature of laws. Canadian Journal of Philosophy 7:667-698.

van Fraassen, B. 1989. Laws and Symmetry. Oxford: Clarendon Press.

Wallace, J. 1965. Sortal predicates and quantification. *Journal of Philosophy* 62:8–13.

Walsh, D. M. 2006. Evolutionary essentialism. British Journal for the Philosophy of Science 57 (2):425–448.

Watson, B. (trans.). 2003. *Zhuangzi: Basic Writings*. New York: Columbia University Press.

Wiggins, D. 1979. Sameness and Substance. Oxford: Blackwell.

Wiggins, D. 1986. On singling out an object determinately. In *Subject, Thought, and Context*, ed. P. Pettit and J. McDowell. Oxford: Clarendon Press.

Wilson, R. A. 1999. Realism, essence, and kind: Resuscitating species essentialism? In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge, Mass.: MIT Press.

Woodward, J. 2001. Law and explanation in biology: Invariance is the kind of stability that matters. *Philosophy of Science* 68:1–20.

Zack, N. 2002. Philosophy of Science and Race. New York: Routledge.

Zemach, E. 1976. Putnam's theory on the reference of substance terms. *Journal of Philosophy* 73 (5):116–127.

2 Induction, Samples, and Kinds

Peter Godfrey-Smith

1 Introduction

This essay will criticize a familiar package of ideas about "inductive" inference, and use that criticism to motivate a different package.

"Induction" is understood here as a pattern of argument or method used to answer questions of the form: "how many *F*s are *G*?" This question is understood as one about a proportion or frequency. So it could also be expressed by asking "what is the rate of *G* in the *F*s?" The question "Are all *F*s *G*?" is a special case. Examples of such questions include:

- 1. How many teenagers smoke?
- 2. How many ravens are black?
- 3. How many emeralds are green?
- 4. How many people in this room are third sons?

5. How many organisms have the amount of bases cytosine and guanine equal, and the amount of adenine and thymine equal, in their DNA? 6. How many electrons have charge of approximately -1.6×10^{-19} coulombs?

There are lots of ways of answering such questions. In "induction," the questions are answered by noting the relation between *F* and *G* in observed cases and making some sort of extrapolation or generalization. This is presumably done with the aid of background knowledge. But the approach taken is one in which the number of *F*s seen is supposed to be epistemically important. The classic case of inductive inference is the one where *all* observed *F*s are found to be *G*, and this is used to conclude that all the unobserved ones are as well.

Here is a view that many people hold: induction is often rational, lest factual knowledge collapse. But as we learned from Nelson Goodman, the *F* and the *G* in a good induction *can't be just anything*. We need a constraint,

P. Godfrey-Smith

probably some sort of "naturalness" constraint, on the kinds or predicates involved. Many philosophers would agree with this even though they do not agree what the constraint is or where it comes from.

I will argue that for the kind of inference that counts as induction in the above sense, and that can be generally justified, the predicates used *can* be just anything—or near enough to anything. In particular, a "naturalness" constraint has no basis. But naturalness does have a role in *another* kind of inference that can answer "how many *Fs* are *G*?" questions. When we get to that point you might say that this other kind of inference is also induction, so we run into an issue that is a bit terminological. But the overall view I will defend is that the familiar philosophical concept of "induction" has conflated two kinds of inference, each of which is successfully exploited by science. For each of the two inference patterns, an account can be given of its *in-principle reliability*. That account is a kind of philosophical justification. The package usually known as "induction" does not have that kind of justification, however. It combines elements from each method without combining parts that give rise to an in-principle reliable combination in its own right.

I also argue for some meta-epistemological ideas. Older work in normative epistemology often included trying to show why a method will or won't *work*—showing whether the method is responsive to the world in a way that leads to us getting the right answers. There was a retreat from this procedural approach in the twentieth century, when much work focused on logical relationships. In the case of induction, there was also a further role for Goodman's *Fact, Fiction, and Forecast* (1955). Goodman seemed to show the complete foolishness of procedural approaches that had been taken up to that point, and also seemed to show the coherence of a quite different orientation to the problem. This reorientation holds that the normative analysis of induction works by means of a mutual adjustment of judgments about cases and judgments about general rules, with the aid of intuitions we have on both matters. This is the method Rawls (1971), crediting Goodman, christened as "reflective equilibrium."

This essay will criticize conclusions reached by that approach and argue for the value of another. In some ways, the alternative offered is a revival of the earlier style, focused on procedures. But it is not foundationalist, and includes a kind of "model-building" orientation. In the case of induction, there are continuities between model-building that has a philosophical level of ambition and abstraction, and model-building of a more practical kind within science. To some extent, problems in the epistemology of induction are exchanged for problems about how model-like theoretical constructs relate to the world and our situation in it. But there is net progress. Reichenbach is a particularly important precursor of the approach I defend here.¹

2 Goodman's Problem and Naturalness Constraints

Suppose we are answering a "how many Fs are G?" question. We note how many Fs we have seen that are G, and extrapolate. We feel better about it if we have seen many Fs rather than a few.

Is this approach justified? Not always. We confront Goodman's "new riddle of induction" (Goodman 1955; Stalker 1994). Though in many cases such extrapolations are surely rational, the predicates used in our argument apparently can't be just anything if the inference is to be a good one. This can be illustrated by introducing a "grue" predicate. Something is grue if and only if it either has been previously observed and is green, or has not been previously observed and is blue. If all observed emeralds are in fact green, we find we can use this data as the basis for two conflicting inferences about the future: the familiar inference leading us to expect newly observed emeralds to be green, and also an inference leading us to expect them to be blue. After all, the previously observed emeralds, being green, are also all grue. If we extrapolate grueness to the presently *unobserved* emeralds, this leads us to expect that those will turn out to be blue. We have encountered an apparent collapse of inductive methods: with suitable choice of grue-like predicates (in F and/or G position), just about anything in our past observations can be used to support just about any hypothesis about the future.

Goodman gave other examples of bad inductions, using less exotic language. One uses the predicate "is a third son": if we find the first few people in this room are third sons, that does not give us reason to think that everyone in the room is a third son. What seems needed is an extra constraint. Perhaps inductive arguments have to use predicates that pick out "natural" properties or kinds (Quine 1969; Lewis 1983). Many different bases for such a constraint have been offered, ranging from the conventional to the metaphysical, and the idea of a naturalness restriction on predicates and sets has spread into many other parts of philosophy, always with Goodman's "grue," and the problems it causes for induction, as the motivating example.

A paper that went against this trend was Frank Jackson's "Grue" (1975). He argued that in a good induction the predicates used *can* be just anything, provided that an additional premise is true. The extra premise he

called the "counterfactual condition." An induction, for Jackson, takes information about objects that have *three* features—they are observed, they are *F*, and are *G*—and draws a conclusion about objects that are unobserved and *F*. A good induction looks like this:

J1. All Fs which are observed are G.

J2. If those Fs had not been observed, they would still have been G.

Therefore (fallibly):

J3. All Fs are G.

The extra premise is true for F = emerald and G = green, but not when G = grue. To say this relies on background knowledge—but background knowledge that, according to Jackson, is reasonably available and does not beg the question. Most importantly, there is no restriction on predicates per se. The predicates used can be anything at all, provided that premise J2 is true. It is *hard* for premise J2 to be true with some predicates in the G position, given normal choices for the "observed," "past," or "sampled" predicate. But naturalness is not an issue, and linguistic "entrenchment" in Goodman's sense is not an issue; those constraints drop out of the story.

Jackson's argument is the beginning of the answer to Goodman's problem. It is not the whole answer. This is for both internal, detailed reasons and more general ones. First, the proposal does not handle all grue-like predicates (especially "emerose" cases). Some were handled in a modified formulation given by Jackson and Pargetter in 1980, and others were not. These problems are discussed in an earlier paper of my own (Godfrey-Smith 2004). Using an idea due to Alexis Burgess we can make more progress. In the present discussion I will ignore additional cases and assume we only have to deal with the original grue problem, where Jackson's 1975 proposal seems to work. But we can also ask a more external question: what is the *status* of Jackson's proposal? If it is right, what makes it right? And what does it then achieve for us?

What Jackson says is that if we put inductions into his form, we see that the good ones are those in which his premises are true, and the bad ones are those where some premise is false. His proposal handles the cases, and it is also the basis for a plausible story about what has gone wrong in the bad inductions. When premise J2 is false, it is *only because those Fs were observed* that they were *G*; their *G*-ness is a special case, not something that can be extrapolated. The methodology here is the usual twentieth-century one: we take logic as far as it goes, and then use intuitions. But suppose

someone now asks: if I go through the world applying the Jackson schema and drawing conclusions, will I do well or badly? What will be its consequences? Can you show me that at least in principle, or in *some* relevant range of circumstances, it should lead me to truth rather than error? Nothing about this is said by Jackson, or could be said within his framework. And in fact we can do more. To see how, let us forget emeralds, and look at the kind of question that would actually be answered using simple extrapolation from a sample.

Suppose you want to know how many teenagers smoke. The obvious way to answer this question is to collect a random sample of teenagers, find the rate of smoking in the sample, and extrapolate to the larger teenage population, in a way that is guided by statistical measures of likely error. Why should we do this rather than something else? Why not find the proportion in the sample and extrapolate half that proportion, or one minus that proportion? (This second option would be a kind of counter-induction.²) *Not* because of an equilibrium between intuitions—or at least, not at this stage in the analysis. Instead, we have a statistical model of why the procedure is in principle a reliable one. The model tells us how samples of different sizes will be distributed in relation to the actual properties of the population being sampled. It tells us when, and the extent to which, the properties of a sample are reliable indicators of the properties of the underlying population.³

I have said that the question about teenage smokers could be answered using inference from a sample. But there are various ways this method might fail. Perhaps you cannot collect a random sample, as the smokers tend to avoid you; or perhaps teenagers will not tell you the truth. There is also a third, more unlikely possibility. Perhaps being asked the question tends to make some teenagers immediately take up smoking. So they truthfully answer that they smoke, but only because they were asked. The process of data-gathering is interfering with the objects you are observing, in a way that makes observed cases an unreliable guide to unobserved cases. This is not a case of "selection bias"; we can assume that the original collection of teenagers asked about their smoking really was a random sample. Some statisticians call this a "Hawthorne effect," after a famous case in the 1930s.⁴ It also has a kinship with the notion of a "confounding" variable," although that term is usually applied in the context of causal inference rather than estimation. The term "observation selection effect" is used in some literature to cover various phenomena, including sample bias, the confounding-like phenomenon discussed here, and maybe others. In any case, that special relationship between surveying and smoking

P. Godfrey-Smith

would make the method fail. The problem has nothing to do with "nonprojectible" predicates; it has to do with the process of collecting our sample and some unwelcome causal relations.

There is a close relation between this phenomenon and the grue problem. The case of the grue emeralds features a *non-causal analog* of the problematic dependence relations seen in the smoking case. Just as the teens surveyed would not be smokers if we had not asked them about it, the emeralds would not have been grue—would not have counted as grue—if we had not observed them. The process of observation is interfering with the properties we are interested in. The semantics of "grue" turn observation itself into a confounding variable. Or to adapt a piece of metaphysical jargon, if the teenage-smoking problem were like a case of confounding, this is a case of *Cambridge-confounding*.⁵ When we try to state a condition that would rule out such problematic dependence relations, and say it in abstract counterfactual terms, we will find ourselves saying what Jackson said in his paper on grue.

Neither Goodman nor Jackson said anything about randomness. They also offered no account of why following arguments that meet their requirements will do us any good. In cases where you *can* collect a random sample, however, there is a model that tells you how and why certain projections will be reliable. The model also tells us that you cannot use random samples to answer grue-questions with grue-observations in the same way you can use samples to answer green-questions with green-observations. The problem arises *as a feature of procedures*. If you had a "non-interfering" way of sampling the emerald population, you *could* estimate the proportion of grue emeralds. How hard this is depends on the exact "grue" predicate used.

The view that emerges diverges sharply from intuitions. Suppose we can randomly sample, and we are keeping an eye out for the confounding role of observation that arises with some predicates. Let us go through the world, sample, and project. The *G*s that the model allows us to project include G = (green or smaller than a fingernail). They also include: G = (jadeite or nephrite), G = (green and born in Iceland), and G = (green or identical to the number 4). This looks all wrong; our intuitions revolt. But the model of reliable estimation from samples allows it. There is no need for the predicates in an inference to be of a kind that can figure in natural laws. Returning to Goodman's "third son" case, there is really no problem here. If there is a rate of third-sonness in this room, and you collect a random sample of people and assay it for third-sonness, you *can* draw reliable inferences, depending on sample size and so on. The same is

true on the *F* side—the population or class whose rate of *G*-ness you are interested in. It can be as arbitrary as you like, as long as it can be sampled.

3 A Second Form of Inference

The previous section described one kind of inference that meets the criteria given earlier for being "inductive," and discussed why the inferences are justifiable. This approach is inapplicable to many induction-like inferences, however. Lots of collections we are interested in cannot be randomly sampled. "Random sampling" here means that every member of the population you are drawing conclusions about has the same chance of making its way into the sample. So a collection containing future individuals (future ravens, future DNA molecules, future third sons) cannot be randomly sampled. It surely seems that we can sometimes gain knowledge of generalizations in such cases, however. In these cases, our observed instances are not a sample drawn from a total population, but are more like a subpopulation "attached" to it. The past, in particular, is attached to the future but not drawn from it.

In the sampling cases, the power of randomness is what gives us a "bridge" from observed to unobserved. In the second kind of case, the bridge—when there is one—is very different. If we want to make inferences about a population that cannot be sampled, we must ask: what *kind* of collection is this? Are these objects the products of a common origin? Do they have a common internal structure? What sort of causal relationship is there likely to be between properties we are projecting from and properties we are projecting to? There need not be "laws of nature" overtly on the scene here, but we are basing the inference on some kind of natural connection—some combination of laws, mechanisms, and etiologies.⁶ I will say less about these inferences than I did about the first category, and will focus primarily on the contrasts between the two.

In the second category of inference, something like the "naturalness" of kinds and properties is central. It is *so* central, in fact, that a crucial feature of the first category drops out—that is, the *number of cases* observed. In inferences from random samples, numbers are epistemically significant. Larger is always better. In the case of inferences of the second sort, based on causal structure and kinds, this is not so. In the purest examples of this sort of investigation, *one* instance of an *F* would be enough, in principle, if you picked the right case and analyzed it well. Ronald Reagan is supposed to have said "once you've seen one redwood, you've seen them all."⁷ When something like this is true, it is a powerful basis for inference. In practice,

one is usually not enough. Numbers often do play a role, but this role is different in character from what we find in the first category of inference.

In simple cases of these inferences based on causal structure and kinds, the Fs all have some feature in common which has a causal relation to G of a kind that makes generalization possible. What the investigation is really aiming to do is assess some sort of dependence relation, which might be expressed by a conditional, linking F and G. This need not describe a causal relation from F (or an underlying feature of F things) to G; it might run from G to F (all redwoods have such-and-such in their DNA) or from a common cause to each of them. Sometimes the causal basis for the dependence might be known, while in others there may just be reason to think there is *some* basis of the right kind. Either way, if such a dependence can be established, it does not matter whether the class of Fs can or cannot be sampled, is small or large, and so on. Seeing a number of cases of F can be helpful, because it may shed light on how F and G are related, and on how the *F*–*G* connection is affected by variation in circumstances. But you may be able to get this knowledge just as well or better by looking at other things, which are not F, and mere repetition of Fs which are G-mere weight of numbers with respect to the F-G association-does no good at all.

One need not be looking for complete uniformity to engage in this second approach to extrapolation. An unpacking of a few representative emeralds, or melanomas, might show that there will be a particular kind of diversity with respect to the properties you are interested in. If you find that emeralds or melanomas are diverse in nomologically comprehensible ways, it will often make sense to undertake separate investigations of each subtype. But it might be that the investigation stops with the claim that given the make-up of *Fs*, some range of alternative features *G*, *H*, . . . (etc.) may each be found with particular probabilities.

At this point a terminological issue arises. Is this second kind of inference also "induction"? I said in the introduction that I was reserving the category "induction" for cases in which the number of *Fs* observed is supposed to be epistemically relevant. I said just above that in the second kind of investigation, having large numbers of *Fs* can be helpful. There is certainly a broad sense in which this helpfulness is "epistemic," but I distinguish it from the stronger sense seen in inferences from random samples. In inference from samples, support for the conclusion goes *via* weight of numbers; there is no way that a sample of ten could in principle do what a sample of one hundred can do, if only you could interact with those ten objects in a less noisy way and learn more about the relationships between their various features. In the second kind of inference, that *is* the case, and it is also true that many ways of seeing new instances are no help at all. So I regard the role of numbers in the second category of inference as *practical* rather than epistemic.

Some people might want to use the term "induction" differently, including both categories. The label itself does not matter much, so let us set aside the label and focus on the picture. That picture is one in which we can recognize two kinds of inference. The first is generalization from random samples. This form of inference has the following features: sample size matters, randomness matters, and "law-likeness" or "naturalness" does not matter. The second kind of inference is generalization based on causal structure and kinds. In these cases sample size per se does not matter, randomness does not matter, but the status of the kinds matters enormously. These two strategies of inference involve distinct "bridges" between observed and unobserved cases: one goes via the power of random sampling, the other via reliable operation of causes and mechanisms. Then we see that the philosopher's concept of induction, especially since Goodman, has often been a hybrid of these. Philosophers have supposed that the crucial category of inference is one in which (i) sample size matters, (ii) randomness is not an issue, and (iii) naturalness of kinds does matter, but weakly. By "weakly" I mean that naturalness is used only in the exclusion of bad kinds and predicates, clearing the way for the weight of numbers to do its work. This combines elements of the two induction strategies, but does so in a way that includes no bridge from observed to unobserved. The link that exploits sampling is not available, nor is there the right kind of role for kinds and causal mechanisms. In good inferences based on causes and kinds, scrutiny of the F and G in question is not aimed at merely excluding bogus collections and pseudo-properties. After all, most combinations of a natural F and a natural G do not show any sort of stable association in which a few cases can be used to draw inferences about many.

So far I have been treating the mistake that has been made about induction as a philosopher's mistake, but perhaps this error has deeper roots. It may be that humans have inductive *habits* which correspond, more-or-less, to the philosophers' picture I am criticizing. That is, we may have habits in which weight of numbers is taken to have a general importance independent of randomness, and a "can't-be-just-anything" principle also modulates how the experience of numbers affects us. Anthropologists have recently become interested in what might be pancultural habits of induction, especially as applied to living things (Medin and Atran 1999). It may turn out that these habits have shaped quite specific features of philosophical treatments of induction. Perhaps this is so—but if so, these are just elements of our psychology, and of our "folk epistemology." These habits are likely to have been reliable enough in the practical domain in which they were developed—or rather, they are likely to have achieved a reasonably good balance with respect to reliability of various kinds, cost, and speed (Gigerenzer and Todd 1999). To show that is to show a kind of in-principle reliability, but of a very local kind. This would not, for example, show that these inductive habits are reliable within science and other contemporary epistemic endeavors.

I will finish this section with a historical note. Given the arguments I have made, it is interesting to look back at an exchange between Hans Reichenbach and John Dewey in Dewey's "Schilpp Volume" (1939). Reichenbach modeled all nondeductive inference on a kind of statistical estimation (not the same kind as that discussed here), and hence saw the number of observed cases as crucial. Dewey spurned traditional concepts of induction, especially with respect to the role of weight of numbers. He thought that in actual scientific generalization, everything hangs on the scientist's ability to find individuals which are representative of their kind. If we can do this, then one individual is often enough: the hard work goes into saying why a particular case should be representative. Reichenbach argued that Dewey did not appreciate the role of probability, and the significance of the possibility of convergence on a limiting value through repeated observations. For Reichenbach the key to projection lies there. I say that both were on the right track with respect to understanding some inferences in science, but that both were too inclined (ironically) to treat one kind of case as the key to all.

4 A Nominalist Challenge

This section discusses an objection to part of the argument presented above. This objection was raised (in discussion) by both Laura Schroeter and Ira Schall. Like Jackson, I claimed there is an asymmetry between green- and grue-inductions that has to do with the fact that observation "affects" the objects sampled with respect to *G*-ness in the grue case, but not in the green case. But is the asymmetry real? Goodman noted, after all, that "blue" and "green" can be defined in terms of "grue" and "bleen," as well as vice versa. He also argued that judgments of similarity are language-dependent. Counterfactual claims, on many views, are

dependent on similarity judgments. Putting these points together, it might seem that if we imagine someone starting out with a language that takes "grue" and "bleen" as basic, the asymmetry for them would go away or reverse. Such a person would apparently be entitled to insist that if our particular observed emeralds had not been observed they would still have been grue, so they would also have had to have been blue. Then, as Goodman held, for linguistically different agents different inductions will be acceptable.

To assess this argument, assume that we have two agents who each have a respective native language, normal English or grue-English, but who also speak the other's language. Their aim is to learn about emerald color by inspecting a random sample. The grue-English speaker picks up the first emerald, and notes that it is both green and grue. Then he asks: "If this thing had not been in the sample, would it still have been grue?" He is a native grue-speaker, but he also knows our words. So he can note to himself that if this thing had not been in the sample, then in order for it to have been grue, it would have had to have been blue. So he is wondering whether *this individual thing in front of him* would have been blue if it had not been one of the emeralds that happened to make its way into the sample—if we had picked up another instead.

Suppose he says that it *would* still have been grue, and hence blue, if unobserved. We ask him what the basis for this claim is. He might answer that two grue things are more similar, *ceteris paribus*, than two green things. He goes on to say that if we ask what things would have been like in a situation in which this emerald had never been observed, that situation would have been one in which this emerald was as similar to its actual state as the assumed difference from actuality permits. So it would have been grue, and hence blue.

If someone says this, we do reach a kind of standoff. But it is different from the standoff that is usually seen as taking hold. The standoff we reach is one that concerns each member of a collection of individual observed things, and how the agreed-upon properties of each thing are causally and counterfactually related. We can pick up each item and discuss its chemistry, the processes that formed it, and its history of interaction with us. For the grue-English speaker to say that each of the sampled emeralds would still have been grue if unobserved, he has to not only use a different *categorization* of things from us, but *also* has to have a large collection of very strange beliefs about chemistry, history, and light.

The situation is again like the teen-smoking case. Suppose you believe that when teenagers are asked about their smoking habits, many of them immediately take up the habit. This is a strange belief about a causal dependence; it is not just a strange way of categorizing things. When a model of sampling is applied in an actual-world case, its application will depend on factual beliefs about other matters. Agents with different beliefs on factual matters will put the same model to different use, and may make divergent projections. *Linguistic* difference is not sufficient to account for this, and neither is a different sense of similarity. Disagreements about causal and counterfactual dependence *will* account for it, but that is not surprising. The analogy between the grue emeralds and the fickle teenage non-smokers helps us to see this.

5 A Discussion of Cases

So far I have discussed two ways in which observed cases can reliably be used as a guide to unobserved cases, and contrasted both with how philosophers usually think of "induction." In this section, I look at how the strategies discussed above relate to some actual cases of generalization.

In actual epistemic practice, especially in science, we see a mixture of the two approaches, plus much more. We see a "mixture" in two senses. First, there are fairly clear examples of work done according to each strategy. Second, the answering of a single question may draw, explicitly or implicitly, on both strategies.

I listed some examples of this at the start of the essay. Let us return here to teen-smoking: here, random sampling is used. Collecting samples is feasible, and its limitations don't matter much; we don't expect a generalization to cover future or very distant cases. The class of teenagers is so locally diverse and causally complicated that it would be fruitless to mount a nomological or mechanistic assault on the problem. In Massachusetts in 2007, about 18 percent of high school students smoked;⁸ this figure was based on a random selection of about 3,000 students.

Close to the other end of the scale, we have the charge on the electron. This was found to be roughly -1.6×10^{-19} coulombs by Robert Millikan in the oil-drop experiment (Millikan 1911, 1917; Franklin 1997). At the time of Millikan's work it was not agreed that there was a fundamental unit of charge for an electron. Some researchers, such as Felix Ehrenhaft, held or suspected that the charge may vary continuously. So Millikan was not just setting a parameter which everyone agreed must hold universally. Millikan suspended tiny individual oil-drops by balancing them between forces due to gravity and an electric field. He then turned the field on and off to see how fast the drops responded. He used that measurement to calculate the

There was no question of his sample being a random one; it was not even a random sample of electrons in Illinois. And Millikan was not looking to support his estimate by sheer weight of numbers. He published results from fifty-eight drops (excluding at least forty-nine others, for reasons that have been queried). His aim was to get a few well-behaved cases that would show the phenomenon clearly and permit a measurement. In the context of background knowledge, the consistent finding of multiples of one number was taken to establish uniformity—not just for the state of Illinois, but for the universe.

This case is very far from the teen-smoking one, but it is not the simplest possible case of a "seen-one, seen-them-all" inference. Millikan was not able to get at individual electrons, and he needed to collect a number of cases to make the argument that there was a *natural* unit being detected. A more overtly mixed case is that of "Chargaff's rules" regarding the composition of DNA. The most important of these rules states that the amount of C (cytosine) is the same as the amount of G (guanine), and the amount of T (thymine) is the same as the amount of A (adenine), in all DNA. Chargaff did this work before the structure of DNA had been discovered by Watson and Crick, and there was no obvious reason why the relation should hold. Chargaff began publishing versions of this finding in the late 1940s, and followed up in the early 1950s. In early papers he gave results for nine kinds of organisms, scattered throughout the major groups, adding a few more later. In comparison to Millikan, Chargaff was looking for some coverage of the total field of living organisms-he was looking outside the genetic analogue of Millikan's Illinois. His list included the genetic material of bacteria, yeasts, wheat, sea urchins, and humans. He took it that equal proportions in one organism, or even in *all* his sample organisms, could be a special case or an accident. In his articles at the time he was quite cautious: "It is noteworthy-whether this is more than accidental, cannot yet be said—that in all desoxypentose nucleic acids examined thus far the molar ratios of total purines and total pyrimidines, and also of adenine to thymine and of guanine to cytosine, were not far from 1" (1951, 206).9 Only with Watson and Crick's work, a few years later, did it become clear that the rule *had* to hold universally, given the structure of the DNA molecule itself.¹⁰

How reasonable would it have been to extrapolate on the basis of Chargaff's data alone? To do this, it seems that one would need to argue that *if* there was variation in DNA composition, Chargaff's sample ought

to have revealed it. There seem to be two ways in which this might be done. One is to mount an argument from features of the biochemistry of life, but no argument of this kind was available in Chargaff's time. The other might be to argue that though the sample was not taken in a truly random way, it might reasonably be taken to have some of the features of a random sample. It is as if someone chose teenagers for a smoking survey who have exactly five letters in their first name. This is not a random sample of American teenagers, but it seems likely to be similar to a random sample in some ways, and is not biased with respect to the smoking question. Claims of this kind often seem very plausible, but they have a lot of problems surrounding them. What we would need to argue is that the relationship between Chargaff's sample and the total set of organisms on earth is one that has some useable similarity to the configuration that is assumed in a model of random sampling. I think that Chargaff was right to be cautious, even though he probably came to regret it. He was not given a share in the Nobel Prize with Watson and Crick, to his great frustration, even though Maurice Wilkins, whose role was fairly minor, was.¹¹ People in later years have wondered why Chargaff did not make more of the regularity he had found, and push its significance harder (Manchester 2008). If he had done so, urging that he had uncovered a general fact that was surely a clue to the structure of the DNA molecule, he might well have won the Nobel.

I have been discussing success stories in this section. We should look at some bad inductions as well as good ones. Here the best-known example is a suitable one: the swans. It is no surprise that the white-swan generalization (if people actually made it) turned out to be false. The sample seen by premodern Europeans was not random (as swans in distant parts of the world were very unlikely to be observed) and there was no support to such a uniformity given by known biological mechanisms (or unknown ones, for that matter). One can be unlucky with a good method, of course, but the swans case was not like that: there was no valid reason to think the generalization was true. No good bridge from observed to unobserved cases was exploited or even available.

6 Conclusion

The approach to the problem of induction taken in this essay has been to ask: how can inferences from observed cases to generalizations be reliable? "Reliability" can be understood in many ways (Goldman 1986). Rather than going into the details of this concept, I have asked several very

coarse-grained questions. What *sort of relation* might be available for an agent to exploit? What bridges between observed and unobserved cases can make reliable inference possible? In the case of a familiar philosophical sense of induction, I argue that there *is* no bridge. Induction in that sense, again, is supposed to be a kind of inference in which weight of numbers matters even without randomness, and the predicates involved are constrained, but only to rule out pseudo-patterns in which the weight of numbers does not have its usual role. I agree that this kind of inference might be an intuitively attractive one, but there is no reason why it should tend to work, at least in worlds like ours. Two relatives of this form of inference can be reliable in principle, however. These are inferences about populations from random samples, and etiologies common to a natural kind.

Science and other parts of epistemic practice often combine or mix these two methods. Sometimes one or the other is used overtly, but sometimes they are used more implicitly, even inadvertently. A researcher might have a rather incoherent epistemic ideology, while it is possible for an observer of their work to say: here the method used in fact approximated an X-based one . . . here it approximated a Y-based one.

The role of approximation seems particularly important here, and poorly understood from a philosophical point of view. In the case of inference from samples, the situation seems like this. If a physical set-up really does meet the requirements for random sampling, then reliable inferences can be drawn. But the requirement that every member of the population has the same chance of making its way into the sample is very strong. Or rather, it is strong if it makes sense at all, and it is not even clear that this application of a physical concept of probability makes sense. So philosophers often try to do without randomness altogether. (An example is seen in the Williams-Stove justification of induction, especially as defended in Campbell and Franklin 2004.¹²) That, I argue, will not work. It is better to try to make sense of the idea that many actual-world set-ups involve a useable approximation to random sampling. That is surely what people designing smoking surveys rely on, and the same can also be said in some cases where an investigator does not realize why what he is doing is in fact likely to work. The task we then have is to philosophically understand these partial matches between abstract probabilistic models and physical configurations.

In cases where a real relation between observed and unobserved is being exploited without the agent realizing this, the beliefs formed have a complicated relation to ordinary patterns of epistemic assessment. In one sense the agent is "justified" in generalizing, and in another sense he is not: our habits of epistemic assessment form a complicated soup. The kind of justification discussed in this essay is not the only one that can be recognized in everyday talk, and not the only one relevant to epistemology.

For some philosophers, this essay will also not have made much contact with what they see as the fundamental problem: inductive skepticism. Why was Millikan, Chargaff, or any Massachusetts public health official entitled to assume that the universe will not change fundamentally from one moment to the next? Why were they entitled to think their offices would still be there the next morning, that sea urchins would not explode, and that teenagers would continue to have lungs to smoke with? Nothing has been said here about questions of that kind, and for some philosophers, *these* are the real problems about induction—not questions, like the ones in this essay, that we ask against a background in which many familiar things are assumed to behave normally. In reply, I accept that more dramatically skeptical questions about induction are worth asking. The skeptical challenge is often posed in a way that relies on a particular model of the universe, however-the Humean model of a "loose and separate" world, or a mosaic of "distinct existences" (Hume 1748; Lewis 1986). Within that sort of model, inductive skepticism is an acute problem; but this is just one possible model of the universe. It is useful for posing certain questions but is not representative of how we must take things to be, or even of a default assumption that must be shown to be wrong before we can think differently. And taking skeptical challenges to be discussionworthy does not prevent us from asking other questions about how observed cases can point beyond themselves to unobserved ones, questions I have tried to make progress on here.

Acknowledgments

I am grateful to Alan Hájek, Russell Payne, Ira Schall, and participants in the 2008 INPC for comments and correspondence.

Notes

1. See Reichenbach 1938. Some could say at this point that other views about induction that might have once been "standard" are being replaced by a Bayesian view. The availability of a Bayesian approach does not affect the main arguments in this essay. I see Bayesianism, in this context, as something like what Strevens (2004) calls an "inductive framework," as opposed to an "inductive logic." The attitude to the

relation between past and future experience exhibited by a Bayesian agent is determined by the agent's assignment of likelihoods—these contain the agent's opinions about what depends on what, and whether past events can make particular future ones more or less probable. It is compatible with Bayesianism that these assignments of likelihoods be inductively skeptical, or reflect counter-inductive attitudes toward experience. A Bayesian framework may be used to represent many different views about the relation between observed and unobserved, including the views in this essay.

2. For a finite population of Fs it would not make sense to do this in all cases; if you see a rate of 100-percent G in a sample, and generalize that exactly 0 percent of Fs are G, you have said something that your own observations show to be false. Other policies nearly as counter-inductive would be possible, though. Further, the conclusion reached by extrapolation from a sample is properly expressed by saying that the true value lies in an interval, whose size is determined by properties of the sample (see note 3). A counter-inductivist might do something similar.

3. If the true proportion of Gs in the population of Fs is p, and the population is sampled randomly with samples of size N, then the number of Gs in the sample will be distributed binomially with parameters N and p. This "sampling distribution" will have a mean of *pN* and a variance of Np(1-p). Let p^* be the observed proportion of Gs (the number observed divided by N). The sampling distribution of p^* will also be binomial, with a mean of p and a variance of p(1-p)/N. The observed proportion p^* is an unbiased estimate of p. An investigator will also calculate a "confidence" interval" around p^* , which will depend on p^* and the sample size. (The likely size of the difference between p and p^* depends on p itself, and this is being estimated from p^* .) A 95-percent confidence interval might be used, in which case the claim made is that if samples were taken repeatedly and the investigator was to claim each time that the true value p lay within that interval around the observed p^* , the investigator would be right 95 percent of the time. The philosophical significance of this sense of "reliability" may be questioned, especially for a real-world setting in which only one sample is taken. Here, as indicated in my introduction, we face problems concerning the relation between an idealized model and real-world cases.

For problems of estimation of the kind discussed in this essay, a 95-percent interval is often calculated as: $p^{\pm}\pm 1.96\sqrt{p^{*}(1-p^{*})/N}$. This method is based on the fact that the binomial distribution approximates a normal distribution when *N* is large. Other methods are also used, which may be more appropriate for small values of *N* and small or large p^{*} . (E.g., when p^{*} is 0 or 1, the interval given by the rule above has 0 size.) There is an exact method (the "Clopper-Pearson" interval).

4. Hawthorne is a place, not a person; the case is sketched in my 2004 paper.

5. Here I draw on Peter Geach's notion of a "Cambridge change" (1972).

6. For discussions of "induction" which bear in different ways on this second kind of inference, see Boyd 1999, Millikan 2000, and Norton 2003.

7. *Eigen's Political and Historical Quotations*, <http://www.politicalquotes.org>, makes this attribution, citing the *New York Times Magazine*, July 4, 1976. An "urban legends" web page, <http://www.snopes.com>, claims that it was attributed to him by Governor Pat Brown, and what Reagan actually said was perhaps relevantly different from the point of view of the epistemology of induction: "If you've looked at a hundred thousand acres or so of trees—you know, a tree is a tree, how many more do you need to look at?"

8. More exactly, about 18 percent had smoked in the previous thirty days. My source here is: *Health and Risk Behaviors of Massachusetts Youth, 2007: The Report,* http://www.doe.mass.edu/cnp/hprograms/yrbs/2007YRBS.pdf>.

9. See also Chargaff et al. 1950, 757. Chargaff's handling of the finding is reviewed in Manchester 2008.

10. Not quite universally: ssDNA viruses, which have a single-stranded form of DNA, are exceptions.

11. Rosalind Franklin, who many think deserved a share in the prize, had died.

12. See also Williams 1947 and Stove 1986.

References

Boyd, R. 1999. Homeostasis, species, and higher taxa. In *Species: New Interdisciplinary Essays*, ed. Robert Wilson. Cambridge, Mass.: MIT Press.

Campbell, S., and J. Franklin. 2004. Randomness and induction. *Synthese* 138: 79–99.

Carnap, R. 1952. *The Continuum of Inductive Methods*. Chicago: University of Chicago Press.

Chargaff, E. 1951. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* 6:201–240.

Chargaff, E., S. Zamenhof, and C. Greene. 1950. Composition of human desoxypentose nucleic acid. *Nature* 165:756–757.

Dewey, J. 1939. Experience, knowledge and value: A rejoinder. In *The Philosophy of John Dewey*, ed. P. A. Schilpp and L. E. Hahn. Library of Living Philosophers. La Salle: Open Court.

Franklin, A. 1997. Millikan's oil-drop experiments. Chemical Educator 2:1-14.

Geach, P. 1972. Logic Matters. Berkeley: University of California Press.

Gigerenzer, G., P. Todd, and ABC Research Group. 1999. *Simple Heuristics That Make Us Smart*. Oxford: Oxford University Press.

Godfrey-Smith, P. 2004. Goodman's problem and scientific methodology. *Journal of Philosophy* 100:573–590.

Goldman, A. 1986. *Epistemology and Cognition*. Cambridge, Mass.: Harvard University Press.

Goodman, N. 1955. *Fact, Fiction, and Forecast*. Cambridge, Mass.: Harvard University Press.

Harman, G. 1965. The inference to the best explanation. *Philosophical Review* 74:88–95.

Hume, D. [1748] 1993. *An Enquiry Concerning Human Understanding*. Indianapolis: Hackett.

Jackson, F. 1975. Grue. Journal of Philosophy 72:113-131.

Jackson, F., and R. Pargetter. 1980. Confirmation and the nomological. *Canadian Journal of Philosophy* 10:415–428.

Kelly, K. 2004. Why probability does not capture the logic of scientific justification. In *Contemporary Debates in the Philosophy of Science*, ed. C. Hitchcock. London: Blackwell.

Lewis, D. 1983. New work for a theory of universals. *Australasian Journal of Philosophy* 61:343–377.

Lewis, D. 1986. Preface. In *Philosophical Papers*, vol. 2. Oxford: Oxford University Press.

Manchester, K. 2008. Erwin Chargaff and his "rules" for the base composition of DNA: Why did he fail to see the possibility of complementarity? *Trends in Biochemical Sciences* 33:65–69.

Medin, D., and S. Atran, eds. 1999. Folkbiology. Cambridge, Mass.: MIT Press.

Millikan, R. A. 1911. The isolation of an ion, a precision measurement of its charge, and the correction of Stokes's Law. *Physical Review* 32:349–398.

Millikan, R. A. 1917. The Electron. Chicago: University of Chicago Press.

Millikan, R. G. 2000. On Clear and Confused Ideas. Cambridge: Cambridge University Press.

Norton, J. 2003. A material theory of induction. Philosophy of Science 70:647-670.

Quine, W. V. 1969. Natural kinds. In his *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Rawls, J. 1971. A Theory of Justice. Cambridge, Mass.: Harvard University Press.

Reichenbach, H. 1938. *Experience and Prediction: An Analysis of the Foundations and the Structure of Knowledge*. Chicago: University of Chicago Press.

Reichenbach, H. 1939. Dewey's theory of science. In *The Philosophy of John Dewey*, ed. P. A. Schilpp and L. E. Hahn. Library of Living Philosophers. La Salle: Open Court.

Stalker, D., ed. 1994. Grue: The New Riddle of Induction. La Salle: Open Court.

Stove, D. C. 1986. The Rationality of Induction. Oxford: Oxford University Press.

Strevens, M. 2004. Bayesian confirmation theory: Inductive logic, or mere inductive framework? *Synthese* 141:365–379.

Williams, D. C. 1947. *The Ground of Induction*. Cambridge, Mass.: Harvard University Press.

3 It Takes More Than All Kinds to Make a World

Marc Lange

1 Introduction

Figure 3.1 shows one picture of what it takes to make a world: the kinds of particles, the kinds of forces, and a vast mixing bowl. I shall be arguing that although, as the saying goes, "it takes all kinds to make a world," there is an important sense in which it takes *more* than all of the *actual* kinds. What the particle species would have been like, had the list of species been different in various ways, is crucial to making the actual classes the natural kinds they are.

The various species of elementary particle (some of which are depicted below, in table 3.1) are in many ways *ideal* cases of natural kinds. Each elementary particle belongs to exactly one of these natural kinds and it essentially belongs to that kind. In terms of certain properties, there are perfect uniformities within each species and sharp distinctions between the species: there are no intermediate cases, much less a continuum of intergradations. Among these characteristic properties, some are fundamental and suffice to give necessary and sufficient conditions for species-membership, while others derive from the fundamental properties via exceptionless uniformities with no *ceteris paribus* escape clauses. The particles belonging to these species are elementary, so there are no worries that these classes are actually heterogeneous at a more basic level. All in all, the elementary-particle species are ideally behaved natural kinds.

(Of course, perhaps charge, charm, and the other properties I have identified in the table as fundamental are not actually fundamental; perhaps the particles I have listed are not even elementary. I will assume they are, and I will assume various other physical details along the way—but merely for the sake of having definite examples. Nothing I have to say will turn on these assumptions.)

Some eler	nentary	particles o	f matter.								
	Funda	umental pro	operties						Some derivative	properties	
quark	spin	charge*	baryon number	s	С	В	Н	mass (MeV/c ²)	magnetic moment (µN)	half-life	decay products (with W)
Up	1/2	2/3	1/3	0	0	0	0	360	+1.852	$>4.6 imes 10^{12} \mathrm{yr}$	q
Down	1/2	-1/3	1/3	0	0	0	0	360	-0.972	>900 s	n
Charm	1/2	2/3	1/3	0	1	0	0	1500	ż	$1.1 imes 10^{-12}~{ m s}$	s(95%), d(5%)
Strange	1/2	-1/3	1/3	1	0	0	0	540	-0.613	$1.24 imes 10^{-8} ext{ s}$	n
Top	1/2	2/3	1/3	0	0	0	1	174000	ż	10 ⁻²⁵ s	b(>99%), s, d
Bottom	1/2	-1/3	1/3	0	0	7	0	5000	ż	$1.3 imes 10^{-12}~{ m s}$	c(>99%), d, u
lepton	spin	charge*	lepton number					mass (MeV/c ²)	magnetic moment (10 ⁻²⁷ J/T)	half-life	decay products
Electron	1/2	-1	1					0.510998918	-9284.764	stable	
Muon	1/2		1					105.658369	-44.90447	$2.2 imes 10^{-6} m s$	e-& ν _e & ν _μ
Taon	1/2	-1	-					1776.99	-2.6801	$2.9 imes 10^{-13}$ s	$\begin{array}{l} e^{-\xi_{r}} \upsilon_{t} \xi_{r} \upsilon_{e} \\ (17.84\%), \\ \mu^{-\xi_{r}} \upsilon_{t} \xi_{r} \upsilon_{\mu} \\ (17.36\%) \ldots \end{array}$

*1 unit = 1.602 176 487(40) × 10⁻¹⁹ C.

Table 3.1



Figure 3.1

The Standard Model of elementary particles and fundamental forces, <http:// www.particleadventure.org/frameless/standard_model.html> (downloaded January 10, 2010). Courtesy of the Particle Data Group of Lawrence Berkeley National Laboratory.

Beyond the sorts of facts depicted in table 3.1, at least three further ingredients are needed to make these classes into natural kinds.

First, the fundamental properties shared within a class and collectively differentiating the classes must be *natural* properties: not gerrymandered, "wildly disjunctive" (Fodor 1974, 103), "grue"-some (Goodman 1983) shadows of predicates. Members of the same class must be genuinely alike and members of different classes genuinely dissimilar in various respects.¹

Second, the combination of fundamental properties characteristic of the muon, for example, must not merely be *universally associated* with each of the muon's derivative properties. Rather, they must be connected *by natural law*. It must be no accident that any two bodies with the muon's charge, mass, and so forth also possess the muon's half-life and magnetic moment.

Third, going beyond the muon or any other particular line of the table, it must be no accident that elementary particles alike in every one of the *fundamental* respects are alike in the *derivative* respects. Of course, that result can be achieved in a cheap way: namely, for the laws

i. to specify the combination of fundamental properties characteristic of the electron, the combination characteristic of the muon, and so forth;

ii. to say that anything with the electron's combination of fundamental properties also has the electron's derivative properties, and likewise for each of the other particle species; and

iii. to require that every elementary particle be an electron or a muon or . . . (insert here the rest of the list of particle species).

That elementary particles which are alike in their spin, charge, and so on are also alike in their half-lives and magnetic moments would then follow

exclusively from nonaccidents and so be no accident itself. But if that is the whole story, then it is just a kind of fluke that elementary particles alike in these various fundamental properties are alike in those derivative properties; this arrangement is just a kind of coincidence, resulting merely from the particular elementary-particle species there happen to be—for example, that no two species have the same spin, charge, and so forth, while having different half-lives. The laws linking the fundamental and derivative properties are supposed to *transcend* the particular particle species there are. The associations between fundamental and derivative properties are supposed to be laws *in their own right*, and not merely consequences of the fact that all bodies with the electron's fundamental properties (happen to) have the electron's derivative properties, and likewise for the muon, and so forth.

Of course, the notions I have just put in play (i.e., natural kind, natural property, fundamental vs. derivative property, law of nature) belong to a much wider network of concepts, such as scientific explanation and inductive projection. Natural properties of concrete things have been thought to figure in natural laws and thereby to explain other properties and behaviors. Our opinions about which properties are natural influence the scope of the inductive projections by which we arrive at our beliefs about natural laws. Philosophers such as David Lewis, Sydney Shoemaker, Jerry Fodor, and Nelson Goodman have taken different members of this network of concepts as being ontologically prior to the others. Let's take a whirlwind tour of some of these options. Lewis (1984; 1986a, 63-69; 1999, 232) takes naturalness to be metaphysically basic, uses it to fund his Best System Account of laws, and then uses lawhood to understand scientific explanations and inductive projections. (The rival account of laws offered by David Armstrong, in terms of nomic-necessitation relations among universals, also presupposes that certain properties already qualify as natural—by virtue of answering to universals.) An approach like Shoemaker's (1984), in contrast, takes properties as being individuated by the contributions their possession can make to the causal powers and susceptibilities of the things which possess them; natural properties are distinguished from unnatural ones by their taking causal-explanatory responsibility, which is a primitive matter. On this view, laws (which express these causal roles) are metaphysically necessary, and that every instance of a given property has the same causal profile supports inductive projections. Alternatively, Fodor (1974, 102) takes a natural property as one that marks out the range of cases covered by the same natural law. Goodman (1983) takes projectibility as responsible for lawhood, and since

a predicate's projectibility arises from our actual *history* of projections, Goodman arrives at a comparatively lightweight distinction between laws and accidents.

In trying to account for the three further ingredients that (I suggested) make the elementary-particle classes into natural kinds, my strategy will be to start by trying to understand natural law—or, more precisely, natural necessity (which sets the laws and their consequences apart from accidents). I shall argue that natural necessity is grounded in various subjunctive facts. I shall then try to use similar subjunctive facts to understand what makes certain properties natural as well as what makes the associations between fundamental and derivative properties constitute laws "in their own right," transcending the particular classes of particles there happen to be.

2 Distinguishing the Laws by Their Stability²

Laws of nature differ from accidents in having greater perseverance under counterfactual suppositions. For instance, suppose that natural law prohibits any body from being accelerated from rest to beyond the speed of light; then this cosmic speed limit could not have been broken even if the Stanford Linear Accelerator had now been cranked up to full power. On the other hand, suppose it is just an accident that all gold cubes are smaller than a cubic mile (to borrow Reichenbach's famous example); then if Bill Gates had desired a gold cube greater than a cubic mile, I dare say there would have been one.

A natural suggestion along these lines is that the *range* of counterfactual suppositions under which a given accident would still have held is *narrower* than a law's range of invariance under counterfactual suppositions. But this is not the case. Suppose we have laid out on a table a large number of electric wires, all of which are made of copper. Had copper been electrically insulating, then the wires on the table would not have been much good for conducting electricity. Now look at what just happened: under the counterfactual supposition that copper *is* an insulator, the law that copper is electrically conductive obviously failed to be preserved; but the accident that all the wires on the table are made of copper *was* preserved. Here is an instant replay. Had copper been electrically insulating, then the wires of copper—would have been useless for conducting electricity. Thus, it is not true that any accident's range of invariance under counterfactual suppositions is strictly narrower than any law's.

Here is a different suggestion. Start with the facts that we are trying to partition into laws and accidents. Put aside all those that could themselves hold at least partly in virtue of which facts are laws and which are not. For example, lay aside the fact that it is a law that all emeralds are green, but retain the bare fact that all emeralds are green. Call what remain the "subnomic facts" and let the "sub-nomic claims" be the claims that in any possible world are true or false exclusively in virtue of the sub-nomic facts there. Stipulate that other species of modality be treated in the same way as natural modality. For example, the sub-nomic facts include the fact that all triangles have three sides, but do not include that this fact has some broadly logical necessity.

Now let's work our way toward a means of picking out the sub-nomic facts that are laws by their special resilience under counterfactual suppositions. Roughly speaking, the laws would still have held under any supposition that is logically consistent with *the laws*. This restriction rules out the supposition "Had copper been electrically insulating . . ." from my example involving the wires on the table. But we still have a problem: had energy conservation not been required by law, then energy might not have been conserved. The supposition that energy conservation is not a law *is* logically consistent with all of the sub-nomic facts that are laws. For instance, it is logically consistent with the fact that the total quantity of energy is conserved. But it is not the case that energy would have been conserved had there been no legal obligation for it to be. So the fact that energy is conserved is not invariant under every counterfactual supposition that is logically consistent with all of the sub-nomic facts that are laws.

One way to solve this problem is to note that the supposition that energy conservation is not *a law* is not a *sub-nomic* supposition; it concerns the laws, and not exclusively the sub-nomic facts. Accordingly, I suggest that the sub-nomic facts that are laws are preserved under every *sub-nomic* supposition that is logically consistent with all of the sub-nomic facts that are laws. Trivially, no accident is preserved under all of these suppositions (since one of these suppositions is the accident's negation). More fully:

It is a law that *m* (where *m* is sub-nomic) if and only if in all conversational contexts, "had *p* been the case, then *m* would still have the case" ($p \square \rightarrow m$) is true for any sub-nomic claim *p* that is logically consistent with all of the sub-nomic claims *n* (taken together) where it is a law that *n*.

The only way that this suggestion stands any chance of being correct is if the logical truths, conceptual truths, mathematical truths, metaphysical truths, and so forth—the "broadly logical truths"—are counted by courtesy among the natural laws, since they have at least the same perseverance under counterfactual suppositions as mere laws do. The suggestion includes a reference to "all conversational contexts" because the truth-values of counterfactual conditionals are notoriously context-sensitive. Of course, the suggestion is that there are *some* limits to the influence that context can have. You might consider whether there are any contexts in which the laws fail to be preserved under a sub-nomic counterfactual supposition that is logically consistent with them.³

But for now, notice that even if the above suggestion is correct, it cannot reveal that the laws bear an especially intimate relation to counterfactuals. That is because in the suggestion, the laws appear on both sides of the "if and only if": the *range* of counterfactual suppositions covered by the suggestion has been expressly tailored to suit the laws. The laws are picked out by their invariance under a range of suppositions which, in turn, is selected by the laws. Unless there is some prior, independent reason why this particular range of suppositions is special, the laws' invariance under this particular range of suppositions fails to reveal anything special about the laws.

Some philosophers have deemed this circularity unavoidable,⁴ but there is a way to avoid it. Roughly speaking, the suggestion we have been looking at says that the laws form a set of truths that would still have held under each counterfactual supposition that is logically consistent with that set. In contrast, take the set containing exactly the logical consequences of the fact that all gold cubes are smaller than one cubic mile. Consider the suppositions that are logically consistent with that set. The set's members are not all preserved under every one of those suppositions. For example, they are not all preserved under the supposition that Bill Gates wants a gold cube larger than a cubic mile.

That is the idea behind what I call "sub-nomic stability":

Consider a non-empty set Γ of sub-nomic truths containing every subnomic logical consequence of its members. Γ possesses *sub-nomic stability* if and only if for each member *m* of Γ (and in every conversational context),

 $\begin{array}{l} \sim (p \Leftrightarrow \rightarrow \sim m), \text{ [entails } p \square \rightarrow m] \\ \sim (q \Leftrightarrow (p \Leftrightarrow \rightarrow \sim m)), \text{ [entails } q \square \rightarrow (p \square \rightarrow m)] \\ \sim (r \Leftrightarrow \rightarrow (q \Leftrightarrow (p \Leftrightarrow \rightarrow \sim m)), \ldots \text{ [entails } r \square \rightarrow (q \square \rightarrow (p \square \rightarrow m))] \\ \text{for any sub-nomic claims } p, q, r, \ldots \text{ where } \Gamma \cup \{p\} \text{ is logically consistent, } \\ \Gamma \cup \{q\} \text{ is logically consistent, } \Gamma \cup \{r\} \text{ is logically consistent } \ldots \end{array}$

The conditionals say that under every sub-nomic supposition that is logically consistent with the set, all of the set's members *m* would still have held—indeed, not one of their negations might have held. ("~ $(p \Leftrightarrow ~ m)$ " means that it is not the case that ~*m* might have held, had *p* held.)

There are also some nested counterfactuals in the definition, requiring that the set's members be preserved under the same range of suppositions, however nested. For instance, had Jones missed his bus this morning, the natural laws would still have held; so had Jones simply clicked his heels and made a wish to arrive instantly at his office, he wouldn't have gotten to work. Later, I will say something about the significance of including these nested counterfactuals in the definition of "sub-nomic stability," but for now, let's set them aside.

The key point is that sub-nomic stability avoids privileging the range of counterfactual suppositions that are logically consistent with the laws. Sub-nomic stability gives each set the opportunity to pick out for itself a convenient range of suppositions. The definition of sub-nomic stability treats the set of laws no differently from any other set: the definition is egalitarian, granting no special privileges to the laws.

Let's return to the sub-nomic stability of the set spanned by the subnomic truths that are laws (call that set " Λ ") in contrast to the set spanned by the gold-cubes accident. Let's look at another example. Take the accident g that whenever a certain car is on a dry, flat road, its acceleration is given by a certain function of how far its gas pedal is being depressed. Had the gas pedal on a certain occasion been depressed a bit further, then g would still have held; had I been wearing an orange shirt, then g would still have held. How can we build a sub-nomically stable set that contains g? Along with g, the set must also include the fact that the car has a fourcylinder engine, since had the engine used six cylinders, g might not still have held. (Once the set includes the fact that the car has a four-cylinder engine, the supposition that the engine has six cylinders is logically inconsistent with the set, so the set does not have to be preserved under that supposition in order to be stable.) But now that the set includes a description of the car's engine, its stability requires that it also include a description of the engine factory, since had that factory been different, the engine might have been different. And had the price of steel been different, the engine might have been different. And so on: this ripple effect will propagate endlessly.

For instance, let's say that all of the apples on my tree are ripe—an accidental truth. Now take the supposition: had either g been false or there been unripe apples on my tree. Is g preserved under that supposition? In

every context? Certainly not. Therefore, to be stable, a set that includes *g* must also include the fact that all of the apples on my tree are ripe, making the set logically inconsistent with the supposition that either *g* is false or there are unripe apples on my tree (so the set does not have to be preserved under that supposition in order to be stable). But if a stable set that includes *g* must also include even the fact about my apple tree, then I conclude that the only set containing *g* that might be sub-nomically stable is the set of *all* sub-nomic truths (the "maximal" set).

Accordingly, I suggest that the laws are distinguished as follows:

No nonmaximal set of sub-nomic truths containing an accident possesses sub-nomic stability, but sub-nomic stability is possessed by Λ (the set containing exactly the sub-nomic truths *m* where it is a law that *m*).

This proposal for distinguishing laws from accidents avoids the circularity afflicting the suggestion that the laws are the truths that would still have held under every supposition that is logically consistent with the laws.

3 Natural Necessity

Besides Λ , are there any other nonmaximal sets that are sub-nomically stable? The sub-nomic, broadly logical truths form a sub-nomically stable set since they would still have held under any broadly logical possibility. Before going further, I will show that for any two sub-nomically stable sets, one must be a proper subset of the other.

Proof that if Γ and Σ are distinct stable sets, then one must be a proper subset of the other:

1. Suppose (for *reductio*) that Γ and Σ are sub-nomically stable, *t* is a member of Γ but not of Σ , and *s* is a member of Σ but not of Γ .

2. Then (~*s* or ~*t*) is logically consistent with Γ .

3. Since Γ is sub-nomically stable, every member of Γ would still have been true, had (~*s* or ~*t*) been the case.

4. In particular, *t* would still have been true, had (~*s* or ~*t*) been the case. That is, (~*s* or ~*t*) $\Box \rightarrow t$.

5. So *t* & (~*s* or ~*t*) would have held, had (~*s* or ~*t*). Hence, (~*s* or ~*t*) $\Box \rightarrow$ ~*s*.

6. Since (~*s* or ~*t*) is logically consistent with Σ , and Σ is sub-nomically stable, no member of Σ would have been false had (~*s* or ~*t*) been the case.
7. In particular, it is not the case that *s* would have been false, had (~*s* or ~*t*) been the case. That is, \sim ((~*s* or ~*t*) $\Box \rightarrow \sim s$).

8. Contradiction from 5 and 7.

Thus, we have our reductio: for any two stable sets, one must be a proper subset of the other.

We were asking whether there are any nonmaximal stable sets besides A. Since any superset of Λ contains accidents, none is stable except for the maximal set. So to find nonmaximal stable sets besides Λ , we must look among Λ 's proper subsets. Many of them are clearly unstable. For example, take Coulomb's law giving the electric repulsive force (F) between point charges q and Q, separated by r, that have been at rest for a timespan of r/c, as equal to qQ/r^2 . Suppose we restrict Coulomb's law to times after today, and we take the set containing exactly the broadly logical consequences of this restricted Coulomb's law. This set of truths is unstable since it is not the case that the restricted Coulomb's law would still have held, had Coulomb's law been violated sometime before today. Under that supposition, Coulomb's law would have been out of the way, and so there would have been nothing to keep Coulomb's law from being violated after today-just as energy would not still have been conserved, had there been no law to make its conservation mandatory. Therefore, the proper subset of Λ consisting of the restricted Coulomb's law and its broadly logical consequences is unstable, since it would not still have held, had Coulomb's law been violated sometime before today.

However, some of Λ 's proper subsets are plausibly stable. Take the fundamental law of dynamics, which (roughly speaking) governs the relation between the forces acting on a body and the body's motion. For simplicity, let's suppose it is Newton's second law of motion (F = ma) relating the net force (F) on a body to its mass (m) and acceleration (a). How would bodies have behaved, had they been subjected to various weird, hypothetical kinds of forces? Naturally, we use Newton's second law of motion to find out. For example, in 1917, Paul Ehrenfest famously showed that had gravity been an inverse-cube force or fallen off with distance at any greater rate, then the planets would eventually have collided with the sun or escaped form the sun's gravity. Ehrenfest's argument presumes that Newton's second law of motion would still have held, had gravity obeyed a different force law. The fundamental dynamical law apparently belongs to a sub-nomically stable set that does not include the force laws, so it would still have held, had the world been populated by different kinds of forces.

It is also frequently suggested that the conservation laws are not mere reflections of the particular kinds of forces there happen to be. As Eugene Wigner remarked:

[F]or those [conservation laws] which derive from the geometrical principles of invariance [that is, conservation of energy, momentum, etc.] it is clear that their validity transcends that of any special theory—gravitational, electromagnetic, etc.— which are only loosely connected. (1972, 13)

I shall return to Wigner's comment, but for the moment, let's take away from it only the lesson that the conservation laws would still have held, had there been different kinds of forces involved or had the actual kinds of forces been different. Richard Feynman has likewise remarked on how the conservation laws are lifted clear of the ruck of the various force laws:

When learning about the laws of physics you find that there are a large number of complicated and detailed laws, laws of gravitation, of electricity and magnetism, nuclear interactions, and so on, but across the variety of these detailed laws there sweep great general principles which all the laws seem to follow. Examples of these are the principles of conservation, certain qualities of symmetry, the general form of quantum mechanical principles. (1967, 59)

Physicists apparently also believe that the law of the composition of forces (the "parallelogram of forces") would still have held, had the force laws been different.

Likewise, consider the Lorentz transformations, which are central to Einstein's special theory of relativity, entailing such famous relativistic results as time-dilation and length-contraction. They specify how an event's spacetime coordinates in one reference frame in a certain family relate to its coordinates in the other frames in that family. In his first relativity paper in 1905, Einstein derived these transformations by using the principle that light travels at the same rate in one of these frames whatever the motion of its source. However, as Einstein later wrote (1935, 223), this derivation is misleading because features of light are not responsible for the coordinate transformations. The widespread recognition that the transformations lie deeper than the particular kinds of fundamental forces there happen to be has led to a tradition, beginning as early as 1909, of deriving the Lorentz transformations without appealing to details of electromagnetism or any other force.⁵ It is only because scientists believe that the Lorentz transformations transcend the fundamental force laws that they believe that were there additional fundamental force laws or had the fundamental force laws been different, the Lorentz transformations would still have held. As Roger Penrose (1987, 21) says, if it is just "a 'fluke'" that certain dynamical laws exhibit Lorentz covariance, then "[t]here is no need to believe that this fluke should continue to hold when additional ingredients of physics are" discovered. But physicists generally do regard special relativity as prior to the force laws. For instance, physicists commonly assert (e.g., Lévy-Leblond 1976, 271) that had the force laws been different, so that photons, gravitons, and other kinds of particles that actually possess zero mass instead possessed non-zero mass, the Lorentz transformations would still have held (though these particles would not have moved with the speed c that famously figures in those transformations).

The idea seems to be that the fundamental dynamical and conservation laws, along with the spacetime coordinate transformations, the parallelogram of forces, and some other elite laws belong to a sub-nomically stable set that omits the various grubby special laws that specify the particular kinds of forces there are. In other words, we seem to have a hierarchy of sub-nomically stable sets, as shown in figure 3.2.67

Now what should we make of a proper subset of Λ that is sub-nomically stable? Let's begin by noticing that for a set to be stable is for that set to have maximum staying-power under counterfactual suppositions-that is to say, for its members all to be preserved under every sub-nomic supposition under which they *could* without contradiction all be preserved. The members of a sub-nomically stable set are *collectively* as resilient under



Figure 3.2

Some good candidates for sub-nomically stable sets.

sub-nomic suppositions as they could *collectively* be. Accordingly, I suggest that a sub-nomic truth is *necessary* exactly when it belongs to a nonmaximal sub-nomically stable set and that for each of these sets, there is a distinct species of necessity possessed by exactly the set's members. This fits nicely, I think, with our pretheoretic conception of what it takes to count as necessary, a conception nicely captured by J. S. Mill in *A System of Logic* (III, 5, vi):

If there be any meaning which confessedly belongs to the term necessity, it is *unconditionalness*. That which is necessary, that which *must* be, means that which will be whatever supposition we make with regard to other things.

The idea here, as I shall appropriate it, is that whether a certain body of truths possesses some characteristic variety of necessity depends on whether those truths would all still have held "whatever supposition we make with regard to *other things*"—that is, under *any* supposition that does not contradict those truths. Not every fact that withstands even a wide range of counterfactual disturbances qualifies as possessing some species of necessity. No variety of necessity is possessed by an accident, since no nonmaximal set containing an accident possesses sub-nomic stability.

This conception of necessity fits nicely with the principle that whatever would have happened, had something possible happened, must also qualify as possible. Suppose that q's truth is possible—that is to say, q is logically consistent with some nonmaximal stable set. Then under the counterfactual supposition that q holds, every member of that set would still have held, since the set is stable. Thus, anything else that would also have held must be logically consistent with the set, and so must possess the same variety of possibility as q. This means that whatever would have happened, had something possible happened, must also be possible.

We were considering what we should make of a stable proper subset of Λ , such as a set containing the fundamental dynamical and conservation laws (and so forth) but none of the force laws. The suggestion at which I have now arrived is that Λ is associated with one variety of necessity, and a stable proper subset of Λ is associated with a stronger variety of necessity. The fundamental dynamical law, then, possesses a stronger variety of natural necessity than the force laws do. This picture of necessity as constituted by sub-nomic stability identifies what is common to broadly logical necessity and to the various grades of natural necessity, in virtue of which they are all species of the same genus. The identification of necessity with sub-nomic stability explains how the laws could possess a variety of *necessity* and nevertheless be *contingent*.

4 The Laws Form a System

On my view, p's necessity involves its belonging to a nonmaximal subnomically stable set. Hence, its necessity (of whatever species) is a collective affair—the accomplishment of an entire team of facts. The range of counterfactual suppositions under which p is preserved, in connection with its necessity, depends not only on p but also on the other facts joining p in the sub-nomically stable set.

Let's consider natural necessity in particular. Because *p*'s lawhood depends on *p*'s membership in a nonmaximal sub-nomically stable set, each natural law is bound up with the others. Each member of the set depends on the others to help specify the range of invariance that it has got to possess in order for it to be a law. That is, each member of the set participates in delimiting the range of counterfactual suppositions under which every other member must be preserved in order for the set to be stable. The laws derive their lawhood collectively; their sub-nomic stability means that they are *together* as resilient under sub-nomic counterfactual suppositions as they could *together* be. They form a unified, integrated whole—a system. (In depicting lawhood as a collective affair rather than an individual achievement, my account agrees with Lewis's and differs from Armstrong's and from Shoemaker-style scientific essentialism.)

That the various laws must "cover" for one another in this way has sometimes helped to lead scientists to discover a heretofore unknown law. Let's look at an example from classical physics. Suppose we have two charged point-bodies—one having been at rest for a long time, the other having been in motion for a long time in a constant direction at a constant speed v. When the moving body streaks past the stationary one, then at its moment of closest approach (at a distance r), their mutual electric repulsion equals the product of their charges divided by $r^2 \sqrt{[1 - (v^2/c^2)]}$. (The force is thus greater than the force-given by Coulomb's law-between point-bodies long at rest at a separation r.) Now although this $\sqrt{1}$ (v^2/c^2)] factor crops up everywhere in Albert Einstein's special theory of relativity, Oliver Heaviside actually discovered this law in 1888 (seventeen years before Einstein published relativity theory) by deriving it from James Clerk Maxwell's electromagnetic-field equations of 1873. (Maxwell's electromagnetic theory was already relativistic before "relativity" came along: it is a bit of twentieth-century physics that was discovered in the nineteenth century.) Several physicists seem to have regarded Heaviside's discovery as suggesting that the speed of light is a cosmic speed limit for charged bodies, since if v exceeds c, then $\sqrt{1 - (v^2/c^2)}$ is the square root

of a negative quantity, and so Heaviside's equation yields an imaginary number as the magnitude of the force. For example, G. F. C. Searle concluded that "it would seem to be impossible to make a charged body move at a greater speed than that of light."⁸ Likewise G. F. FitzGerald wrote to Heaviside: "You ask 'what if the velocity be greater than that of light?' I have often asked myself that but got no satisfactory answer. The most obvious thing to ask in reply is 'Is it possible?'"⁹

How should we reconstruct their argument for this conclusion? Here is my suggestion. It seems doubtful that all charged bodies would still have accorded with Heaviside's equation had there been a charged body moving at superluminal speed. But a law must be preserved under every p that is logically consistent with Λ . Hence, unless Heaviside's equation is not a law (or applies only to subluminal speeds¹⁰), there must be a law prohibiting charged bodies from moving superluminally. Thanks to that law, the failure of Heaviside's equation to be preserved under the supposition of superluminal charged bodies fails to impugn that equation's lawhood.¹¹ Heaviside's law depends on the law prohibiting superluminal speeds to limit the range of invariance that it has to have in order to qualify as a law.

That the laws have to form a *system* helps to account for certain other features of lawhood, such as the relation between deterministic laws and chancy facts. Suppose there is a rare radioactive isotope Is. Let C be the chancy fact that each Is atom has a half-life of 7 microseconds. Thus, it could happen that regularity R holds: each Is atom decays within 7 microseconds of being created. But given C's nonvacuous truth, R cannot be a law; R, if true, must be an accident. Why is that? After all, R's nonvacuous truth permits C to hold as a law. Why does C's nonvacuous truth prohibit R from holding as a law?

Here is an explanation—supplied by the way in which the laws must form a system. Suppose R is a law; C's nonvacuous truth is logically consistent with R. Suppose for the sake of *reductio* that C's nonvacuous truth is logically consistent not just with R, but with all of the laws taken together. Then because the laws form a stable set, the laws would have to be preserved under the counterfactual supposition that C is nonvacuously true. But had C been nonvacuously true, then R might *not* still have held; each of the Is atoms *might* have decayed within 7 microseconds, but then again, some might just as well not have done so. Thus we have our *reductio*: if R is a law, there must be some other law that prohibits C's nonvacuous truth. Therefore, if R is a law, C cannot be nonvacuously true. Hence, if C is nonvacuously true, then R cannot be a law. This explanation captures the intuition that R's lawhood would mean that it is impossible for an Is atom to last more than 7 microseconds, contrary to C.

Lewis, of course, also portrays the laws as forming an integrated system. But the above relation between deterministic laws and chancy facts does not fall easily out of his account. Lewis must impose a requirement for a system's being eligible to compete for Best System; Lewis (1986b, 128, cf. 125; 1999, 233–234) calls this a requirement of "coherence." The coherence-requirement says that each candidate for Best System must imply that the chances are such as to give that very system no chance at any time of being false. Hence, if the system contains R, then it must entail that no Is atom ever had a chance of lasting longer than 7 microseconds, and so the system must entail not-C. This "requirement of coherence" gets the right answer, but it seems like an ad hoc device. It would be far better for the relation between deterministic laws and chancy facts to fall out nicely from the rest of the account, rather than to have been inserted by hand.

Let me turn briefly to a feature of sub-nomic stability that I set aside earlier: the nested counterfactuals that stability requires. These are needed to capture the laws' characteristic invariance. For instance, had we tried to accelerate a body from rest to beyond the speed of light, we would have failed—and moreover, even if we had access to twenty-third-century technology, had we tried to accelerate a body from rest to beyond the speed of light, we would have failed. (That was a nested counterfactual.)

By way of these nested counterfactuals, the relation between lawhood and stability has a further consequence: it explains why the laws would still have been *laws* under any sub-nomic counterfactual supposition *p* that is logically consistent with the laws. Let's see how. Suppose that m is a member of G, a sub-nomically stable set, and that each of q, r, s . . . is individually logically consistent with G. Then G's stability ensures that ~ $(q \Leftrightarrow am), amplitude (r \Leftrightarrow am)), amplitude (r \Leftrightarrow am)), and so$ forth. One of the connections between might-conditionals $(\diamond \rightarrow)$ and would-conditionals $(\Box \rightarrow)$ is that ~ $(q \diamond \rightarrow \neg m)$ logically entails $(q \Box \rightarrow m)$. So the counterfactuals in the above sequence respectively entail $(q \square \to m), (q \square \to \sim (r \diamond \to \sim m)), (q \square \to \sim (r \diamond \to (s \diamond \to \sim m))), and so forth.$ Therefore, had *q* been the case (that is, in the "closest *q*-world"), the following all hold: $m_r \sim (r \Leftrightarrow am)_r \sim (r \Leftrightarrow am)_r \sim (r \Leftrightarrow am)_r \ldots$ simply each of the earlier conditionals with its opening ' $q \square \rightarrow \ldots$ ' lopped off (since we are talking about what's true in the closest q-world). This sequence supplies exactly what is needed for G to be sub-nomically stable in that world: *m* and its colleagues in G are true and preserved under every counterfactual supposition that is logically consistent with G.

Hence, if G is actually sub-nomically stable, then G would still have been sub-nomically stable had q been the case, for any q that is logically consistent with G. Therefore, if in any possible world the natural laws *there* are exactly the members of at least one nonmaximal sub-nomically stable set *there*, then the actual laws would still have been laws had q been the case, for any q that is logically consistent with the actual laws. (Moreover, any stratum of laws would still have constituted a stratum of laws, had any such q been the case.) We thereby save the intuition that had Uranus's axis not been so nearly aligned with its orbital plane, then although conditions on Uranus would have been different, the actual laws of nature would still have been laws—which is *why* conditions on Uranus would have been so different. The laws' collective invariance is itself no "accident"; the laws' invariance (and hence their lawhood) is invariant.

5 How Some Laws Can Transcend Others

We are now, at last, in a position to return to the list of elementary-particle species in table 3.1. I have suggested what makes the associations in this table constitute laws rather than accidents. But as we saw in section 1, more than these laws are needed for the classes of elementary particle to be natural kinds. Not only must it be no accident that all elementary particles alike in various fundamental properties are alike in various derivative properties, but this fact must transcend the particular kinds of particles there are. The associations between fundamental and derivative properties must be "laws in their own right"—which we can now understand in terms of the *strata* of stable sets.

The associations between fundamental and derivative properties transcend the particular kinds of elementary particles there happen to be, in that had there been elementary particles *not* belonging to the actual kinds of quarks and leptons, then the same laws connecting the fundamental properties with the derivative properties would still have held; the force laws would have been no different. For example, the laws by which a particle's lifetime and magnetic moment are fixed by its mass, charge, spin, and so forth would still have held even if there had been different species of elementary particle. That is why had there been an elementary particle with the up quark's mass and spin but twice its charge, it would have possessed twice the up quark's magnetic moment.

Such subjunctive facts make stable a proper subset of Λ containing the laws connecting the fundamental properties with the derivative properties. Accordingly, these laws possess a stronger variety of natural necessity than

Broadly logical truths that are sub-nomic truths

The above and the fundamental dynamical law, the law of the composition of forces, the spacetime transformations, and the conservation laws, but not the force laws or the laws giving the physical characteristics of various kinds of particles

The above and the force laws (together with any other laws responsible for the associations between the fundamental and derivative properties of the particles)

All of the sub-nomic truths m where it is a law that $m(\Lambda)$



All sub-nomic truths?

Figure 3.3

The laws linking the fundamental and derivative properties as transcending the particular kinds of elementary particles there are.

the laws that specify the combinations of fundamental properties characteristic of the particular elementary-particle species or the "closure law" that all matter consists fundamentally of just those kinds of quarks, leptons, and so forth.

In the pyramidal hierarchy of stable sets, the force laws (and any other laws responsible for the associations between fundamental and derivative properties) occupy a stratum above Λ but below the stratum containing the fundamental dynamical and conservation laws, the parallelogram of forces, and so forth. Thus, the associations between fundamental and derivative properties belong to a higher stratum than the laws describing whatever particle species there happen to be. Note well: whatever particle species there *happen* to be. Although which particle species there *are* is a matter of natural necessity, the closure law specifying those species as well as the laws identifying their characteristic properties lack the stronger variety of natural necessity possessed by the laws responsible for the associations between fundamental and derivative properties. In terms of that stronger necessity, the inventory of particle species is a matter of what happens to be, not a matter of what must be; in terms of the variety of necessity that is characteristic of the force laws, the closure law is an "accident."

Since the laws linking the fundamental and derivative properties belong to a more elite stable set than does the closure law, it is not a mere coincidence of the particular kinds of elementary particles there are that no two kinds are alike in their mass, charge, and spin but not in their magnetic moments. It is not a coincidence of the universe's inventory that a particle's mass, charge, and so forth suffice to entail a vast array of its other properties. The links between those fundamental determinables and their derivative properties transcend the determinates of those determinables that happen to be instantiated.

This transcendence is crucial to making the particle species qualify as members of the same family of natural kinds, because it entails that were there (or had there been) fundamental particles *other* than these particular kinds of quarks and leptons, their mass, charge, and so forth would likewise suffice (or would have sufficed) to distinguish them. For example, they would not possess (or would not have possessed) an up quark's mass, charge, and so forth but a different magnetic moment or half-life. Had there been fundamental particles falling outside the classes listed in table 3.1, their kinds would have been distinguished in the same way as actual kinds of fundamental particles are distinguished. The natural kinds of fundamental particle are themselves members of the same kind.

The pyramidal hierarchy crucial to this picture is funded by various subjunctive facts expressed by counterlegals with antecedents such as "Had there been a fundamental particle with the up quark's mass and spin but twice its charge," or "Had there been another kind of fundamental particle." Accounts of natural law according to which the laws are *metaphysically* necessary—such as a Shoemaker-style scientific essentialism, different versions of which Brian Ellis (2001) and Alexander Bird (2007) defend—will have some difficulty dealing with such counterlegals. Whereas I have emphasized that different laws exhibit different grades of natural necessity, and that these differences make certain laws transcend the idiosyncrasies of others, accounts of laws as metaphysically necessary assign to all laws the same, very strong grade of necessity. Such a view threatens to squash the pyramidal hierarchy flat by treating all laws as on a modal par.¹²

For example, suppose that Coulomb's law is essential to the electromagnetic force (or to electric charge); in any possible world with electromagnetic forces (or electric charges), they obey Coulomb's law. The essence of the electromagnetic force (or electric charge) is then responsible for the truth of various counterfactuals, such as "Had those two charged particles been twice as far apart, their electrostatic repulsion would have been onequarter as great." The essence of the electromagnetic force (or electric charge), being metaphysically necessary, is preserved under the metaphysically possible supposition of the particles being farther apart.

But suppose that the fundamental dynamical law, the conservation laws, the spacetime transformations, and so forth transcend the particular kinds of forces there happen to be. They belong to a stable set that omits Coulomb's and the other force laws. These other laws would still, then, have held even if Coulomb's law had not held-even if the electromagnetic force had not declined with the inverse square of the distance, let's say. What essence is available to fund *that* subjunctive fact? And more importantly, why is that essence preserved under the metaphysically impossible supposition of Coulomb's law being violated? That the laws are all metaphysically necessary is not enough to account for the different grades of necessity possessed by different laws. Of course, we could write in-by hand, as it were-that mass's essence (which includes the fundamental dynamical law) and momentum's essence (which includes its conservation) would withstand certain metaphysically impossible changes to the various kinds of forces there are. But this kind of maneuver seems hopelessly ad hoc; the game is no longer worth the candle.

By the same token, had there been different particle species, the actual laws by which the derivative properties are fixed by the fundamental properties would still have held. The antecedent of this counterfactual violates the closure law, so if that law is metaphysically necessary, then this antecedent posits a metaphysical impossibility. How can the essences of the fundamental properties explain why it is the case that *had* there been different particle species, the world would still have been populated by particles possessing charges, masses, magnetic moments, and so forth, and would still have been governed by the actual laws relating the fundamental and derivative properties? Those essences are not preserved under every metaphysically impossible supposition. Why are they preserved under this one?

An essentialist might try to account for the counterlegals I have mentioned by positing that under a given metaphysical impossibility, as many essences are preserved as is possible without contradiction. But this principle is dubious. Return to the table of elementary particles: the electron is the lightest charged particle, and electric charge is a conserved quantity. Therefore, the electron is stable; there is nothing that it can decay into. Suppose that it is part of the electron's essence to be stable and that it is part of the essences of electric charge and lepton number to be conserved. Had there been a charged species of lepton lighter than the electron, the electron would have been unstable; there would have been something into which it could decay without violating the conservation laws, which would still have held since those laws transcend the particular kinds of elementary particles there happen to be. Can the fact that the electron would have been unstable, had there been a lighter charged lepton, be accounted for by the proposal that under the countermetaphysical supposition of a charged lepton which is lighter than the electron, as many essences are preserved as is possible without contradiction? There is no contradiction in the electron's remaining stable, lepton number's remaining conserved, and electric charge's remaining conserved. But then we will not get the right answer: that it is *not* the case that the electron would still have been stable, had there been a charged lepton lighter than the electron—just as it is *not* the case that energy would still have been conserved, had there been no law to make its conservation mandatory.¹³

To understand the sense in which the various particle species belong to the same family, and the sense in which the associations between fundamental and derivative properties are laws in their own right, I have appealed to facts about what would have happened, had there been particles *not* belonging to the actual kinds. This is part of what I meant at the outset when I declared that it takes *more* than all of the actual kinds to make the world.

6 Natural Properties

I shall now turn to the final ingredient I identified in section 1 as making the classes of elementary particles into natural kinds: that the fundamental properties shared within a class and collectively differentiating the classes are *natural* properties. My remarks regarding natural properties will remain tentative and sketchy. Nevertheless, I believe it worthwhile to explore how far the foregoing considerations may take us toward understanding the notion of a "natural property," since even a glimmer of progress on this vexed topic would be valuable.

Since natural necessity is closed under logical consequence, it is a natural necessity that all things that are either electrons or up quarks have either a charge of 2/3 or a mass of about half an MeV. But that wildly disjunctive property is unnatural, despite figuring in this natural necessity.

One option is for the natural properties to be distinguished somehow *prior* to the natural necessities, as Lewis and Armstrong and (in a different way) Shoemaker believe them to be. Another option is for the natural properties to be the properties figuring in the laws, but then the laws must somehow be set apart from the other natural necessities; not all logical

consequences of laws are then laws (Fodor 1974, 109). If we pursue this option, we will have to regard membership in a nonmaximal stable set as associated with natural necessity, not lawhood; some members of such a set are naturally necessary, but are not laws.

Nevertheless, I find this option attractive, since the distinction between laws and mere natural necessities appears to play a role in scientific practice. For example, many physicists in the first decade of the twentieth century believed it to be no law, but merely a coincidence (albeit a logical consequence of the laws), that the correct equation for the black-body spectrum can be derived from the light-quantum hypothesis. Hence, these physicists did not regard other equations derivable from the light-quantum hypothesis as being confirmed to any degree by the empirical success of the equation for the black-body spectrum. Likewise, in the nineteenth century, some relations among atomic weights were regarded as no laws themselves, but as following from the laws. For instance, von Pettenkofer believed it to be a law that all alkane hydrocarbons differ in their atomic weights by integral multiples of 14 atomic-weight units, and a law that the atomic weight of nitrogen is 14 units, but a mere coincidence that all alkane hydrocarbons differ in their atomic weights by integral multiples of the atomic weight of nitrogen. We can readily understand why von Pettenkofer deemed this fact coincidental (albeit naturally necessary): alkane hydrocarbons contain no nitrogen.¹⁴

However, it is difficult to see how the laws can be picked out from among the natural necessities except by virtue of their involving natural properties or their underwriting scientific explanations. For example, neutrinos are electrically neutral: they do not repel each other electrostatically. Two neutrinos attract each other gravitationally or repel each other electrostatically not because they have masses or like charges, but rather simply because they have masses. This thought is closely related to the notorious footnote 28 in Hempel's and Oppenheim's landmark 1948 essay on explanation, where the conjunction of Kepler's and Boyle's laws is described as entailing but not explaining Kepler's laws.¹⁵ Whether a property's naturalness makes it explanatory or its explanatory power makes it natural, properties figuring in natural necessities may lack both naturalness *and* explanatory power, so the natural necessities alone seem unable to distinguish the natural properties.

I have appealed to certain subjunctive facts as being responsible for making certain truths naturally necessary. Can similar subjunctive facts be responsible for making certain properties natural by picking out the laws from all the natural necessities? Let's explore this line of thought. Earlier, I suggested that had Coulomb's law been violated sometime *before* today, then with Coulomb's law out of the way, there would have been nothing to keep Coulomb's law from being violated *after* today—just as it is not the case that energy would still have been conserved, had there been no law making such conservation mandatory. Coulomb's law restricted to times after today is merely a natural necessity, not a law. Coulomb's law is responsible for its truth, so it is not the case that it would still have held had Coulomb's law been violated sometime before today. This suggests that if Coulomb's law restricted to times after today were both genuine laws, so that being electrically charged before today and being electrically charged after today were somehow *both* natural properties, then had one of these laws *not* held, the other *would* still have held. They would be independent if they were genuine laws.

The mutual independence of various fundamental force laws is frequently assumed by philosophers and physicists describing how the universe would have been different had a given force been slightly different. For example, it is frequently remarked that had the strong nuclear force been somewhat weaker, it would have been insufficient to bind two protons together into one nucleus. This remark presupposes that had the strong nuclear force been slightly weaker, the mutual electromagnetic repulsion of two protons would have been no different. By the same token, had the gravitational force been slightly stronger, none of the stars in the "main sequence" (except the bluest) could have formed—but only because the strengths of the *other* forces would have been no different (Barrow and Tipler 1986, 322–327, Carter 1990, 132–133). Intuitively, under these counterlegal antecedents, "all other things" are held fixed—and the other forces count as genuinely *other* things.

Can such relations of mutual independence be used somehow to distinguish the laws from the other natural necessities? Take the stable set in the middle of the pyramid (in the previous figure)—a set that includes the force laws but omits the closure law giving the various natural kinds of particles there are. To generate this set from the stable set immediately above it in the hierarchy (which lacks the force laws), that higher set needs to have only a few members of the lower set added to it; the rest then follow logically. Let's call any few members of the lower set that would suffice to generate it from the higher set a "generator" for the lower set. What would be a generator for that set?

One possible generator is such that each member covers exactly one kind of force. Pretending (for the sake of simplicity) that the world is governed by something like classical physics, this generator's members would be such facts as

• At any moment, between any two point bodies with masses m and M, distance r apart, there is a force $F_g = GMm/r^2$

• At any moment, between any two point bodies with charges q and Q, distance r apart, there is a force $F_e = Qq/r^2$, etc.¹⁶

This generator's members are mutually independent in that had one member not held, the others would still have held.¹⁷

• At any moment, between any two point bodies at a distance r apart, with charges q and Q if the moment precedes today or with masses m and M otherwise, there is a force $F_e = Qq/r^2$ before today or there is a force $F_g = GMm/r^2$ otherwise

• At any moment, between any two point bodies at a distance r apart, with masses m and M if the moment precedes today or with charges q and Q otherwise, there is a force $F_g = GMm/r^2$ before today or there is a force $F_e = Qq/r^2$ otherwise

This generator's members are not mutually independent. It is not the case that had this generator's first member been violated, then all of its other members would still have held. Rather, were its first member violated, then either Coulomb's law or Newton's gravitational-force law would not be in force, so there would be nothing to enforce that law's remaining fragment, figuring in other members of the generator.

Another generator has members that subdivide the kinds of force too finely—members such as

- At any moment before today, between any two point-bodies with charges q and Q, distance r apart, there is a force $F_e = Qq/r^2$
- At any moment not before today, between any two point-bodies with charges q and Q, distance r apart, there is a force $F_e = Qq/r^2$

This generator's members also fail to be mutually independent. It is not the case that Coulomb's law would still have held after today, had it been violated sometime before today.

A final option is for a generator's members to subdivide the kinds too coarsely.¹⁸ One of its members might be the conjunction of Coulomb's law and Newton's gravitational-force law. That member *is* independent of the others (which concern other forces altogether). But this member is equivalent to the conjunction of two truths—one for each kind of

force—that are mutually independent, each belonging to the stable set being generated but not to the stable set (sitting one rung higher in the pyramid) from which it is being generated. The same applies to a generator consisting of a single truth, p, having all of the force laws as logical consequences: it can be split into mutually independent components.¹⁹

Here, then, we have a procedure for using the subjunctive truths to privilege a generator: by virtue of its members being mutually independent and incapable of decomposition into mutually independent components that belong to the stable set being generated but not to the stable set from which it is being generated. By belonging to a privileged generator, certain natural necessities qualify as laws; others are mere consequences of those laws. We would like it to turn out that roughly speaking, the natural properties (at least, those such as the "fundamental properties" in our table of elementary particles from section 1) are the determinates (e.g., 5.0 statcoulombs of electric charge) of the determinables (e.g., electric charge) figuring in a privileged generator. Wildly disjunctive properties, such as having a mass of 3 grams or an electric charge of 5 statcoulombs, will fail to qualify as natural. Insofar as subjunctive facts expressed by counterlegals (such as "Had Coulomb's law not held, then Newton's gravitational-force law would still have held") make certain properties natural, we have here another respect in which it takes more than the actual kinds to make the world.

Notice, by the way, that a generator subdividing the kinds too coarsely returns us to the problem raised by Hempel and Oppenheim's notorious footnote: why does a natural necessity that arbitrarily conjoins Kepler's laws and Boyle's law fail to function as a law in scientific explanations? Suppose that a given momentary force, occurring today, is electrostatic. The force's magnitude and direction follow from resting bodies' locations and charges at the given moment today, together with Coulomb's law. Of course, that deduction still goes through if Coulomb's law as a premise is replaced by a conjunction of Coulomb's law and Newton's gravitationalforce law. Admittedly, the Newtonian conjunct is superfluous for the conclusion to follow from the premises—but by the same token, the original deduction would still go through if Coulomb's law is replaced among the premises by Coulomb's law restricted to today. What makes the Newtonian conjunct explanatorily otiose whereas the part of Coulomb's law concerning the universe's history outside today contributes to the explanationwhen neither premise is needed in order for the explanandum to follow? An important part of the solution, I speculate, is that Newton's law and Coulomb's law are mutually independent, unlike Coulomb's law restricted to today and Coulomb's law restricted to the universe's history apart from today. Here we have the germ of a connection between natural properties, explanatory significance, unification, and the distinction between natural laws proper and mere logical consequences of those laws.

In emphasizing the mutual independence of the privileged generator's members, my approach exploits the fact that the various natural kinds of force are only "loosely connected," as Wigner put it in the passage I quoted in section 3. However, although the procedure I have just given rules out some generators, it does not privilege exactly one. Any truth logically equivalent to one of the truths in a privileged generator could substitute for it, and the result would still be a generator that passes muster according to our procedure; its members would be mutually independent and incapable of decomposition into mutually independent components. How, then, do we finish picking out laws, and hence natural properties, from among the natural necessities?

One possible solution is to consider the simplest of the generators that pass muster according to our procedure; perhaps the properties figuring in that generator (or the determinates of those determinables) are the natural properties. For instance, Coulomb's law can be expressed very simply in terms of the property of being a pair of point-charges repelling each other with a force equal to the product of their charges divided by the square of their separation r, if the charges have been at rest for a timespan of r/c. (Coulomb's law is that this property is possessed by every pair of point-charges.) But that property will not figure in simple expressions of other laws. On the other hand, although Coulomb's law is a bit more complicated when expressed in terms of properties such as separation and charge, these properties will also figure in simple expressions of some of the generator's other members. Taken as a whole, a privileged generator involving these properties is simplest. Perhaps the properties figuring in the simplest privileged generator are natural. Of course, several privileged generators may be equally, optimally simple. One uses mass and volume; another density and volume. One uses mass and velocity; another momentum and velocity. Perhaps figuring in any one of them suffices for naturalness.

Let me close with some caveats.

I do not mean to suggest that all actual natural kinds acquire that status by behaving just like the elementary-particle kinds I have been examining. Nor do I mean to suggest that it is metaphysically necessary for there to be laws of nature, much less natural kinds of elementary particles. For that matter, I do not contend that the actual elementary-particle kinds must derive their status as natural kinds in the manner I have described. It seems to me merely one possible route to this status. Alternatively, the force laws may determine the particular combinations of mass, charge, and so forth that species of elementary particles could have. (Perhaps by inserting into the force laws any combinations of fundamental properties other than those characteristic of the actual particle species, we end up dividing by zero or encountering probabilities greater than 1 when we try to compute particle trajectories.) In that event, it remains the case that the associations among the fundamental and derivative properties transcend the particular kinds of particles there are—but not because those associations would still have held, had there been other kinds of particles. Those associations are not byproducts of the particular particle species there happen to be; rather, those associations explain the kinds of particles there are.

And finally, if (as I tentatively suggested) a property's naturalness is fixed by various contingent subjunctive facts, then a property can be natural in one possible world but unnatural in another. This runs contrary to the view, shared by Shoemaker and Lewis (1986a, 60–61 n.44), that a natural property is necessarily natural. The motivation behind this view seems to be that whether some genuine feature is possessed by all and only a property's possible instantiations is a trans-world matter. On the other hand, perhaps any property's instantiations have *some* feature distinguishing them—being instantiations of that property—but whether a given property is natural depends on whether it is privileged by the laws (or their lawmakers). Natural properties, on this view, carve up the world as nature does. This seems to make good sense of the way we discover the natural properties in the course of discovering the laws. But as yet, I see no good way to ascertain whether a natural property is *necessarily* natural except as "spoils to the victor."

Notes

1. Lewis (1986a, 60; 1986b, ix–x) gives mass, charge, quark flavors, quark colors, and spin as among the perfectly natural properties, according to our best physics.

2. For greater elaboration of the argument in sections 2-4, see Lange 2009.

3. I examine possible examples in my 1993, 2000, and 2009. Lewis is a notable dissenter to this principle; I examine his arguments in my 2000 and 2009.

4. For example: Bennett 2003, 224; Foster 2004, 90–91; Pollock 1974, 201; Stalnaker 1984, 155–156; Swartz 1985, 53–54; van Inwagen 1979, 449–450.

5. For references to much of the early literature, see Berzi and Gorini 1969, 1518; for more recent references, see Pal 2003.

6. According to many proposed logics of counterfactuals, $p \square \rightarrow q$ is true trivially whenever p&q is true (a principle known as "centering"), and likewise for nested counterfactuals. If centering is correct, then each member of the set of all sub-nomic truths is trivially preserved under every sub-nomic supposition p that is true. But there are no sub-nomic suppositions p that are false and logically consistent with the set. (If p is a sub-nomic truths trivially possesses sub-nomic stability. On the other hand, if centering is false, then even the set of all sub-nomic truths may lack sub-nomic stability. I believe that centering fails in a universe where the laws are deterministic.

7. That the force laws belong to a stratum of law located (in the pyramidal hierarchy) beneath a stratum without them, but containing the fundamental dynamical law and its mates, is part of what motivated unease with the equality of inertial and gravitational mass in classical mechanics. Their equality (more precisely: their standing in the same ratio for all bodies—which, with suitable choice of units, can be made equal to 1) connects a force law to the fundamental dynamical law-a connection *between* the levels, threatening to disrupt the stability of the higher stratum: had the gravitational force been different, would the fundamental dynamical law have been different to compensate, preserving the fact that the gravitational acceleration of falling bodies is independent of their masses? In terms of the stronger variety of necessity possessed by the fundamental dynamical law, the gravitationalforce law (and hence the equality of inertial and gravitational mass) is an accident. ("[T]he equality of the two masses . . . was quite accidental from the point of view of classical mechanics" (Einstein and Infeld 1951, 227); "[It is] an accidental gift of nature" (Born 1961, 313).) But the equality of inertial and gravitational mass seems suspiciously nonaccidental—and not merely by hindsight assisted by general relativity. Hertz wrote (about 1884): "This correspondence must mean more than being just a miracle" (Blaser 2001, 2395).

8. Searle 1897, 341; cf. Hunt 1991, 91. Searle was an associate of J. J. Thomson at the Cavendish Laboratory in Cambridge.

9. In an undated letter (probably from early 1889) quoted in Nahin (1987, 126). Today FitzGerald is often remembered for his "contraction hypothesis" advanced to deal with the null result of the Michelson-Morley experiment.

10. That was Heaviside's view. He tried to derive an equation for the superluminal case.

11. Here is an alternative reconstruction of the argument that apparently involves no counterfactuals: it is a law that all forces have real-valued (rather than imaginary-valued) magnitudes (since F = ma, and accelerations and masses are real-valued),

It Takes More Than All Kinds

so it follows from Heaviside's law that a charged body's speed cannot exceed c—at least if it has been moving uniformly for a long time in the presence of another charged body.

But how does this argument arrive at the more general moral drawn by Searle and FitzGerald: that no charged body can move superluminally? The argument might continue: if the laws prohibit a charged body's speed from exceeding c when it is moving uniformly for a long time in the presence of another charged body, then presumably this is not a special case but applies even to non-uniform motion, and even in the absence of another charged body. I suggest that this thought ought to be cashed out as follows. Had there been a charged body moving superluminally but non-uniformly or in the absence of another charged body, then there might have been a charged body in uniform superluminal motion in the presence of another charged body. In other words: let *m* be that no charged body's speed exceeds c when it has been moving uniformly for along time in the presence of another charged body. (That *m* is a law is the narrower conclusion that followed above from Heaviside's law.) The following is false: had there been a charged body moving superluminally but non-uniformly or in the absence of another charged body, then *m* would still have held. The laws' stability is thereby threatened. So for *m* to be a law, the laws must prohibit charged bodies from any superluminal motion, not merely from uniform superluminal motion in the presence of another charged body.

Thus, by looking at the counterfactuals, we capture how the evidence for Heaviside's law counts also as evidence for a law prohibiting charged bodies from moving superluminally. The extent to which scientists were prepared to generalize the prohibition against charged bodies engaging in uniform superluminal motion in the presence of another charged body corresponds to the counterfactuals that scientists were prepared to accept specifying conditions under which there might have been a charged body in uniform superluminal motion in the presence of another charged body. (Interestingly, Searle, FitzGerald, et al. limited the prohibition to *charged* bodies. They seem not to have held that had an uncharged body moved superluminally, then some charged body might have done so too. Counterfactuals thus help to reveal the limits of their argument.)

12. I offer additional critiques of Ellis 2001 in Lange 2004, 2005.

13. That the force laws (and the other laws responsible for the associations between the particles' fundamental and derivative properties) would still have held, had there been a charged lepton species lighter than the electron, might appear to conflict with the fact that the electron would no longer have been stable, had there been a charged lepton species lighter than the electron. But there is no conflict here, since the laws responsible for the associations between the fundamental and derivative properties need not give a particle's derivative properties as a function solely of its own fundamental properties. For instance, the laws responsible for the associations between the fundamental and derivative properties may entail, given all of the kinds of particle there are, how likely it is that a given particle (considering its fundamental properties) will decay into one or another of the other kinds of particle over the course of a given timespan. An addition to the particle species there are may then add a decay mode and alter the distribution of probabilities among the former decay modes, but those chances would still be determined by the same laws.

14. For more discussion and references, see Lange 2000, chap. 5.

15. In the 1965 reprint, it is note 33, on page 273.

16. This is clearly a vast oversimplification even of classical physics, but I think it is harmless for my purposes here.

17. Of course, under some suppositions that are logically inconsistent with the generator's first member, the generator's second member is not preserved. The suppositions relevant to the members' "mutual independence" are simply, "Had [some other member of the generator] not held."

18. Of course, a single generator could also cross-cut some kinds while subdividing others too finely.

19. Compare Feynman on the bogus unification created by "the unworldliness of mechanical effects" (see Lange 2002, 202).

References

Barrow, J. D., and F. J. Tipler. 1986. *The Anthropic Cosmological Principle*. Oxford: Clarendon.

Berzi, V., and V. Gorini. 1969. Reciprocity principle and Lorentz transformations. *Journal of Mathematical Physics* 10:1518–1524.

Bennett, J. 2003. A Philosophical Guide to Conditionals. Oxford: Clarendon.

Bird, A. 2007. Nature's Metaphysics: Laws and Properties. Oxford: Clarendon.

Blaser, J. P. 2001. Remarks by Heinrich Hertz (1857–94) on the Equivalence Principle. *Classical and Quantum Gravity* 18:2393–2395.

Born, M. 1961. Einstein's Theory of Relativity. New York: Dover.

Carter, B. 1990. Large number coincidences and the anthropic principle in cosmology. In *Physical Cosmology and Philosophy*, ed. J. Leslie. New York: Macmillan.

Einstein, A. 1935. Elementary derivation of the equivalence of mass and energy. *Bulletin of the American Mathematical Society* 41:223–230.

Einstein, A., and L. Infeld. 1951. *The Evolution of Physics*. New York: Simon & Schuster.

Ellis, B. 2001. Scientific Essentialism. Cambridge: Cambridge University Press.

Feynman, R. 1967. The Character of Physical Law. Cambridge, Mass.: MIT Press.

Fodor, J. 1974. Special sciences (Or: The disunity of science as a working hypothesis). *Synthese* 28:97–115.

Foster, J. 2004. The Divine Lawmaker. Oxford: Clarendon.

Goodman, N. 1983. *Fact, Fiction, and Forecast,* 4th ed. Cambridge, Mass.: Harvard University Press.

Hempel, C. G., and P. Oppenheim. 1948. Studies in the logic of explanation. *Philosophy of Science* 15:135–175. Reprinted with some amendments in Hempel, *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press.

Hunt, B. 1991. The Maxwellians. Ithaca: Cornell University Press.

Lange, M. 1993. When would natural laws have been broken? *Analysis* 53:262–269.

Lange, M. 2000. Natural Laws in Scientific Practice. New York: Oxford University Press.

Lange, M. 2002. An Introduction to the Philosophy of Physics: Locality, Fields, Energy, and Mass. Malden, Mass.: Blackwell.

Lange, M. 2004. A note on scientific essentialism, laws of nature, and counterfactual conditionals. *Australasian Journal of Philosophy* 82:227–241.

Lange, M. 2005. Reply to Ellis and to Handfield on essentialism, laws, and counterfactuals. *Australasian Journal of Philosophy* 83:581–588.

Lange, M. 2009. Laws and Lawmakers. New York: Oxford University Press.

Lévy-Leblond, J. 1976. One more derivation of the Lorentz transformations. *American Journal of Physics* 44:271–277.

Lewis, D. 1984. New work for a theory of universals. *Australasian Journal of Philosophy* 61:343–377.

Lewis, D. 1986a. On the Plurality of Worlds. Malden, Mass.: Blackwell.

Lewis, D. 1986b. Philosophical Papers, vol. 2. Oxford: Oxford University Press.

Lewis, D. 1999. Humean supervenience debugged. In *Papers in Metaphysics and Epistemology*. Cambridge: Cambridge University Press.

Nahin, P. 1987. Oliver Heaviside: Sage in Solitude. New York: IEEE Press.

Pal, P. 2003. Nothing but relativity. European Journal of Physics 24:315–319.

Penrose, R. 1987. Newton, quantum theory, and reality. In *Three Hundred Years of Gravitation*, ed. S. Hawking and W. Israel. Cambridge: Cambridge University Press.

Pollock, J. 1974. Subjunctive generalizations. Synthese 28:199–214.

Searle, G. F. C. 1897. On the steady motion of an electrified ellipsoid. *Philosophical Magazine* 44 (ser. 5):329–341.

Shoemaker, S. 1984. Causality and properties. In *Identity, Cause, and Mind*. Cambridge: Cambridge University Press.

Stalnaker, R. 1984. Inquiry. Cambridge, Mass.: MIT Press.

Swartz, N. 1985. The Concept of Physical Law. Cambridge: Cambridge University Press.

van Inwagen, P. 1979. Laws and counterfactuals. Noûs 13:439-453.

Wigner, E. 1972. Events, laws of nature, and invariance principles. In *Nobel Lectures: Physics 1963–1970*. Amsterdam: Elsevier.

4 Lange and Laws, Kinds, and Counterfactuals

Alexander Bird

In this essay I examine and question Marc Lange's account of laws, and his claim in the preceding chapter that the law delineating the range of natural kinds of fundamental particle has a lesser grade of necessity than do laws connecting the fundamental properties of those kinds with their derived properties.

1 Lange on Laws

Regularity theorists about laws face the following problem: many regularities are true, but only some of them correspond to laws. Consider:

- (S1) All bits of copper conduct electricity;
- (S2) All lumps of gold have a volume of less than a cubic mile.

These two generalizations are both true. However, of

(N1) It is a law that all bits of copper conduct electricity;

(N2) It is a law that all lumps of gold have a volume of less than a cubic mile,

only the first is true. The trick for the regularity theorist is to pick from among the regularities those that are like (S1), for which the addition of the *nomic* operator 'it is a law that . . . ' is truth-preserving, while excluding those like (S2), for which that operator yields a falsehood.

Marc Lange is no regularity theorist, but his starting question is very similar to that posed for the regularity theorist. Consider all the true propositions that are *not* facts about which laws there are—propositions that can be expressed without using vocabulary such as 'law', 'nomological' and the like. Call this set ' Σ '. Propositions such as (S1) and (S2) are in Σ because they are true and don't concerns laws. (N1) is excluded because although true, it concerns a law; (N2) is doubly excluded because it concerns an alleged law and is false. Σ is the set of what Lange call the "sub-nomic"

facts. They are sub-nomic because they don't concern which laws there are; they are not expressed using nomic vocabulary. Some of the facts in Σ are those like (S1) for which the nomic operator is truth-preserving. These are, of course, the *laws*.¹ I will use 'accident' as the term for the remaining members of Σ that are like (S2), for which the nomic operator takes us from a truth to a falsehood.

How do we separate the laws in Σ from the accidents? Lange's approach is to exploit their superior stability under counterfactual suppositions. Consider:

(G1) Were Bill Gates to want some non-conducting copper, all bits of copper would still conduct electricity;

(G2) Were Bill Gates to want a lump of gold with a volume greater than a cubic mile, all lumps of gold would still have a volume of less than a cubic mile.

These are counterfactual (or better, subjunctive) conditionals of a mildly unusual sort, where although the antecedent is counter-to-fact (i.e., false), their consequents are true. But they do not require any special kind of treatment. We can see that while (G1) is certainly true, (G2) is quite possibly false.

Powerful though Bill Gates is, there are some accidents that would remain true, whatever he wanted. Using the Gates antecedent allows us to divide Σ into two sets: one includes the laws and other things that Gates is not powerful enough to change, the other set includes the things Gates could possibly change. We may be able to generalize the strategy that distinguishes between (G1) and (G2) by considering more powerful counterfactual antecedents in conditionals such as (G1) and (G2). On the other hand, we don't want the counterfactual supposition to be too powerful:

The transition elements are all non-conductors $\Box \rightarrow$ all bits of copper conduct electricity

looks to be false. Think of our conditionals as having the following form:

 $(T) \quad C \square \rightarrow F$

where 'C' denotes a counterfactual supposition and 'F' denotes an actual fact. The generalized Gates test suggests that (T) will always be true when F is a legal truth and C is an accident, but will sometimes be false when both are nomological facts or both are accidents. Lange's account of lawhood proposes that the laws are those that are stable under *any* counterfactual supposition that is an accident (or its negation).² As an account

of law, that looks to be problematic because it employs a concept 'accident' that is itself a concept of the kind we are wishing to elucidate: accidents are precisely the members of Σ that are not legal truths.

One of Lange's (many) smart moves is that he shows how to avoid this circularity. He does this by defining 'sub-nomic stability'. Let Γ be a subset of Σ that includes all its consequences in Σ . Consider propositions of the form:

$p \square \rightarrow m$

where *p* is the negation of some member of the complement of Γ in Σ and *m* is some member of Γ . The first pass at defining sub-nomic stability is to say that Γ is sub-nomically stable iff every proposition of this form is true. We can spell out sub-nomic stability this way: when Γ is sub-nomically stable, its members would still be true under every counterfactual supposition that is consistent with Γ . (We're still confining ourselves to propositions that are members of Σ , the sub-nomic truths, plus their negations.)

Lange argues that if *g* is an accident, than the only sub-nomically stable set containing *g* is Σ . Do the laws form a special sub-nomically stable set? Let Λ include all the laws, plus the metaphysical, mathematical, and logical necessities, but excluding any accidents. This set is sub-nomically stable for the reasons we have seen above. The members of Σ *not* in Λ are just the accidental truths. The counterfactual supposition that any of these is false cannot change the laws—not the mathematical or logical truths, etc. So we see that Λ is a nontrivially, sub-nomically stable set. Does that suffice to pick out Λ ? No, because there are other sub-nomically stable sets. Take the mathematical and logical truths (but not the laws): these are stable under any supposition that an accident or law is false. So how do we pick out Λ ? And furthermore, how do we pick out the laws from within Λ (in contrast to the logical and mathematical truths, and so on)?

Lange's next clever move is to show that the sub-nomically stable sets form a hierarchy. That is, for any pair of sub-nomically stable sets, one is a subset of the other. We can thus pick out Λ as the largest sub-nomically stable set that is not Σ .

Lange's proof is important, and it is worth seeing how it works. Consider a world in which there are two basic sub-nomic contingent facts, F and G. Trivially, the set {F, G} (i.e., Σ for this world) is sub-nomically stable. Could both {F} or {G} be sub-nomically stable? We shall see that one or the other may be sub-nomically stable, but not both. Let us ask what would be the case were one of F or G false, that is,

$\neg F \lor \neg G \Box \rightarrow ?$

Would F be false or would G be false? Or both? This is to ask: which of F or G is more *modally fragile*? In the actual world, both F and G are true. We consider worlds at increasing distance from the actual world; eventually we will come to a world in which one or the other of F and G is false. That is the nearest world in which $\neg F \lor \neg G$ is true. If F is false but G true in this world, then F is more modally fragile (less modally robust) than G. For example, note that it is true, F, that I am wearing red socks, and it is also true, G, that I am under 2.5m tall (in fact, I am less than 1.9m tall). Which is the nearest possible world in which one or the other of these is false? Clearly it is a world in which I do not wear red socks, but in which I remain under 2.5m tall. In this case the following are true:

- $(i) \quad \neg F \lor \neg G \ \square \rightarrow \neg F$

F is thus more modally fragile than G.

Note that $\neg F \lor \neg G$ is consistent with F and is consistent with G. Let's now ask whether either of {F} or {G} is sub-nomically stable. (i) above tells us that {F} is not sub-nomically stable. If both {F} and {G} were to be sub-nomically stable, we would need:

(iii)
$$\neg F \lor \neg G \square \rightarrow F$$

(iv)
$$\neg F \lor \neg G \square \to G$$
.

But clearly it cannot be that both (iii) and (iv) are true. It is this impossibility that means that at most one of {F} or {G} is sub-nomically stable. Lange's proof depends on this principle, extending the case to any sets that are not related by the subset relation.

2 Lange on Kinds

The fact proved by Lange, that the sub-nomically stable sets form a hierarchy ordered by the subset relation, means that we can order different kinds of (sub-nomic) propositions by their place in the hierarchy. So Σ , the set of all sub-nomic truths is, trivially, a sub-nomically stable set. So also is Λ , the sub-nomically stable set next-largest in size to Σ . This contains the laws, but also the mathematical and logical truths, and so forth. M, the set of metaphysical, mathematical, and logical truths also forms a subnomically stable set. The different levels in this hierarchy correspond to different grades of necessity. 'Natural necessity' is the grade possessed by all of Λ , 'metaphysical necessity' is the grade possessed by M, and so on. It may be that the laws properly speaking are the truths in Λ that are not in M: that is, the laws of nature are those truths that are naturally necessary *without* being metaphysically necessary. Call the laws-properly-speaking ' Λ ''.

More interestingly, within the laws, Λ' , Lange claims there are divisions that correspond to different *kinds* of law with differing grades of natural necessity. For example, Lange thinks that there is a proper subset of Λ' that contains the fundamental dynamical laws, such as Newton's second law of motion and the force-composition law (the 'parallelogram of forces'), but that excludes the laws such as Coulomb's law and the law of gravitation which tell us what particular forces there are. Similarly, he holds that the conservation laws have a special status that transcends the particular force laws.

Something similar seems to be true concerning the natural kinds of fundamental particle. These particles have certain basic, fundamental properties and certain derived properties; these basic and derived properties are linked by natural law. Thus the muon has a certain spin, charge, and mass; these are its fundamental properties. Among its derived properties are its magnetic moment and half-life. The laws linking the fundamental and derived properties in some sense transcend the actual fundamental particles there are. Even if there had been other fundamental particles, these laws would still have held. That allows us to think counterfactually about what would have been the case had there been other fundamental particles. For example, Lange informs us, had there been a lepton less massive than the electron, then the electron would not have been stable, because there would have been some other particle into which it could decay. Let Ω be the true proposition that lists all the fundamental particle kinds there are and asserts that these are all the kinds of fundamental particle there are. Ω looks to be in $\Lambda'\!-\!\Omega$ would remain true whatever accidents happened to be otherwise. On the other hand, as we have just seen, the laws linking fundamental properties of the members of Ω to their derived properties are laws of a more robust sort, which would remain true even if Ω were false. This suggests that Ω possesses a low or event the lowest grade of natural necessity.

While I regard Lange's remarks on kinds to be stimulating and instructive, I think that their significance as an account of law and natural necessity remains up for discussion. First, the fact that

If *p* were not the case, *q* would still be the case

does not show that q has a higher grade of necessity than p. The truth of this conditional is consistent with the propositions just being independent. Lange doesn't suggest otherwise, but much of his discussion is framed, as above, in terms of one set of laws remaining as they are even if some other set were different. Consider the following claims made by Lange:

(A) Had there been elementary particles not belonging to the actual kinds of quarks and leptons, then the same laws connecting the fundamental properties with the derivative properties of elementary particles would still have held; the force laws would have been no different.

(B) Such subjunctive facts make stable a proper subset of Λ containing the laws connecting the fundamental properties with the derivative properties of elementary particles.

I want first to question whether (A) is true, and second, whether (B) follows from (A), and third, whether (B) is as significant as Lange's presentation seems to suggest.

Is (A) in fact true? One reason why it might not be is the following: we do not yet know what the fundamental laws of the universe are. Perhaps they form a very small, highly integrated set of laws, such that they together explain both the range of kinds of particle and the laws that connect the fundamental properties with the derivative properties. That is not implausible. The standard model was itself able to predict the existence of the W and Z bosons, the gluon, and the top and charm quarks, and their properties—indicating that Ω is not a fundamental fact, but a derived law. The force-mediating particles in particular are not independent of the force laws and other laws that govern the behavior and nature of the fundamental particles. If one set of laws was responsible for both the existence of the particles and their natures (including their derived properties), then (A) would be false. Different fundamental particles would have required different fundamental laws, as a consequence of which the laws connecting the fundamental properties with the derivative properties would probably also be different.³ Something similar may be said of other sets of kinds. There are only so many chemical elements and their isotopes; there are conceivable isotopes of elements that do not exist. The existing isotopes have various physical properties that derive from their respective structures. What explains why there are just *these* isotopes and not others? It is primarily the strong nuclear force. For example, there is no helium-2 (diproton, an isotope with just two protons and no neutrons) because the actual value of the strength of the strong force prevents its existence. Had the strong force been 2 percent more powerful, then helium-2 would have existed. It is also the strong force that explains the half-lives of the isotopes that do exist.⁴ So statements such as:

Had helium-2 existed, the half-life of polonium-210 would not have been any different

are false. For helium-2 to exist, the strong force would have to be stronger, and as a consequence the half-life of polonium-210 would also be different.

One response to the above is to regard my argument against (A) as invalid, since it implicitly appeals to a backtracking counterfactual. If I say 'if Lucy had not been at the party, I would have had a dull evening', that claim cannot be denied on the ground that in any (nearby) circumstances that Lucy did not go to the party she would have told me in advance and I would have not gone to the party, but would have done something else of interest. That reasoning is backtracking and deemed illegitimate by Lewis and others. Likewise, the reasoning that if kinds of particles were different then the fundamental laws would also be different, seems to have a backtracking element to it—and thus ought to be outlawed.

Should the argument against (A) be outlawed on backtracking grounds? First, the rejection of all backtracking reasoning and counterfactuals is contentious. Second, it is difficult to reject it in this case. Lewis avoids backtrackers by introducing small miracles-for example, a change in the firing of Lucy's neurons that causes her to not enter the party just before she arrives. This enables the antecedent of the counterfactual to hold, without changing the world's earlier history. That thus allows me to be at the party, unaccompanied and bored. The small miracle (which incidentally, makes some limited backtrackers true) occurs at a particular moment in time, too late for it to affect my attendance at the party. But what is the analogue to a small miracle in the case of (A)? The antecedent and consequent concern eternal truths, so it is difficult to see how anything like a small miracle could occur. An analogue would be something that allowed the antecedent to be true—that is, for there to be different particle kinds but without this happening as a result of any change to the fundamental laws. But on the supposition that Ω , the actual truth about the particle kinds, is a logical (or at least metaphysically necessary) consequence of the fundamental laws, nothing can fill that role. In short, small miracles may be able to break a chain of *causal* consequence but they cannot break a chain of *logical* consequence, and that is what would be required here. So I am not sure that the backtracking argument against (A) can be rejected on that ground. That argument is speculative, but it does suggest that we cannot take the truth of (A) for granted.

Note also that were the argument against (A) rejected on backtracking grounds, it would undermine either the inference from (A) to (B) or the significance of (B) itself. Let us say that a conditional:

If p had not been the case, q would still have been the case

is held to be true on the grounds that its denial would require a backtracking argument. If so, it is difficult to see that the truth of the conditional shows that q is modally more robust than p and that p is modally more fragile. As pointed out above, the truth of that conditional is consistent with the truth of:

If q had not been the case, p would still have been the case.

Perhaps the denial of *either* counterfactual requires backtracking. The relative modal fragility of p and q is not decided by such conditionals. What we need is an answer to:

If one or the other of *p* or *q* were false, which would it be?

In our case, that is the question:

If the particle kinds were different, or the (conjunctive) law relating the fundamental and derived properties of those kinds were different, which would it be?

It is not clear that the answer is that the particle kinds would be different (but the connecting law would be the same). And if we are allowed backtrackers, the answer might be that in any world where one is different, the other is different also.

Note also that the exclusion of backtrackers, even temporal ones, threatens to undermine the significance of (B). Let e be the first event in the history of the universe—an event that is uncaused, unexplained, etc. (Perhaps the Big Bang is such, perhaps not.) For any other event, d, in the history of the universe, the following is true:

If *d* had not occurred, *e* would still have occurred.

As a result, if Λ is a sub-nomically stable set, then the set that is the union of Λ and $\{e\}$ (and their consequences in Σ) is also sub-nomically stable. Thus *e* would have a grade of necessity not shared by any other event in history. While the idea that the initial conditions of the universe have a special status akin to that of the laws is not itself objectionable, it would be odd if that conclusion were reached simply on the basis of the rejection of backtracking counterfactuals.

Let us now look a little more at the idea of a sub-nomically stable set. According to Lange, the idea is revealing because, thanks to the proof sketched above, it sorts sub-nomic claims into a hierarchy while the different levels in this hierarchy correspond to different grades of necessity. Putting aside the caveats made above, let us concede that Ω , the statement about what particle kinds there are, is a member of Λ , and so has some grade of natural necessity and is not an accident. On the other hand, if Lange is right, it has a very low grade of natural necessity—a lower grade than, for example, the symmetry and conservation laws. Is that fact revealing?

I want to suggest that it is perhaps less revealing than it might at first appear to be. If the hierarchy of sub-nomic sets sorted them into a smallish number of grades of necessity, corresponding to where we intuitively see potential differences of necessity (accidents, natural laws, metaphysical truths, logical truths), then that might well be significant. And if we found a small number of additional divisions—say, sorting the laws into two or three different levels—then that would be an interesting discovery. On the other hand, if there are *many more* levels than we intuitively perceive, then one might find the hierarchy less significant. I will suggest that the latter might be the case.

The proof that sub-nomically stable sets form a hierarchy turned on the fact that the relation of 'being modally as robust or more robust than' provides a total order on all facts. Now let us suppose that it is also a well-ordering. Let Δ be a sub-nomically stable set. Let D be that member of Δ with the property that is it more modally fragile (i.e., less robust) than any other member of Δ . That is, the nearest possible world in which *any* of Δ is false, is a world where D is false *but* the remaining members of Δ are all true. Now consider Δ –D. Let E be any member of this set. D will be more modally fragile than E (by definition of D), so:

 $\neg D \lor \neg E \square \rightarrow \neg D$

$$\neg D \lor \neg E \square \rightarrow E.$$

This shows that Δ –D is also sub-nomically stable. Now consider the least robust member of Δ –D: D'. By the same reasoning, the set Δ –D–D' is also sub-nomically stable. Repeating this construction leads to a series of sub-nomically stable sets, each one a subset of its predecessor formed by removing its most fragile member. It might seem that we can generalize this for cases which are not well-ordered by robustness (but are still totally ordered).

Let λ mark some position in the ordering of facts by modal robustness, and let Π_{λ} be the subset of the sub-nomically stable set Π that is formed by removing all facts that are at level λ or below: then Π_{λ} looks as if it is sub-nomically stable also. For every fact *not* in Π_{λ} is more modally fragile than every fact *in* Π . If that is correct, then there will not be just a few sub-nomically stable sets—but vast numbers of them. And correspondingly, *grades* of necessity will be multiplied to the same extent as such sets.

There is a fallacy in the preceding reasoning, however. I have assumed that it if both A and B are more fragile than every member of some set Π , that is sufficient to show that we can draw a distinction between the modal grade of A and B on the one hand and the members of Π on the other. But matters are not so simple, because the following inference is invalid:

 $A \square \rightarrow C \land B \square \rightarrow C \text{ therefore } A \land B \square \rightarrow C.$

For example, I may suffer from an illness for which one pill of a powerful drug will cure me, while I know that taking two pills would result in overdose and kill me. I have two pills, A and B, in front of me. 'If I were to take pill A, I would recover' and 'If I were to take pill B, I would recover' are both true, but 'If I were to take pill A and to take pill B, I would recover' is false.

More generally, from:

 $\forall q \; \forall p \; (q \text{ is atomic } \land q \text{ is consistent with } \Pi \land p \in \Pi) \rightarrow q \Box \rightarrow p$

it does not follow that:

 $\forall q' \forall p \ (q' \text{ is compound } \land q' \text{ is consistent with } \Pi \land p \in \Pi) \rightarrow q' \Box \rightarrow p.$

That is, although every member p of Π is such that for every *atomic* proposition q consistent with Π , it is true that if q were the case, p would still be the case, it does not follow that the same is true for every *compound* proposition q'. As a consequence, sets of propositions of a range of modal fragility must be banded together to form a sub-nomically stable set: there isn't as much variation in grades of necessity as there is in modal robustness or fragility.

Nonetheless, this fact revealed in the preceding paragraph seems to be a consequence of the causal complexity of the actual world. Consider a much simpler world in which basic events happen at one-second intervals, such that e_i causes e_{i+1} and so on, governed by a simple causal law. In such a world it would seem that

If e_m had not occurred, e_n would still have occurred

is false if $m \le n$ and true otherwise. In such a world, later events are more fragile than earlier ones. Furthermore, let Σ be the set of all sub-nomic facts. Let Σ_i be the set formed from Σ by excluding all events after e_i . The sequence of sets {... Σ_{i+1} , Σ_i , Σ_{i-1} , ...} is a hierarchy of sub-nomically stable sets.

What are we to say then about laws concerning the kinds there are, such as Ω ? Assume that Ω is a member of a sub-nomically stable set that is lower in the hierarchy than some other sub-nomically stable sets that include other, more robust laws (such as the symmetry laws). What does this show? It might just show that the different laws have different degrees of modal fragility; that in turn might reflect not a different grade of necessity but the logical structure of a law. It might seem that a force law, such as Coulomb's law, could easily fail to hold—for example, in a world with a slightly different value of the permittivity of free space. Likewise laws concerning the range of kinds there are might, with relative ease, have been otherwise—as mentioned above, helium-2 would exist in a world which (it seems) is only slightly different from ours. In these cases there are similar, 'nearby' laws that would be true if the actual ones were not true. But for the structurally simpler symmetry laws, this is arguably not true.⁵

3 Conclusion

Marc Lange has given us an elegant account of what the laws are which brings with it the prospect of deep insights into the nature of necessity, showing how there are different grades of necessity. The laws are necessary, but less so than the metaphysical truths. Within the laws there are different grades of necessity. And it appears that laws stating which kinds there are have a lower grade of natural necessity than some of other species of law—for example, the apparently more general laws linking the fundamental properties of these kinds and their derived properties.

In this essay, I have not rejected this view outright, but I have given several reasons to wonder whether it has been firmly established. First, one might wonder whether our intuitions concerning the relative modal robustness of the laws really reflect their natures. Could these intuitions rather reflect our ignorance of the natures of the laws? It seems to me that there is ample room for the possibility that the very same laws that are responsible for the existence and non-existence of kinds are also the laws responsible for fixing the derived properties of those kinds. That is in fact the case for the existence of isotopes and properties such as their half-lives, both of which are determined by the strong force: it seems at least plausible that the same is true for the fundamental particles.

Second, let us grant that it is true that were there different kinds of particle, the laws linking their fundamental and derived properties would not be different. While suggestive, it is far from clear that this supports the idea that the kinds law belongs to a different sub-nomically stable set from the property laws.

Third, we need to consider carefully the exact import of the hierarchy of sub-nomically stable sets. That hierarchy is closely related to the ranking of facts by their modal robustness. If the two were exactly the same, then the hierarchy of sub-nomically stable sets would be of little interest, and there would be infinitely many grades of necessity. But the two are not the same. The accidents, which vary regarding their modal robustness all have the same grade of necessity. On the other hand, this fact may reflect the causal complexity of the accidents in our world. In simpler worlds, there are different grades of necessity among the accidents corresponding to their differing degrees of modal robustness. It is not clear to me that the same isn't true of the laws of the actual world.

Notes

1. Note that for Lange, like Lewis and other regularity theorists, but unlike Armstrong, laws do not have a different logical form from other kinds of sub-nomic facts.

2. I say "or its negation" since I defined 'accident' to be factive, as a result of which the negations of accidents are not themselves accidents.

3. I say "probably" because a difference in fundamental laws does not logically necessitate a difference in all derivative laws. But the important ones will differ.

4. In cases of fission and alpha-decay, not beta-decay, which is explained by the weak nuclear force.

5. Compare this case: John is 210 cm tall. 'John is tall' is more robust than 'John is 210 cm \pm 5 mm'.

5 Are Fundamental Laws Necessary or Contingent?

Noa Latham

A central issue in metaphysics is whether there are any necessary relations in nature. The Humean answer is that there are not, so that laws of nature are contingent. The opposing view, that laws of nature are such necessary relations, offers a very different metaphysical picture of the universe. In this essay I primarily address a dispute among non-Humeans as to whether laws of nature are metaphysically necessary, or metaphysically contingent with a weaker kind of necessity, commonly labeled natural, nomological, or nomic necessity. I call the parties to this dispute necessitarians¹ and contingentists.² Restricting the scope of the discussion to fundamental laws, I take up the debate under the assumption that all fundamental properties are dispositional or role properties in part I, arguing that the dispute would then be purely verbal. In part II, I assume that there are categorical intrinsic properties as well as dispositional properties and examine the relation between them. And in part III, I return to the debate between necessitarians and contingentists under the assumption that there are both dispositional and categorical fundamental properties. I conclude that these necessitarian positions can again be recast as contingentist, but that there are some unequivocally contingentist positions that despite being ontologically more complex are to be preferred because they are less mysterious.

I take laws to be facts construed as true propositions, rather than statements expressing those facts, though sometimes I explicitly discuss lawstatements. Fundamental facts comprise fundamental laws and facts presenting the distribution of fundamental properties in spacetime. To be fundamental, these facts must provide a metaphysical supervenience base for all other facts, i.e., the remaining facts must hold in all possible worlds with these facts. It looks quite plausible that all the fundamental dynamic laws of our world take the form envisaged by field theories that concern the distribution of fundamental nonkind properties such as mass and
charge throughout spacetime. If so, these laws do not involve fundamental kinds, though kinds such as electrons may feature as nonfundamental aspects of fields. If there were fundamental dynamic laws involving fundamental kinds, they would have to be integrated with the fundamental laws of field theory involving nonkind properties, which presumably also exist. It is hard to see how this could be accomplished other than by treating laws involving kind properties as hierarchically prior and overriding the laws of field theory, rather as laws involving biological and mental kinds have sometimes been thought to override fundamental physical laws.³ It looks unlikely that empirical reasons will be encountered to warrant embracing such a bulky hierarchical system of fundamental dynamic laws for our world. So carving nature at its most fundamental joints will most likely not require sorting particulars into kinds.

Throughout the essay I focus on the example of the particle mechanics gravitational law (PMG), which does not involve fields but has the merits of simplicity and familiarity.

(PMG) All particles of mass m attract particles of mass m* with a force of Gmm*/d², where G = 6.673×10^{-11} m³ kg⁻¹ s⁻².

I shall not examine to what extent the discussion would generalize to the case of fields and laws involving fundamental kinds.

My discussion henceforth is thus narrower than familiar discussions of laws of nature. I shall not examine nonfundamental laws such as special science laws, ceteris paribus laws of physics, and thermodynamic laws. I see it as preferable in discussing laws of nature not to run together what I take to be very different types of laws in quest of a unified account.

I

Throughout this essay I assume a distinction between two types of properties akin to those commonly referred to as dispositional and categorical. Some use has been made of this in distinguishing two types of fundamental physical property and assessing the relation of each to phenomenal properties.⁴ A common thought is that physical theory deals only with dispositional properties, but that dispositional properties require categorical bases.

I prefer to speak here of intrinsic natures, or as C. B. Martin puts it, the qualitative side of properties, ⁵ rather than of categorical bases or properties. That is because shape and size are often given as paradigmatic categorical properties, and I take these to be definitive of fundamental particulars. The intrinsic natures, if they exist, do not include the size and shape of spatial

regions, but are the properties manifested by spatial regions of a given shape and size, or manifested by particulars fully occupying those regions. In particle mechanics and field theories, intrinsic natures, if they exist, are manifested by spatiotemporal points or particulars fully occupying those points.

I am not completely happy using the term 'dispositional' either, preferring to talk of "role properties" where the roles are constituted by relations, including temporal relations. Dispositional properties are taken to involve temporal relations to the future but not the past, and I would like to talk more generally of properties of particulars that relate them to particulars at different times, leaving it open whether those times are future or past. I also wish to leave the modal status of these relational properties open, allowing that they could involve the simple material conditional favored by a Humean, metaphysical necessity, or some intermediate strong nomic necessity. I shall briefly return at the end of this essay to consider the modal status of these properties. But the main focus of the essay will be two other loci where I take the modal status of laws to be at issue.

A further qualm I have with the term 'dispositional' is that it is often denied that dispositional properties are relational. A well-known presentation of this position is given by C. B. Martin in arguing that dispositions are not relational properties because they cannot be analyzed as conditionals.⁶ An object *x*'s property of breaking if dropped over a hard surface does not relate *x* to any other specified particular, but relates it to other unspecified particulars and so counts as a relational property. According to Martin, fragility cannot be analyzed as such a relational property. I take his arguments to show that nonfundamental dispositions, such as being electrically live, or being fragile, are not so analyzable.

But things look different when it comes to fundamental properties such as mass.⁷ In classical particle mechanics a particle *x* has determinate gravitational mass m iff $\forall y \forall m^*(y \text{ has gravitational mass } m^* \rightarrow x \text{ and } y \text{ attract}$ each other with a force of Gmm*/d²). On this account, the property of having a determinate gravitational mass m appears on the left-hand side of the biconditional, and on the right-hand side of the biconditional the property of having a gravitational mass of m* appears too. This is a different determinate mass property, so the account does not constitute a selfcontained analysis of having determinate mass m. But that does not make it uninformative or inadequate. Rather, it has the merit of systematically linking the different determinate mass properties. Particle *x*'s property of having gravitational mass m does not relate *x* to any other specified particular, but relates it to other unspecified particulars, so counts as a relational property. Thus fundamental dispositional properties can be relational and so can be included as (temporally one-directional) cases of what I am calling "fundamental role properties."

Inertial mass is a tendency of a particle to resist being accelerated by a force. More formally, x has determinate inertial mass m iff x has a tendency when acted upon by force F to move with an acceleration of F/m. This is also a relational property of x, relating it to unspecified external forces. We can see from the definitions of these role properties that inertial mass does not entail gravitational mass. But gravitational mass does entail inertial mass if, as I shall assume, force is defined in terms of Newton's first and second laws.⁸

Let us begin now to tackle the question of the status of fundamental laws, taking PMG as example. I shall begin by supposing the view, recently designated ontic structural realism,⁹ that fundamental properties have no intrinsic natures, i.e., have no intrinsic properties as components, but are given purely in terms of roles. I do not know of any discussions of the status of fundamental laws by ontic structural realists. But for reasons that will become clear. I find it instructive to consider first how the discussion would go under this supposition. If 'mass' is construed here as referring to gravitational mass, then the statement of PMG expresses a tautological, necessary, a priori truth. This should not be dismissed as the daft suggestion that scientists need do no empirical investigation to determine the laws of nature they use in explaining natural occurrences. Rather, it can be seen as suggesting that all the empirical work goes into establishing which properties are instantiated in the world, and it is the distribution of these properties at a certain time that is used to explain the distribution of these properties at other times. The fact that the statement of PMG expresses a tautology nicely captures the sense in which laws have no work to do once such role properties are instantiated. Accordingly, necessitarians are inclined to hold that properties have a primacy over laws.¹⁰ On such a necessitarian view it would be said that "there would be no mass if the gravitational constant were different." This sounds odd.

A more linguistically natural view is to take 'mass' in PMG as referring to inertial mass. Using R_i for the inertial mass role and R_g for the gravitational mass role, we can now construe PMG as asserting a relation between determinables $R_i \rightarrow R_g$. This is effectively a family of laws. It is the fact that for all determinate mass roles, anything with determinate inertial mass role R_{im} must have determinate gravitational mass role R_{gm} . This assimilates one example of what Armstrong calls functional laws¹¹ to the familiar form of laws asserting relations between properties. The possibility that the statement of PMG could express a necessary a posteriori truth has no Kripkean understanding when there is no underlying nature for the terms to be designating.¹² The necessity of such laws could then only be sui generis. On this view, it would be said that "a world in which massive particles do not satisfy an inverse square law of attraction is conceivable but not metaphysically possible." The counterfactual intuition expressed here is held by very few. And to maintain such a sui generis, a posteriori necessity, would mean abandoning the hope of assimilating the necessity of natural laws to some well-understood kind of necessity.

Far more plausible, if 'mass' refers to inertial mass, is the option that PMG, construed as $R_i \rightarrow R_{g'}$ is contingent a posteriori. On this view, one would say that "massive particles could have attracted each other with a different force, but couldn't have failed to respond to forces." This expresses a genuinely different and far more plausible counterfactual intuition than the necessary a posteriori option.

But could there be a genuine dispute between someone who takes PMG to be necessary a priori and someone who takes it to be contingent a posteriori? The apparent difference in status of PMG arises entirely from a difference in how 'mass' is taken in the statement of PMG. It seems linguistically more natural to take it as referring to inertial mass in the statements of law and counterfactual intuitions. But both sides would agree on what counterfactual intuitions hold and what status the statement of PMG has on each reading of 'mass'.

Disagreement may be thought to arise as to which reading captures the true metaphysics of natural laws. However, I believe that when each position is uncontroversially supplemented, they can be shown to be equivalent. By adding the modest further assumption that R_i is instantiated to the contingentist claim that $R_i \rightarrow R_g$, the necessitarian claim that R_g is instantiated follows. Conversely, the necessitarian view cannot be successfully captured in the mere fact that R_g is instantiated, without also adding that there are no instances of inertial mass that aren't also instances of gravitational mass. To give a complete dynamics of mass, the necessitarian must assert $\exists x(R_g x)$ and $\neg \exists x(R_i x \text{ and } \neg R_g x)$. This is formally equivalent to $\exists x(R_g x)$ and $R_i \supset R_g$. Since $\exists x(R_i x)$ follows from $\exists x(R_g x)$, the supplemented necessitarian position entails the supplemented contingentist position. Thus we have yet to identify a substantive disagreement between necessitarian and contingentist.

Necessitarians might now object that their assertion of which fundamental properties are instantiated identifies a different and superior position because contingentists need a way to distinguish the nomic necessity of $R_i \rightarrow R_g$ from the merely material conditional fact $R_i \supset R_g.$ I think contingentists should respond here that the fundamental laws are just the material conditional facts relating fundamental properties, and that nomically necessary truths are those that hold in all worlds with those fundamental laws. This way of demarcating the nomically necessary truths from the larger class of metaphysically contingent truths is equivalent to a necessitarian demarcation that takes nomically necessary truths to be those holding in all worlds in which all and only the fundamental properties of the actual world are instantiated. There need be no worry that this way of defining nomic necessity fails to distinguish the contingentist non-Humean from the Humean, since Humeanism cannot be formulated under the supposition that there are only role properties. Under this supposition, I conclude that there is no genuine difference between necessitarian and contingentist (setting aside the far less plausible necessary a posteriori position). Contingentists may claim that their way of formulating the position is more linguistically natural, but this is not an argument that their position itself is different and superior to the necessitarian position.

Ш

Now let's suppose ontic structural realism false, i.e., that fundamental properties cannot all be role properties and hence cannot all be purely relational. I favor this view but shall not attempt to argue for it here. A highly plausible assumption is that role properties supervene on intrinsic natures, that is, that there can be no difference in distribution of role properties without an underlying difference in distribution of intrinsic natures. This is analogous to the highly plausible supervenience of phenomenal properties on physical properties. In both cases, this supervenience entails that if some region or particular manifests some supervening property then it must have some subvening property that guarantees the presence of the supervening property.¹³ And in both cases, we have two questions to settle-the status of the supervenience laws from subvening to supervening properties, and whether such laws are one-one or manyone. So now we have a more complex set of issues to discuss-the status of both the fundamental laws of nature and the fundamental supervenience laws. We need fundamental supervenience laws revealing what roles are played by things with a given intrinsic nature; we also need laws revealing how the distribution of intrinsic natures in space changes over time, and laws revealing how the distribution of role properties in space changes over time, one or both of which sets of laws must be fundamental.

This conceptual framework seems plausible and accommodates positions in which some of this complexity collapses. It may prove to be flawed, but I shall adopt it for the remainder of this essay. Given that I am suggesting three different places where the question of modal status arises in discussing fundamental laws, the blanket terms 'necessitarian' and 'contingentist' are too blunt to capture all the potential hybrid positions. Nevertheless, I shall continue to use them, offering more precise characterizations when necessary.

Consider first the supervenience laws from intrinsic natures to role properties. They state that if some particular, spatial point, or region manifests some fundamental role property, then it must have some intrinsic nature that guarantees the presence of the role property. The question to be addressed now is this: In what way are role properties guaranteed by (i.e., supervene on) intrinsic natures? Is the link metaphysically necessary or contingent? A major obstacle to answering this question stems from a couple of ways in which intrinsic natures are disanalogous to phenomenal properties. First, there is no special class of concepts for intrinsic natures which is analogous to the class of phenomenal concepts. And second, we have no direct access to intrinsic natures, so that any terms we use to designate them must have their reference fixed by role properties. So 'mass' can now also be interpreted as rigidly designating an intrinsic nature for inertial mass N_i or an intrinsic nature for gravitational mass N_g, where this reference is fixed by a role property R_i or R_g. Our question then becomes: What is the status of the supervenience conditionals $N_i \rightarrow R_i$ and $N_g \rightarrow R_g$? (This question also determines whether the reference of a rigid designator 'N' is fixed by a contingent feature of the nature or a necessary one.)

One possibility is that the supervenience is metaphysical. That is, in all possible worlds in which a spatial point, region, or some other particular has a given fundamental intrinsic nature, it has all the same role properties. It may be that there are intrinsic-nature concepts which we can never grasp from which the link to role properties would be a priori. So the status of necessary a priori is a candidate for the supervenience of role properties on intrinsic natures, but as it is one that is opaque to us, I will focus on the choice between the other two candidates—necessary a posteriori and contingent a posteriori.

Metaphysically necessary supervenience laws taking physical to phenomenal properties have been thought mysterious if they are not a priori and cannot be understood along the lines of Kripkean a posteriori necessities, and a similar mysteriousness might be thought to attach to necessary a posteriori laws taking intrinsic natures to role properties. The view that intrinsic natures are identical to role properties is usually invoked to dissolve this mystery.¹⁴ This is analogous to the claim that phenomenal properties are identical to physical properties. A principal motivating thought in each case is that a posteriori metaphysically necessary supervenience links are puzzling, but a posteriori metaphysically necessary identity relations are not. This strikes me as smoothing a bump in the carpet only to create another one in a different place. In both cases we are left with the alleged identity of two properties that are on their face radically different, and no way of understanding how such different-looking properties could actually be one and the same. Or to put the problem as it is more commonly presented: we are told that one and the same property can be picked out by two radically different concepts, but are offered no clue as to how such different concepts could actually be picking out the same property. So I don't think identity helps dispel the mystery of a posteriori metaphysically necessary supervenience links. Armstrong puts it this way:

I confess that I find this totally incredible. If anything is a category mistake, it is a category mistake to identify a quality—a categorical property—and a power, essentially something that points to a certain effect. They are just different, that's all. An identity here seems like identifying a raven with a writing desk. (Armstrong 2005, 315)

In the case of psychophysical identity, the position has a further burden. Our epistemic access to both phenomenal and physical properties presents us with a seeming many-one relation between physical and phenomenal properties which has to be shown to be illusory. In the case of the identity of intrinsic natures with role properties, we lack direct access to intrinsic natures and so are in no position to say whether the relation seems to be one-one or many-one. The identity theorist doesn't therefore have to contend with the apparent existence of many-one relations, but still needs an argument as to why they are to be ruled out. The supervenience of role properties on intrinsic natures doesn't of course rule out different natures within a world guaranteeing the same role property. To accommodate the possibility that many intrinsic natures guarantee a given role property, property terms like 'mass', when taken to designate intrinsic natures by means of a reference-fixing role property, can be understood as designating an intrinsic nature or natures. Alternatively, they could be regarded as designating the unique intrinsic nature consisting of a heterogeneous disjunction of natures.¹⁵ This latter option would allow a Pyrrhic victory to the identity thesis but would not help its claim to be mystery-removing. The upshot is that metaphysically necessary supervenience links are mysterious.

An alternative possibility for the status of the supervenience link, commonly referred to as quidditism, is the view that intrinsic natures metaphysically guarantee no role properties at all, so that the supervenience of role properties on intrinsic natures is contingent. On this view, "mass could have played the role of charge" is true since whatever intrinsic nature or natures contingently guarantee the gravitational-mass role property in our world could have contingently guaranteed the charge role property in another world. This statement is thought by many to be unacceptable,¹⁶ and indeed, it would be incoherent if we took the terms to pick out role properties. But it isn't obviously amiss if we allow terms to designate intrinsic natures.

The view that this supervenience is contingent, as I have defined it, actually covers a spectrum of possibilities. For it might be that intrinsic natures metaphysically guarantee some but not all role properties. For example, it might be that a certain mass-nature metaphysically guarantees that things possessing this nature will attract each other with a force proportional to their masses and inversely proportional to the distance between them, but does not guarantee how strong the gravitational constant is. If such is the case, it would be true that "mass could not have played the charge role, but massive particles could have attracted each other with a weaker force." This statement may seem to reflect a more feasible contingent status for supervenience laws. I don't think much importance should be attached to such counterfactual intuitions compared with considerations of mystery and parsimony in assessing the overall merits of a view. But it seems to me that both parts of this modal intuition are equally strong. So those who advocate a metaphysically necessary supervenience link and thereby fail to capture the second part of this intuition cannot consistently complain that quidditists fail to capture the first part. The upshot is that quidditism has a counterintuitive air but lacks the mystery of metaphysically necessary supervenience links.

When supervenience laws are construed as contingent they are commonly referred to as "nomically necessary." This seems perfectly acceptable. No stronger notion of necessity is required than that which is associated with material conditional facts and shared by Humeans. When the discussion is restricted to fundamental supervenience laws, there is no danger of failing to distinguish merely accidental regularities from regularities regarded as laws. In the case of fundamental, psychophysical supervenience laws, the consequents of the conditionals are fundamental phenomenal properties and the antecedents are (perhaps complex distributions of) fundamental physical properties. In the case of fundamental supervenience laws between natures and role properties, the antecedents of the conditionals are fundamental intrinsic natures and the consequents are (perhaps complex arrangements of) fundamental role properties.

Ш

Now let's return to the status of the fundamental laws of nature. The conclusions drawn earlier—that we have a contingent relation $R_i \supset R_g$ between determinable inertial-mass role and determinable gravitational-mass role, as well as a conceptually necessary relation $R_g \rightarrow R_i$ —were established on the assumption that there are no intrinsic natures. I see no reason why they should not still hold if we assume there is an underlying level of intrinsic natures. However, that new level introduces new lawlike relations such as the supervenience laws we have just been considering and dynamic laws relating intrinsic natures. Let us consider now whether all three types of fundamental laws are irreducible, or whether one type is reducible to the others.

If there are many-one relations from natures to role properties then these supervenience laws together with dynamic laws relating role propeties do not tell us how distributions of intrinsic natures change over time. That is because if distinct natures N_1 and N_2 guarantee the same role property R, then there is no specification of how spatial regions containing N_1 or N_2 evolve over time. All kinds of bizarre patterns of alternation between N_1 and N_2 would be consistent with the dynamic laws relating roles. So we would need dynamic laws relating intrinsic natures.

Taking VN_i to be the disjunction of all the natures guaranteeing R_i and VN_g to be the disjunction of all the natures guaranteeing R_g, we get two-way relations between natures and role properties VN_i \leftrightarrow R_i and VN_g \leftrightarrow R_g. Together with the dynamic law R_i \supset R_g this yields VN_i \supset VN_g, which constrains the dynamic laws relating intrinsic natures but, as we have seen, does not entail them. It would be neat if these dynamic laws were simply relations between intrinsic natures N_{i1} \supset N_{g1}, N_{i2} \supset N_{g2}, etc., but these laws would still not rule out the troublesome alternations between natures. It looks as though a dynamic law concerning intrinsic natures must be interpreted as an irreducibly functional relation among natures.

But starting with these dynamic laws we would be able to derive $VN_i \supset VN_g$, which together with the relations between natures and role properties would yield the dynamic law $R_i \supset R_g$. Thus we can take fundamental laws relating role properties to be reducible to fundamental laws

relating natures and fundamental supervenience laws from natures to role properties.

One option then for a package of fundamental laws would consist of many-one contingent $N \rightarrow R$ laws and contingent dynamic laws relating natures. Since the Humean shares the view of intrinsic natures contingently related to role properties, we do not yet have a distinctly non-Humean contingentist package. That is because the nomic necessity here is no stronger than mere holding in all possible worlds with the material conditional facts that relate fundamental properties of the actual world, which can also be offered by a Humean as an ersatz account of nomic necessity.

A second option would consist of one-one contingent $N\leftrightarrow R$ laws, and fundamental laws relating natures, which this time are fully captured by $N_i\supset N_g$. This option similarly invokes a nomic necessity that is available to a Humean.

If we assume that natures are identical to role properties, then the only fundamental law is $N_i \supseteq N_g$. This third option has the definite advantage of ontological sleekness. Necessitarians assuming the identity of natures with role properties may be inclined to take as a primary statement of their view the instantiation of R_g , now expressible also as N_g . But this still needs to be supplemented with the claim that there is nothing instantiating N_i that doesn't also instantiate N_g , which can be recast as the claim that $N_i \supseteq N_g$, equivalent to $R_i \supseteq R_g$. So there is a similar verbal component here to the dispute between contingentists and necessitarians about the status of the fundamental dynamic laws as we encountered when examining role properties alone. What might have been thought of as a necessitarian package of fundamental laws and properties thus turns out to consist of contingent dynamic laws and metaphysically necessary identities N = R that are not an ontological addition.

The view that role properties metaphysically supervene on intrinsic natures without being identical to them provides the fourth option for a package of fundamental laws. And here the dynamic law consisting of an irreducibly functional relation among natures is all that's needed. As all facts involving role properties would metaphysically supervene on facts about natures, there would be no need for an irreducible category of fundamental role properties and thus no need for an irreducible category of fundamental supervenience laws. So this view is also ontologically sleek.

On the third and fourth options, the only fundamental laws are contingent relations among natures that have a weak nomic necessity consisting just in holding in all possible worlds with the material conditional facts that relate fundamental properties of the actual world. These options are distinctly non-Humean since we are currently assuming metaphysically necessary links between natures and role properties, which a Humean rejects.

I have argued that if there are no intrinsic natures, then necessitarian and contingentist views of fundamental laws turn out to be equivalent when plausibly supplemented. If there are intrinsic natures, then in addition to fundamental dynamic laws of nature there are also fundamental supervenience laws from intrinsic natures to roles. These supervenience laws may be many-one or one-one; they may be metaphysically necessary or contingent. Worries about mystery favor contingency; worries about complexity favor necessity. As I cannot see a plausible, intelligible way of avoiding contingent supervenience laws $R \rightarrow N$, I favor a wholly contingentist package of fundamental laws.

No way has yet been offered to distinguish a purely contingentist package of fundamental laws from Humeanism. At this point I return to consider the modal status of the role properties R_i and R_g. Both these properties involve a tendency to respond to forces. So having such a role property will amount to having some conditional property of being such that if it is A then it is B. Clear sense can be made of such a conditional $A \rightarrow B$ being construed as purely material. But with this interpretation the contingentist packages would have nothing to distinguish them from a Humean view. Clear sense can also be made of $A \rightarrow B$ being construed as holding in all possible worlds in which the antecedent A applies, and this is how dispositional essentialists would construe dispositional properties. What I've been calling a contingentist package will in this case contain an important necessitarian component. I find this an unattractive position because I suspect that the difficulties encountered for the view that $N \rightarrow R$ holds with metaphysical necessity will transfer to the view that $N\&A \rightarrow B$ can hold with metaphysical necessity.

Non-Humean contingentists will typically claim that such conditionals $A \rightarrow B$ can hold with a nomic necessity that is stronger than a Humean ersatz account of nomic necessity. And here they have been widely criticized for invoking a mysterious and ontologically extravagant fundamental notion of nomic necessity.¹⁷ To the charge of mystery I think there is an adequate response that cannot be given on all contingentist non-Humean views, which it is beyond the scope of this paper to discuss, namely that we can make good sense of a primitive quasi-causal necessitation relation among laws and particular states of affairs when a primitive

direction of time is built into the conception of fundamental law.¹⁸ But it must be conceded that this notion of nomic necessity is indeed an onto-logically distinct species of necessity.

One important consequence of this debate is the viability of the Humean view of laws as supervenient on the distribution of fundamental properties in spacetime. It would appear that this requires a quidditistic view of fundamental properties, since on all other views laws are already entailed by the instantiation of the fundamental properties prior to examining their distribution in spacetime.¹⁹ For the Humean wants to deny that there are any such necessities in nature. Thus Humeanism requires quidditism (which requires that there be intrinsic natures in addition to role properties). In finding no decisive considerations against quidditism I have therefore found no appreciable obstacle to Humeanism based on concerns about properties. Its strengths and weaknesses are to be judged on the basis of other criteria.

Acknowledgments

Special thanks to Tony Dardis, Travis Dumsday, Mark Migotti, and Jonathan Schaffer for their comments, and to Russ Payne and Robert Rupert for the written comments they delivered at presentations of early drafts of this essay.

Notes

1. A prominent exponent is Shoemaker. See Shoemaker 1980; see also Swoyer 1982; Bigelow, Ellis, and Lierse 2004. Necessitarians should also include dispositionalists who are inclined to be eliminativists about laws, such as Martin in Armstrong, Martin, and Place 1996; Heil 2003; and Mumford 2004.

2. This position is advocated most famously by Dretske (1977), Tooley (1977), and Armstrong (1983).

3. Lewis decries such "emergent natural properties of more-than-point-sized things" (1980, x).

- 4. See, for example, Chalmers 1996, 153-155; Stoljar 2001.
- 5. See Martin 2007, chap. 6.
- 6. See Martin 1994 or Martin 2007, chap. 2.

7. For an object's having a mass is not a disposition that is waiting for a partner in order for the conditional to apply, thereby allowing for the possibility that there

could be a fink that does the job of attracting another massive body when such a partner should arise. There is never a moment when an object lacks massive disposition partners, so there is never a moment when a massless object could satisfy the conditional analysis solely because of the operation of a fink. If the fink is necessarily always involved in arranging the truth of the conditional clause, then we should grant that this is what having the dispositional property of being massive consists in.

8. If there are also pre-Newtonian folk concepts of force then on such concepts force could equal some alternative to inertial mass times acceleration, such as inertial mass times velocity, so that gravitational mass on this understanding of force would not entail inertial mass. But then there would be a corresponding pre-Newtonian folk concept of inertial mass such that gravitational mass would entail this folk inertial mass, and my discussion would apply mutatis mutandis replacing inertial mass by folk inertial mass. So I will proceed without worry on the understanding that force and inertial mass presuppose Newton's first and second laws.

9. For an extended defense see Ladyman et al. 2007.

10. For example, Bigelow, Ellis, and Lierse (2004, 147) argue that it is more plausible to take the direction of explanation to go from the dispositional properties of things to laws. Armstrong, Martin, and Place (1996, 65) take claims about laws to be parasitical upon claims about dispositions of particulars.

11. See, e.g., Armstrong 1983, chap. 7.

12. Later I will offer a classical Kripkean necessary a posteriori reading of the statement of PMG in which 'mass' rigidly designates an intrinsic nature. One could construct another reading of the statement of PMG loosely analogous to an a posteriori necessity on the Kripkean model by construing 'mass' as rigidly designating a mass role property, so that it means something like that mass role property player in all possible worlds which plays the mass role in the actual world. I think it very unlikely that anyone would adopt such a reading so I do not pursue it.

13. In this essay I set aside the property counterpart theorist's denial that a property can be instantiated in more than one world, and hence that there can be any context-invariant truth about how properties correlate across worlds.

14. This resembles the view that dispositional properties or powers are identical to intrinsic properties advocated, e.g., by Martin (1997), Martin (2007, 81), and Heil (2003, chap. 11).

15. These alternatives would serve as a response to anyone who took the possibility that reference-fixing descriptions such as 'the nature guaranteeing the electron role' would fail to refer as a reductio argument against many-one relations between natures and roles. An argument resembling this is given by Bird (2007, 77–78).

16. See, e.g., Mumford 2004, 104; Bird 2007, 70-81.

17. See, e.g., Bigelow, Ellis, and Lierse 2004, 147–148, and the famous criticism in Lewis 1986, xii.

18. An account of fundamental laws along these lines is developed by Tim Maudlin (2007).

19. Jonathan Schaffer has suggested to me that this could be rejected on property counterpart theory, which holds that there is a distribution of natures to be found at each world, summarized by the Humean laws, but denies that there is ever identity of nature between properties at different worlds. This may be the view held by Robert Black (2000) in arguing that Humean supervenience can still be formulated without quidditism "as the thesis that the laws of nature supervene on the pattern of instantiation of qualities in spacetime."

References

Armstrong, D. M. 1983. *What Is a Law of Nature?* Cambridge: Cambridge University Press.

Armstrong, D. M. 2005. Four disputes about properties. Synthese 144:309–320.

Armstrong, D. M., C. B. Martin, and U. T. Place. 1996. *Dispositions: A Debate*. London: Routledge.

Bigelow, J., B. Ellis, and C. Lierse. 2004. The world as one of a kind. In *Readings in the Laws of Nature*, ed. J. Carroll. Pittsburgh: University of Pittsburgh Press.

Bird, A. 2007. Nature's Metaphysics. Oxford: Oxford University Press.

Black, R. 2000. Against quidditism. Australasian Journal of Philosophy 78:87-104.

Chalmers, D. 1996. The Conscious Mind. Oxford: Oxford University Press.

Dretske, F. 1977. Laws of nature. Philosophy of Science 44:248-268.

Heil, J. 2003. From an Ontological Point of View. Oxford: Oxford University Press.

Ladyman, J., D. Ross, D. Spurrett, and J. Collier. 2007. *Everything Must Go*. Oxford: Oxford University Press.

Lewis, D. 1986. Philosophical Papers, vol. II. Oxford: Oxford University Press.

Martin, C. B. 1994. Dispositionals and conditionals. Philosophical Quarterly 44:1-8.

Martin, C. B. 1997. On the need for properties: The road to Pythagoreanism and back. *Synthese* 112:193–231.

Martin, C. B. 2007. Mind in Nature. Oxford: Oxford University Press.

Maudlin, T. 2007. The Metaphysics within Physics. Oxford: Oxford University Press.

Mumford, S. 2004. Laws in Nature. Oxford: Oxford University Press.

Shoemaker, S. 1980. Causality and properties. In *Time and Cause*, ed. P. van Inwagen. Dordrecht: Reidel. Reprinted in *Properties*, ed. D. H. Mellor and A. Oliver. Oxford: Oxford University Press, 1997.

Stoljar, D. 2001. Two conceptions of the physical. *Philosophy and Phenomenological Research* 62:253–281.

Swoyer, C. 1982. The nature of natural laws. *Australasian Journal of Philosophy* 60:203–223.

Tooley, M. 1977. The nature of laws. Canadian Journal of Philosophy 7:667-698.

6 Para-Natural Kinds

Roy Sorensen

A para-natural kind is an absence defined by a natural kind. For instance, cold is defined as the absence of heat and shadow as an absence of light.

There is a folk taxonomy for para-natural kinds. Scientists refine this taxonomy by improving the alignment of our commonsense categories with the corresponding para-natural kinds. After William Hershel discovered infrared light, physicists acquiesced to his talk of invisible shadows (despite the locution echoing the oxymoronic "invisible light"). When William Ritter went on to discover ultraviolet light, physicists continued the process of extending 'shadow' across the electromagnetic spectrum.

Just as scientists discover new natural kinds such as the platypus and the electron, scientists discover new para-natural kinds such as biological niches (Smith and Varzi 1999).

Elsewhere I have argued that scientists have overlooked para-reflections and para-refractions (Sorensen 2003): picture a white beach ball and a black beach ball resting on the surface of a placid pond, as shown in figure 6.1.

The sun is shining from behind you and to your left. The white ball has a reflection because light is bouncing off the ball and onto the flat water. What about the black ball? Since the black patch is due to the light that does *not* reflect off the ball, that patch cannot be a reflection of the ball. The distinction between para-reflections and shadows is salient in cases of objects that have both simultaneously, as shown in figure 6.2.

The black ball has a para-reflection on the side near the sun and a shadow on its far side. If a hot duck wants to cool off, it will paddle to the ball's shadow, not to its para-reflection.

There is little pressure to correct the marginal confusions between reflections and para-reflections because the generalizations that govern the host work well for the parasite. For instance, the virtual image of a fish's shadow is displaced in accordance with Snell's law even though Snell's law concerns light rather than the absence of it.



Figure 6.1

A reflection and a para-reflection.



Figure 6.2 Simultaneous reflection and para-reflection.

Para-Natural Kinds

The positive transfer of learning from host to parasite is copious—and symmetrical. Students often learn about the host by studying its parasites.

Absences are not substances and so are not natural kinds. But they have so many other features of natural kinds (lawfulness, mind-independence, projectibility) that they are frequently mistaken as natural kinds.

Since money is not a natural kind, bankruptcy is not a para-natural kind. Nor are voids that form *fortuitously*. Consider the hidden figure in *The Tomb and Shade of Washington* (figure 6.3). The "silhouette" is depicted as forming coincidentally; the figure does not inherit the structural integrity of the trees and plants which frame it.



Figure 6.3 James Merritt Ives, The Tomb and Shade of Washington.

Para-natural kinds also differ from *substances* that get defined in opposition to other natural kinds. Consider pure samples. To borrow an example from Sarah Sawyer (2003), corundum is a colorless mineral with the pure chemical constitution Al_2O_3 . If mixed with a little chromium (less than 1 percent is enough), the mineral is a ruby. (The chromium makes it red.) If corundum is adulterated with titanium and iron, the mineral is a sapphire. (These trace contaminants make it blue.) If absences never mattered for kind individuation, mineralogists would not be able to distinguish pure corundum from rubies and sapphires.

Extending Plato's Metaphor

Plato compared definers to butchers (*Phaedrus* 265d–266a). The amateur hacks the carcass, the expert carves at its joints. Similarly, a good definition reflects pre-existing divisions.

In the segmental sense, a joint is a body part joined to another body part. In the privational sense, a joint is the gap between the two segments. This opening promotes actions such as bending, gliding, and rotation.

Natural kinds correspond to the segments. Para-natural kinds correspond to the openings.

The Greek atomists show early appreciation of the mechanical liberties conferred by para-natural kinds: a plenum makes for a static universe in which objects gridlock. By placing atoms in a void, the atomists could explain the locomotion of a cart, the absorption of a sponge, the filtering of a sieve, and the "suction" of a siphon.

Contrast with Artifacts

Natural kinds owe their features to their internal nature. They have characteristic origins, and characteristic patterns of change (such as the phase structure of water). This regularity is underwritten by a rich network of causal interactions that are kept stable. Similarly, shadows wax and wane systematically. Sundials exploit this regularity.

Artifacts contrast with natural kinds. Clocks are alike in that their purpose is to measure time, but this end can be achieved by diverse physical principles. Thus there are sundials, water-clocks, hourglasses, pendulum clocks, and digital clocks. Artifacts morph easily into other artifacts such as when a door becomes a table. This flexibility makes sense given that artifacts are so dependent on intentions.

Para-natural kinds depend on natural kinds but otherwise share the contrasts with artificial absences. For instance, origins are important: your shadow cannot have been cast by someone else.

In some respects, shadows meet hopes about natural kinds better than any natural kind. For instance, shadows are homoromerous: any part of a shadow is a shadow. Consequently, a shadow can be made of nothing but shadow (any light in a shadow is pollution). So each part of a shadow is homogenous with respect to its intrinsic properties.

Shadows are so well behaved that the founders of projective geometry have precisely axiomatized their movements. Given all this lawfulness and predictability, shadows have been of considerable interest to scientists especially astronomers who can only examine objects telescopically.

Nomic Legacies

Para-natural kinds inherit the lawfulness and projectibility of the natural kinds that shape them. A black raven confirms "All ravens are black" but a white handkerchief does not (despite the logical equivalence of "All ravens are black" and "All non-black things are non-ravens"). W. V. Quine's explanation is that 'black' and 'raven' pick out natural kinds but 'non-black' and 'non-raven' do not.

I agree that 'raven' picks out a natural kind. I agree that 'non-raven' does not. But I think 'black' picks out a para-natural kind. ('Non-raven' is not a para-natural-kind term because it is merely the complement of 'raven'.) Black is an absence of hues arising from a relative absence of light flow (whereas white is an absence of hues arising from an indiscriminate flow) (Sorensen 2008, 217–218). Blackness can be achieved in a variety of ways just as blue can arise from disparate physical mechanisms.

Many would be reluctant to go along with Quine's assimilation of colors to natural kinds. Colors are secondary qualities, properties that arise from a combination of psychology and primary qualities. Whatever this does to the status of black as a para-natural kind, it does raise an interesting issue about secondary qualities. Black is the appropriate visual response to an absence of light-flow, so the physical component of the secondary quality is an absence rather than the presence of physical stimulus.

Human beings prefer to picture reality positively. Philosophers precisify this conviction in various ways, such as the linguistic claim that negative statements can be paraphrased as positive statements. This preference for the positive leads scientists to interpret absences as positive things. For instance, black is taken to be the color of maximally efficient light-absorbers. They insist that the blackness of a lightless cave is only darkness, not blackness. In unguarded moments they will describe sunspots as black; but once the lack of a light-absorber is noted, they will redescribe sunspots as merely dark.

The distinction between positive and negative is often relative to a scale. 'Sobriety' is a negative term insofar as it indicates absence of drunkenness, but it is a positive term with respect to character assessment.

Grammar is only an imperfect guide to the positive-negative distinction. Logicians have tried to put the distinction on a more principled basis by associating positiveness with specificity. In keeping with the philosophical tendency to associate the less real with the psychological, philosophers have also tried to characterize negative terms in terms of disbelief or frustrated expectations.

Anything less than a chapter-length summary of these efforts (some worthy descendent of Gale's 1976 review) would fail to do justice to the sophistication and insightfulness of these efforts. Yet in my opinion, these efforts are premature because a patient study of absences shows that they come in a rich assortment. People do not carefully sort items they plan to discard. More specifically, metaphysicians have not carefully classified the absences they wish to expel from their ontologies. Just as a physical anthropologist learns much by meticulously examining trash, a metaphysical anthropologist can learn much by studying what philosophers wish to dismiss as unreal.

Scientists agree that they sometimes mistake the negative for the positive. Benjamin Franklin believed that electric current flowed from where it is to where it is not. His factual errors led to the entrenched misnomers about the positive and negative poles of batteries.

A *shadow* conforms to many of the laws governing its host, light. As a bonus, a shadow echoes the boundaries of its caster. Galileo exploited this fidelity when measuring the heights of lunar mountains.

But a shadow is an *absence* of light. The laws of light only indirectly apply to shadows. If shadows were directly under the jurisdiction of the law, they would be counterexamples to important corollaries about how things behave. During a lunar eclipse, the earth's shadow appears on the far side of the moon without penetrating the moon. Shadows have no mass or momentum. They move, but not because one stage of the object causes the next stage of the object. Yet shadows propagate at speeds that precisely relate to the speed of light in the medium in question. The proper path to understanding a shadow is always a *detour* off the main road of physics.

As is evident from the controversial career of the vacuum, Aristotelian philosophers and many hardheaded scientists have denied that paranatural kinds are appropriate objects of study. After all, absences are not even objects!

However, para-natural kinds inevitably come along for the ride once we accept a metaphysics of natural kinds. Just as citizens underestimate the contribution of illegal immigrants to their economy and overestimate the ease with which they can be expelled, theoreticians underestimate the services of absences—and their entrenchment.

Natural Kind or Para-Natural Kind?

Sometimes there is uncertainty about whether something is a natural kind or a para-natural kind. Early astronomers modeled stars as holes in the celestial sphere. A celestial fire burns on the other side of the sphere. As the sphere moves, the holes move in unison. So instead of postulating many light sources that manage to move while maintaining their relative positions, the hole-theorist postulates a single perforated sphere and a single source of light.

Sometimes scientists are confident that one member of a pair of terms denotes an absence but are unsure which is the positive term and which is the negative. Is heat the absence of cold or cold the absence of heat? Ernst Mach (1896/1986) believed that the question made no empirical difference and so regarded the issue as merely verbal—and perhaps a little comical—anticipating a bit of Gerald Grow's humor (figure 6.4).

Atomists eventually persuaded most physicists that cold is indeed the absence of heat. The issue was not verbal—and, contrary to Mach's positivism, it would not have been verbal even if there were no way to resolve the dilemma.

The intense chemical activity of chlorine can be modeled as the effect of an electron hole in its outer shell. The absence of this electron behaves like a positive charge because electrons are attracted to the gap.

Indeed, Paul Dirac modeled the proton as a hole in a sea of negativeenergy electrons. Such a hole would act like a positive energy electron with a reversed charge.

Hole-theorists offer you something for nothing. Instead of postulating a new entity, you postulate a gap in an entity to which you are antecedently committed. You solve the problem by using *less* of the stuff you already have!



Yesterday, I found out that I am a cartoon character, and I've spent all night trying to decide if I'm the lines or the spaces in between.

Figure 6.4 Cartoon by Gerald Grow.

In addition to this parsimony with substances, there is parsimony with respect to forces. As illustrated by the least-effort principle in mechanics, we prefer explanations that postulate fewer forces. We prefer explanations that appeal to omissions rather than commissions. Failures to prevent the natural course of events provide a simpler explanation than interventions.

Epistemologically, para-natural kinds have the same potential to solve confirmation puzzles, in particular, Carl Hempel's Raven paradox and Nelson Goodman's new riddle of induction. Psychologically, developmental studies should reveal para-natural-kind thinking in toddlers just as developmental psychologists have discovered natural-kind thinking in children almost as soon as they can be interrogated.

Kripke Tests

Para-natural kinds pass many of the tests we employ to distinguish natural kinds from nominal kinds. Saul Kripke emphasizes tests which demonstrate that beliefs about the object are not *necessary* for reference. Consider the ignorance test. If agnostic or misinformed speakers still manage to refer, then their knowledge about the kind cannot be responsible for reference. A more objective link, such as causation, is needed.

'Shadow' passes Kripke's ignorance test. Speakers make errors of omission and commission about the nature of shadows and yet still successfully refer to them.

Kripke's account of how natural-kind terms achieve reference extends smoothly to para-natural-kind terms. In lecture three of *Naming and Necessity*, Kripke illustrates his concept of a reference-fixing description with 'heat'. The sensation of heat was used to lock on to a natural kind, but 'heat' does not mean the sensation of heat. If we apply the reference-fixing account to 'sound' we get an interesting answer to the riddle: "If a tree falls in a forest and no one hears it, does it make a sound?" Physicists *dissolve* the riddle by distinguishing between two senses of 'sound', one for sound waves and the other for auditory sensations. But Kripke would not admit that there is a sense of 'sound' in which it means the auditory sensation. Instead, auditory sensations merely fix the reference of sound. So Kripke's answer to the riddle is that the tree would make a sound. This underscores the objectivity of sounds.

A parallel story applies for 'cold'. The sensation of cold was used to lock on to the para-natural-kind of cold (the absence of heat, or more specifically, thermal energy). The reference-fixing description can be positive without the corresponding kind being positive. Ice would be cold even if there were no one around to feel cold. Coldness is objective.

Natural kinds are still needed to differentiate absences from one another. Mice die from cold, not darkness or silence.

Putnam Tests

Hilary Putnam emphasizes tests that are designed to show that nominal essences are not *sufficient* to pick out natural kinds. If speakers with the same beliefs (individuated narrowly) pick out different things, then their *environments* must determine the meaning. Imagine a Twin Earth in which there is a smoky type of substance that projects from objects that block

light sources. To the prescientific eye, the dark gas-like projections behave just like shadows: they assume the shapes of their casters, defy gravity, and obey the laws of projective geometry. But the dark projections are actually substances rather than absences of light. The dark things people see would not be shadows even though the speakers on Twin Earth use a word that sounds just like 'shadow'.

What really determines the reference of 'shadow' are the samples of shadows used by founding speakers. The word 'shadow' is an effect of those shadows and it is the causal connection that fixes the reference, not the stereotypes and superstitions that guided early discourse about shadows.

The good performance of shadows on these Putnam tests and Kripke tests should be expected: para-natural kinds are "copying" the test results of their corresponding natural kinds!

Back-up Strategy

If an apparent natural kind does not seem to fulfill the requirements, then one should entertain the possibility that it is actually a para-natural kind.

Consider the issue of whether diseases are natural kinds. On the one hand, diseases have many of the marks of natural kinds:

The diverse systems of classification in the history of medicine proceed not, in most cases, from mere convenience or arbitrary choice, but from a first reliance on descriptive accounts worked out prior to some access to the underlying causal processes of these conditions. Once the underlying nature is discovered, ideally a specific virus or bacteria, the previous descriptive account is either reformed or adjusted to the now discovered underlying disease process. (D'Amico 1995, 557)

Patients defer to physicians to learn what diseases they have. We all engage in this linguistic division of labor when wielding terms such as 'arthritis'.

On the other hand, diseases appear to be deficiencies, departures from the norm that compromise survival and reproduction. These privations lack integrity. Pathogens can exist separately from the organisms they infect (and are themselves subject to diseases). But the disease itself depends on its sufferer: diseases depend on the organism in the manner moth holes depends on a moth-eaten shirt.

One way to reconcile this dependency with the resemblance to a natural kind is to characterize diseases as para-natural kinds. Given the analogy between diseases and irrationalities, fallacies and fads, there may be scope for para-natural kinds in the social sciences.

Second Order Para-Natural Kinds

First order para-natural kinds are absences defined by natural kinds. Second order para-natural kinds are absences defined by para-natural kinds. Nasal passages are para-natural kinds defined in terms of holes in the nose. Choanal atresia is an absence of these holes, in particular a congenital failure of one or both nasal passages to open. There is a whole family of congenital disorders that consists of absences of holes: esophageal atresia, anal atresia, aortic atresia, and so on. These holes have vital functions. People with an absence of these holes must have them surgically supplied.

Some types of fossils are second order para-natural kinds. Archeologists have found footprints with an important human characteristic: the absence of a gap between the big toe and the next toes. This second order absence is evidence of a stable longitudinal arch. This can be seen in the impressions of the heel and of the ball of the foot left by humans walking with a normal gait on soft ground.

Or consider the "waves" that pass through shadows. Although they obey nearly all the laws that govern normal waves, they cannot be real waves because they do not transfer energy. Their medium is a shadow. These para-waves are second order para-natural kinds.

There can be third order para-natural kinds, fourth order, and so on. This hierarchy may stimulate the traditional doubts about an unparsimonious proliferation of absences.

This worry may be tempered by the thought that if one is going to be unparsimonious, then be unparsimonious about absences! There is no *substantive* addition to your ontology. This bookkeeping loophole is analogous to that touted by believers in the principle of unrestricted composition. Consider the creative accounting of David Lewis (1991, 81). Yes, Lewis counts any combination of objects as an object. No, he has not added anything to his base ontology. Lewis starts with *n* cats. He applies the principle of unrestricted composition and gets $2^n - 1$ cat fusions. But not any more cats.

Lewis (2004) counts absences as causes. So his initial stock of cats must be supplemented with many absences of cats. But he does not have any more substances than he began with.

One limit on this precedent is that mereological fusions are extensional. Many absences may have a psychological aspect. In 1838 the explorer Charles Wilkes mapped the South Seas. Spreading out all five vessels from north to south, so that an estimated twenty miles of latitude could be continuously scanned on a clear day, they sailed over the coordinates of these "vigias," or doubtful shoals. Invariably they found no sign of any hazard, and Wilkes would later send a list of these phantom shoals to the secretary of the navy. As the Ex. Ex. was proving, exploration was much about discovering what did *not* exist as it was about finding something new. (Philbrick 2003, 77)

Absent shoals are creatures of expectation. These psychological absences are prone to the general problems of mental overpopulation. But paranatural kinds are objective absences. Shadows and craters are not mind-dependent.

Historical Controversy

Parmenides objected to absences on the grounds that non-being cannot exist. Atomists were accused of smuggling non-being back into reality with their claim that empty space is available for the movements of atoms. The history of science has involved various Parmenidean purges in which absences get pushed out the front door. The resulting absence of absences creates negative pressure which then draws them back in—usually through the back door.

There are important exceptions. With his edict of 1277, Pope John XXI escorted the void *in* through the front door. Although nominally aimed at the Arab infidel Averroes, the real target was known to be the excesses of the Aristotelians. Exclusion of the vacuum was forbidden (to make room for the Christian doctrine of creation from nothing), and this promoted interesting speculations about the behavior of objects in vacuums. Can they be seen? Might they move less than infinitely fast?

Vacuum research became empirical in the seventeenth century. The devout Blaise Pascal defied René Descartes's plenism with explicit barometric arguments for a vacuum (leading Descartes to confide to Christian Huygens that Pascal had too much vacuum in his head).

Absences are traditionally viewed as a special problem for materialism. The Lewis dialogue on "Holes" features a materialist trying to parry objections based on holes. Roberto Casati's and Achille Varzi's *Holes and Other Superficialities* (1994) characterize holes as immaterial beings.

But absences are as much a difficulty for the idealist as the materialist. The mental realm has vanishing points, objectless emotions, lexical gaps, and other embarrassing whatnots (Sorensen 2007). Absences, like existence and time, pose a *generic* ontological challenge.

Some Parting Imagery

A patriot about natural kinds would object to an outsider's advocacy of aboriginal rights for absences. Their sentiments parallel those voiced by Prime Minister Paul Keating immediately preceding the Australian Bicentenary in 1988 (figure 6.5): "I do not believe that the symbols and the expression of the full sovereignty of Australian nationhood can ever be complete while we have a flag with the flag of another country on the corner of it." What is the Union Jack of absences doing on the masthead of realism?

The seven points of the star below the Union Flag represent each of the six original states of Australia plus a star for the Commonwealth's internal and external territories. The right half depicts the Southern Cross, the brightest constellation visible from the Southern Hemisphere. Although the Southern Cross is large in spatial extent, it is composed of only five stars or perhaps only four if you count only the four prominent stars needed to form the Latin cross. Some Australians claim the Southern Cross is the smallest constellation.

But Northern Hemispherians might claim smaller constellations. The three brightest stars of Triangulum form an isosceles triangle. Canis Minor has only two stars, Procyon and Gameisa. (Canis Minor is one of Orion's hunting dogs.)

A constellation is an extended pattern of stars. This rules out single-star constellations: a single point cannot constitute a shape. However, the Australians could still claim to have the smallest constellation, and, simultaneously, the biggest. Next to the Southern Cross is a dark shadow called the Coal Sack. This is the head of the Emu, a negative constellation. It is



Figure 6.5 The Australian flag.

an immense expanse of blackness between stars and the bands of the Milky Way caused by the dust and gas clouds of space.

Ancient aborigines knew that the male emu sits on his mate's eggs. The bird symbolizes the role of Aboriginal Elders in the initiation of boys into manhood. The constellation may have also played a role in Aboriginal navigation and their calendar. In any case, the Emu is the largest and longest known constellation.

Some may resist calling the Emu a constellation. But bear in mind that constellations are not natural kinds. The stars only look to be closely situated and to have shapes from a terrestrial perspective. Constellations are visual mnemonics that help us remember the positions of randomly distributed stars. Negative constellations serve this mnemonic function just as well as positive constellations.

The Emu is not a para-natural kind. It is a fortuitous formation like the hidden figure in *The Tomb and Shade of Washington*. Even so, the Emu is a grand absence, worthy of note on a map of the night sky, and if it were a genuine para-natural kind the Emu would be of still greater interest.

Summary of Three Theses

My parity thesis has been that, by and large, what goes for natural kinds goes for para-natural kinds. So there is linkage; if you adopt a metaphysics of natural kinds, then you will also be committed to para-natural kinds.

Conventionalists may interpret this linkage as a new destructive premise: if the interesting features of natural kinds are also features of para-natural kinds, then natural kinds are no more real than shadows.

However, I am a believer in natural kinds. My more specialized work on shadows, para-reflections, and their ilk has been guided by natural-kind reasoning. It is no coincidence that the parasitical kinds have the same interesting features as the host. They owe their order to natural kinds, not the other way around.

The recommended attitude toward para-natural kinds is that they are a bonus, an unanticipated byproduct of a metaphysical system that has appealed to common sense and scientific thinkers since the era of Aristotle.

Acknowledgments

I thank Gerald Grow, at http://longleaf.net/ggrow, for permission to reprint his cartoon.

References

Casati, R., and A. Varzi. 1994. *Holes and Other Superficialities*. Cambridge, Mass.: MIT Press.

D'Amico, R. 1995. Is disease a natural kind? *Journal of Medicine and Philosophy* 20:551–569.

Gale, R. M. 1976. Negation and Non-Being. Oxford: Blackwell.

Lewis, D. 1991. Parts of Classes. Oxford: Blackwell.

Lewis, D. 2004. Void and object. In *Causation and Conditionals*, ed. J. Collins, N. Hall, and L. A. Paul. Cambridge, Mass.: MIT Press.

Lewis, D., and S. Lewis. 1970. Holes. Australasian Journal of Philosophy 48:206-212.

Mach, E. 1896/1986. *Principles of the Theory of Heat*. Trans. P. E. B. Jourdain. Dordrecht: Reidel.

Philbrick, N. 2003. The Sea of Glory: America's Voyage of Discovery, the U.S. Exploring Expedition, 1838–1842. New York: Viking Press.

Putnam, H. 1975. The meaning of 'meaning'. In *Minnesota Studies in the Philosophy of Science*, vol. VII, ed. K. Gunderson. Minneapolis: University of Minnesota Press.

Quine, W. V. O. 1969. Natural kinds. In *Ontological Reality and Other Essays*. New York: Columbia University Press.

Sawyer, S. 2003. Sufficient absences. Analysis 63:202-208.

Smith, B., and A. Varzi. 1999. The niche. Noûs 33:214-238.

Sorensen, R. 2003. Para-reflections. *British Journal for the Philosophy of Science* 54:93–101.

Sorensen, R. 2007. The vanishing point: Modeling the self as an absence. *Monist* 90:432–456.

Sorensen, R. 2008. *Seeing Dark Things: The Philosophy of Shadows*. New York: Oxford University Press.

7 Boundaries, Conventions, and Realism

Achille C. Varzi

1 Natural vs. Artificial Boundaries

If you have driven in Europe recently you may have had that strange feeling. You see a sign that says 'Deutschland' or 'France' or 'España'-and just drive through. No customs barrier, no passport control-just a sign. You say "Ah!" and carry on; the sign could be a hundred yards further out and it would make no difference. Yet by crossing that line you enter a different world-district, magically separated from its surroundings-you enter a region where people suddenly speak another language, rely on their own authorities, share a different heritage, and struggle to solve *their* problems and to improve the quality of their common life. The line is there, even if you don't see it. That sign conceals a long history, perhaps even a thread of blood, though all you see today is a spread of asphalt, souvenir shops, motels, gas stations, abandoned customs houses. It is more difficult to get that feeling as you drive across the United States. Most drivers feel nothing at all as they pass the border between Wyoming and Idaho, a line whose embarrassing geometric straightness says very little about its history (or says it all). Yet even here there are differences, and Idahoans are proud of their license plates just as Wyomingites are proud of theirs. Such is the magic of boundary lines: they are thin yet powerful; they separate, and thereby unite; they are invisible yet a lot depends on them, including one's sense of belonging to a country, a people, a place. They are abstract, in a way—and yet people take them seriously and some states expend huge sums of money and sacrifice soldiers' lives to protect or redraw them. (Kashmir is one example where the drawing of boundarieseven the precise drawing of the Line of Control—is still central to the conflict.)

Not all boundaries are so magical, though—are they? As I was flying over Yellowstone National Park I did not, in fact, see the Idaho-Wyoming boundary, just as you don't see the boundary between Germany and France when you fly over Europe. Nor did I see the boundaries of the Park itself, or those of the Missouri Plateau earlier on. But I did have a clear and distinct impression of seeing *other* boundaries: the shoreline of Lake Erie, for instance, or the edges of the Missouri River. I saw the boundaries of Long Island when my flight took off. And I think at some point I saw the crater of a volcano, probably Bear Butte, though I am not positive about that. (It may have just been a small lake.) In his celebrated Romanes Lecture of 1907, the British Viceroy of India, Lord Curzon of Kedleston, introduced an important distinction in this regard, a distinction that is so intuitive as to be part of common sense, and that geographers have embraced ever since. And it fits the bill. It's the distinction between *artificial* boundaries or frontiers, on the one hand, and *natural* boundaries or frontiers on the other (Curzon 1907).

The boundaries I didn't see during my flight would be of the first sort. National and state borders are artificial insofar as they are the product of human decisions and stipulations, an expression of collective intentionality that translates into political, social, and legal agreements whereby it is determined where a certain territory begins and where it ends. So, too, are the boundaries of many other geographic entities, such as plateaus or wetlands or areas of a given soil type, though these may be induced by cognitive or cultural processes, or by scientific stipulation, rather than by legal or political practices. Such artificial boundaries may be drawn with great accuracy (for example, a national border) or left somewhat vague, fuzzy, underspecified (for example, the boundaries of the Missouri Plateau); it depends on the importance we attribute to the relevant demarcations, on the role they play in our lives. But whether sharp or vague, they all qualify as "artificial" precisely because the demarcations are humaninduced: they need not correspond to any genuine physical or otherwise objective differentiations in the underlying territory. They are *de dicto*, so to speak, not *de re* boundaries.

Geographic boundaries of the second sort—the natural, or *de re* boundaries—would by contrast be characterized precisely by their apparent independence from our organizing activity. We can stipulate that one half of Lake Erie belongs to Canada and the rest to the United States, and that dividing line will be an artifact. But the shoreline—the border of the whole lake—does not seem to depend on us. It's there regardless, it exists "on its own." The same can be said for the boundaries of certain political or administrative entities, such as the Region of Sicily, whose limits are for the most part identified with the limits of the Sicilian *island*; or

such as Spain, which although connected to continental Europe, is separated from it by the admirably fashioned Pyrenees ("the most obvious of features," wrote Joseph Calmette, "the plainest of lines, designed by nature in her boldest manner" [1913, 1]). Artificial boundaries may be subject to controversy. They can be ignored or deleted, and thereby go out of existence; they can be drawn anew, and thereby come into being. Not so with natural boundaries. We are free to ignore natural boundaries for certain purposes, but we cannot ask a cartographer to omit them from a map of the world. In a physical map we may omit all political boundaries; but a political map will perforce include all physical boundaries—at least, physical boundaries that are visible at the relevant scale.

Now it is, of course, an open question whether Lord Curzon's intuitive distinction is well grounded. That is precisely the question I want to address. First, however, let me emphasize that this question does not only arise relative to the large-scale geographic world that we find depicted in ordinary maps and atlases. It also arises, for instance, in the smaller-scale world featured in a cadastre. Here too, the parceling of land into real estate is not simply a geometrical affair. In some cases it would seem to rely on natural, preexisting physical discontinuities such as creeks, rocks, cliffs, or ditches; in other cases, it is crucial that people believe that whoever fenced off a plot of land is the person who actually owns it. So collective intentionality appears to be necessary to explain the difference between landed property and raw land. And what goes for the cadastre goes for everything: boundaries play a central role at *any* level of representation or organization of the world, and so does the relevant artificial/natural distinction. We think of a boundary *every* time we think of an object as of something separated from or distinct within its surroundings. There is a boundary (artificial) separating my part of this shared desk from my colleague's, my head from the rest of my body, or the sirloin from the rump on a butcher's chart; there is a boundary (natural) demarcating the interior of an apple from its exterior, the hole from a donut, or shadow from light. Events also have boundaries, including temporal ones, and the distinction appears to apply equally well: the end of a war or my turning twenty-one would be examples of artificial boundaries; my birth and death or the point in the cooling process when water begins to solidify would be obvious candidates for natural boundaries. Even abstract entities, such as concepts or properties, may be said to have boundaries. Those expressed by disjunctive predicates such as 'emerose' and 'grue', or by phase sortals such as 'student' and 'jobless', would have *de dicto* boundaries. Those expressed by

substance sortals and so-called natural-kind terms, such as 'cow' or 'water', would have genuine, *de re* boundaries.

It is not an exaggeration to say that boundaries are at work in articulating every aspect of the reality with which we have to deal. They stand out in every map we draw of the world—not only the cartographic world but the world of nature at large, as well as the sociocultural world that emerges through the weaves of our social and individual practices. And this ubiquity of boundaries goes hand in hand with Curzon's artificial/natural distinction, the apparent contrast between merely *de dicto* and genuinely *de re* demarcations, the opposition—in Barry Smith's more recent terminology—between *fiat* articulations and *bona fide* joints of reality.¹ It is not, therefore, an exaggeration to say that our question bites deeply: how tenable is this distinction? And how does its answer affect our overall metaphysical picture of the world? How does it affect our understanding of the identity and survival conditions of the very things that boundaries demarcate?

2 Boundaries and Things

To begin with, the relationship between a boundary and the entity it bounds demands clarification. Brentano, following a tradition that goes back to Abelard if not to Aristotle,² held that the distinguishing feature of boundaries lies in their being ontologically dependent on the entities they bound: a boundary "can never exist except . . . as belonging to a continuum which possesses a larger number of dimensions" (1914/1976, 14). And it is true: there are, in reality, no isolated points, lines, or surfaces. We cannot eat all the interior parts of an apple and just keep the surface-not the *skin* (which is a bulky part) but the perfectly two-dimensional entity that circumscribes the skin from the outside. We cannot display the boundary of our country in a museum, or steal the point of intersection between the equator and the Greenwich meridian. Not even God could work such marvels, pace Suarez (1856-1866, XL, 5, 41). However, this relation of dependency is symmetrical: it is equally impossible to have an apple without a surface, or a nation without a border (with few exceptions, such as Poland during the Era of Partition). Indeed, when it comes to de dicto boundaries, it would seem that the latter sort of dependency is especially strong, precisely insofar as those boundaries emerge from our social or cognitive *fiats*. Some entities begin to exist only when we draw their boundaries. Think of the states of the so-called Northwest Ordinance, as they were literally brought into being by Thomas Jefferson's pencil (and

ruler). Or think of when the colonial powers relied on cartography to subdivide the "heathen lands": a few lines of ink was all it took to legitimize—and simplify—their territorial conquests in spite of any preexisting social and political structures. As Mark Monmonier put it, sometimes the pen really is mightier than the sword (1991, 90). But even when the sword prevails, the outcome is a boundary which, though conforming with Brentano's thesis, bears witness to the double-barreled nature of the relevant dependency: were it not for those boundaries, those states would never have existed.

In fact, it is not even correct to speak of the relation between a boundary and the extended entity to which it must belong: every time we have a boundary, we have two entities, one on each side. Boundaries separate, but they separate two entities (or two parts of the same entity) which are continuous with each other. The Idaho-Wyoming boundary is thus a boundary of Idaho-but also a boundary of Wyoming. Who gets to claim ownership? Surely the boundary does not belong to both, for the states do not overlap. And we cannot simply say that it belongs to neither: these states occupy the whole territory by definition-no boundary can be left as a thin, unclaimed space between them. So? This is an old problem, and it goes without saying that it is not peculiar to geography. Euclid defined a boundary as "that which is an extremity of anything" (1908, bk. I, def. 13), and Aristotle made this more precise by defining the extremity of a thing *x* as "the first point beyond which it is not possible to find any part [of x], and the first point within which every part [of x] is" (Aristotle 1984a, V, 1022a). It is a definition that today we may also find in a dictionary. But what about the extremity itself—does it belong to x or to the complement of x? Consider the dilemma raised by Leonardo da Vinci in his *Notebooks*: what is it that divides the atmosphere from the water? Is it air or is it water (1938, 75-76)? Or think of the boundary of a black spot on a white surface, as Peirce wrote in his Logic of Quantity (1893/1933, 98): What color is this boundary-black or white? The puzzle also arises in relation to temporal boundaries. As Aristotle himself asked in the *Physics*, at the instant when an object begins to move, is it in motion or at rest (1984b, V, 234a-b)?

Here one cannot simply dispose of the puzzle by treating it as an artifact of the modeling process. As Antony Galton points out, properties like color or material constitution only apply to extended bodies, so in a way it makes no sense to ask whether a lower-dimensional entity is air, water, or colored (2003, 167). Yet at bottom the problem is one of *ownership*, not of physical characterization. The puzzle is purely topological and
originates in the fact that space and time, and hence the entities that according to common sense occupy space and endure through time, have the dense geometry of the continuum. There are no *adjacent* points, lines, or surfaces: either they coincide, or they are separated by an infinity of further points, lines, surfaces. In terms of the classic doctrine that goes back to Bolzano, this means that when two regions or bodies are adjacent one must be "closed"—that is, include the boundary among its parts—and the other "open" (1851, sec. 66). It's no help to stigmatize this doctrine as "monstrous," as Brentano did (1906/1976, 174). What happens when we cut an apple in half?—he asked. *Which* half will come out "closed," leaving the other half "open" and bleeding? Alas, this is the very problem at issue, and one cannot solve it by jeering at the apparent counterintuitiveness of the continuum.

Now, I like to think that precisely the distinction between *de dicto* and de re boundaries may help us solve the puzzle.³ On the one hand, with regard to artificial boundaries of the first sort, it is true that the question of their ownership can hardly be settled without contravening to the principle of sufficient reason. To assign the Idaho-Wyoming border to Idaho, or to Wyoming, would amount to a peculiar privileging of one state over the other. But precisely insofar as such boundaries are of our own making, it is also true that their actual ownership is no real issue. Simply put, we have not *decided* which state gets to own the border, as we have not decided which hemisphere gets to own the equator, or who gets to own the line separating my part of the desk from my colleague's. We have not decided this because the decision would be of no practical consequence whatsoever. And this sort of indeterminacy is of no metaphysical consequence either. To say that there is no "fact of the matter" here is not to endorse worldly indeterminacy, precisely because we are not dealing with bona fide facts; the indeterminacy pertains exclusively to our *fiat* practices. In this regard, the picture is no different from the one we get as we consider the vagueness of certain boundaries, such as the boundary of the Missouri Plateau. There is no clear-cut line separating the interior of the plateau from its exterior, but that doesn't mean that the plateau is a vague entity. It means that what we mean by 'the Missouri Plateau' is vague: nobody has been "fool enough," in David Lewis's phrase (1986, 212), to draw a precise line around its intended referent. It means that it is indeterminate (wide-scope) whether certain parcels of land belong to the Missouri Plateau, not that the Missouri Plateau is indeterminate (narrow-scope) with regard to the inclusion of those parcels of land.

On the other hand, when it comes to *de re* boundaries—such as the margin of a black spot on a white panel, or the line separating water from the atmosphere—one might think that their ownership need not be up for grabs. After all, in such cases we are confronted with two entities, one of which is *figure* while the other is *ground*, and it would not be implausible to resolve the asymmetry in favor of the former. This "ecological" intuition may be found, for instance, in Ray Jackendoff (1987, App. B). The black spot wins the status of "figure" over its white exterior, hence the relevant boundary belongs to the spot, not the exterior. Water wins over air, which is only "ground," hence the relevant boundary belongs to the ocean and not the atmosphere. And the same could be said of the surface of an apple. We never have two closed bodies in contact with each other, only one body embedded in its surroundings, and it is the body—one could say that gets to own the boundary. Only the apple is topologically closed; the background is open, and that is why the two can be genuinely in touch. As for Brentano's problem of explaining what goes on when we cut the apple in half, one could say that the dilemma betrays an incorrect model of the cutting process. Surely it would be arbitrary to elect one half as figure and the other as ground: after the cutting, each half is equally figure with respect to their common background, hence each will be enveloped by a complete surface. But that is not to say that such surfaces were already hidden inside the apple before the cutting. By dissecting an object we do not "bring to light new surfaces," as Ernest Adams has it (1984, 400),⁴ nor do we convert a de dicto boundary into two de re boundaries. Rather, topologically, when the cutting takes place the extant outer surface of the object is progressively deformed. A long, continuous process suddenly results in an abrupt topological change: there was one thing; now there are two. (Think of a splitting soap bubble.) There is indeed something deeply problematic about the abruptness of such a change, but never mind: whatever the explanation, the figure/ground account would go through.⁵

Still, I don't think this can be the whole story: figure wins, ground loses. For what happens when two figures compete? Think of the Dover cliffs hard to construe them as an "open" background for the waters of the Channel. Even harder if we consider that the cliffs themselves stand out on the horizon: should we say that they are open against the water but closed against the sky? And what about the line along which they all meet—water, rock, air? Whichever item gets the honor of "figure," the other two should be "ground," and hence "open." But how then could *they* meet? The topology of the continuum forbids contact between two closed bodies—but also between two open bodies. Clearly something is going astray. Brentano would say that this is yet another proof of the inadequacy of classical topology, and would begin to speak of *plerosis* and other complicated things that supposedly yield a better fit with intuition and common sense. I would rather say that here intuition and common sense begin to show their limits, and there is a serious possibility that we embarked on a wrong path.

Indeed, to me it seems obvious that for certain *de re* boundaries, especially those that mark the limits of material bodies, the ownership problem does not even arise, for on closer look such boundaries are not what we think. On closer look, as we know, an apple is not a solid, continuous object. On closer look, material objects are just swarms of subatomic particles frantically dancing in an otherwise empty space (the "material" volume of an apple is really only one billionth of what we commonly measure), and speaking of their "surfaces" is like speaking of the "flat top" of a fakir's bed of nails, as Peter Simons put it (1991, 91). On closer look, therefore, it makes little sense to speak of continuous objects separated by a common de re boundary. It makes no sense to ask who gets to own that boundary. All there is are smudgy bunches of hadrons and leptons, and if we really wish to insist, we can say that each such thing is figure against an empty ground. But the background is *empty*: there is *nothing* else that could claim ownership. To put it differently, on closer look the spatial boundaries of common material bodies involve the same degree of arbitrariness as those of any mathematical graph smoothed out of scattered and inexact data, the same degree of idealization as a drawing obtained by "connecting the dots," the same degree of abstraction as the figures' contours in a Seurat painting. It makes no sense to inquire about the owners of those boundaries, or rather, it only makes sense insofar as we recognize their ephemeral status as *fiat* demarcations that exist in virtue of our cognitive acts but that are not genuinely present in the autonomous-which is to say, mind-independent-physical world: hence, as de dicto boundaries. And we have seen that the ownership of such boundaries can be left indeterminate. It is the same when we move from objects to events. On closer look, as we know, a body's being at rest amounts to the fact that the vector sum of the motions of the trillions of restless atoms of which the body is composed, averaged over time, equals zero: hence, it makes no sense to speak of the *instant* at which an object begins to move.⁶ Either we are dealing with a *de dicto* boundary, in which case we know how to handle the problem (or rather, leave it unresolved), or we are dealing with particles that are *restlessly* in motion, in which case the problem does not even arise.

3 From Boundaries to Things

At this point, however, the very distinction between *de dicto* (artificial) and de re (natural) boundaries demands reexamination, and it is here that the question of its tenability bites deeply. Were it just a matter of boundaries, the exact extension of these two concepts might leave us unmoved; it is their intension that takes care of the problems. But once the de dicto/de re distinction has been recognized, it can be drawn across the board: not merely in relation to boundaries but also in relation to all those entities that may be said to have boundaries. If a certain entity enjoys natural boundaries, it is reasonable to suppose that its identity and survival conditions do not depend on us; it is a bona fide entity of its own. By contrast, if (some of) its boundaries are artificial—if they reflect the articulation of reality that is effected through human cognition and social practices—then the entity itself is to some degree a *fiat* entity, a product of our worldmaking. This is not to say that *fiat* entities of the second sort are imaginary or otherwise irreal entities. As Frege put it, the objectivity of the North Sea "is not affected by the fact that it is a matter of our arbitrary choice which part of all the water on the earth's surface we mark off and elect to call the 'North Sea'" (1884, sec. 26). It does, however, mean that such entities would only enjoy an individuality as a result of our cognitive and/or social practices, like the cookies carved out of the dough: their objectivity is independent, but their *individuality*—their being what they are, including their having the identity and survival conditions they have-depends on the baker's action. (In the terminology of John Searle [1995], they are "social objects": from a God's-eye point of view you don't see a cookie just as you don't see a nation or the North Sea; you only see us treating certain chunks of reality as cookies, nations, or the North Sea.)

Now, the existence of *de dicto* boundaries, and hence of *fiat* entities, is uncontroversial. We can even be more fine-grained, if we wish. We can, for instance, draw a further distinction between *fiat* entities that owe their existence to collective intentionality, or to the beliefs and habits of a community, as with geopolitical or social entities at large, and *fiat* entities that emerge instead from the cognitive acts of single individuals, beginning with perception, which as we know from our experience of Seurat paintings has the function of articulating reality in terms of continuous boundaries even when such boundaries are not genuinely present. Individual *fiats* are more ephemeral than social *fiats* because they are rigidly dependent (on *these* acts, taking place *now*) rather than generically dependent (on the existence of relevant acts of a certain kind). We can also distinguish

between those cases where a *fiat* entity is delineated or carved out as a proper part of a larger entity, as with our initial examples, and those cases where a *fiat* entity is obtained by circumcluding a number of smaller entities within larger wholes, as with Polynesia, the constellation Orion, or ordinary material bodies such as apples upon recognition of their microscopic structure, or as when we describe the world as consisting of forests, physaliae, schools of fish, swarms of bees, fleets of ships, pairs of shoes.⁷ (Natural language contributes to the generation of *fiat* entities also through the opposition between mass nouns, such as 'cattle', and count nouns, such as 'cow'. A hungry carnivore points toward a field and pronounces, "There is cattle over there." How does this pronouncement differ, in its object, from "There are cows over there"? Not, certainly, in the underlying real bovine material.) Regardless of all such refinements, it is clear that in each case *de dicto* boundaries are at work in articulating the reality with which we have to deal, and the entities obtained thereby are themselves of the *fiat* sort. Nor will such entities acquire a *bona fide* status upon further work on our side. As we have seen, no preexistent inner surface is brought to light by a process of cutting. Indeed, administrative or political borders may in course of time come to involve boundary-markers (barriers, walls, barbed-wire fences or electronic devices) that will tend, in accumulation, to replace what is initially a pure *de dicto* boundary with something more substantial. Yet again, this is not a process of ontic transformation. Such markers can be very robust, but not so robust as to turn the artificial into the natural. The Chinese Great Wall has survived for centuries-but the Berlin Wall lasted a mere fifty years, and Israel's "Separation Wall" in the West Bank has not been recognized by the International Court of Justice. On 22 September 2005, the U.S. Secretary of Homeland Security used his power to "waive in their entirety" the Endangered Species Act, the Migratory Bird Treaty Act, the National Environmental Policy Act, the Coastal Zone Management Act, the Clean Water Act, the Clean Air Act, and the National Historic Preservation Act to extend triple fencing along the U.S.-Mexico border through the Tijuana River estuary; a few months later, the San Diegans were playing volleyball with the Tijuanans using the fence at the beach border as a net. Even Romulus's plowshare could not make natural what natural is not; the blade cuts the soil, tears the grass, uproots all that lies on its *fiat* path.

Let us rather ask whether, and to what extent, it is appropriate to countenance the existence of natural, *de re* boundaries, and hence of natural, *bona fide* entities that do not depend on our deeds. By itself, the concept is perfectly intelligible, and we should be thankful to Lord Curzon for emphasizing its political significance. As Lucien Febvre famously argued, the very notion of a natural frontier began to emerge in the nineteenth century precisely as an expression of the idea that in some cases the limits within which we are allowed to act are set by Nature itself: natural boundaries are "fixed by destiny," they represent "ideals to conquer and realize" (1922, 354). It is one thing to take up arms on the crest line of the Pyrenees, quite another to sacrifice your life for a Jeffersonian line. It is not surprising that the Irish Republican Army wants the whole *island* of Ireland, but who would fight for the Wyoming Independentist Party? Still, even in relation to the geopolitical world, the natural/artificial distinction is all but robust. It is true that I had the impression of *seeing* the shoreline of Long Island from my plane; but it is also true that when you actually go there, groundlevel, things look very different. What looked from the air like a sharp line turns out to be an intricate disarray of stones, sand, algae, piers, boardwalks, concrete blocks, musk sediments, marshy spots, putrid waters, decayed fish. The same could be said for the much-celebrated Irish coast. And it's not just a matter of our disrespect for Nature; the worry would not change significantly if we took a close look at the coast of a virgin island in the middle of the ocean. Suppose we locate the boundary of the island at the water/sand interface. That boundary is constantly in flux, and it is only by filtering it through our cognitive apparatus—it is only by interpolating objects and concepts—that a clear-cut line may emerge. Even if both water and soil were perfectly still and each were materially homogeneous, it would be hard to locate the boundary with precision. One is reminded here of a question familiar from the early literature on fractals: just where, and how long, is the coastline of Britain? When measured with increased precision, this coastline could furnish lengths ten, a hundred or a thousand times greater than the length read off a schoolroom map. As ever-finer features are taken account of the total measured length increases, and Benoît Mandelbrot concluded that "there is usually no clear-cut gap between the realm of geography and details with which geography need not be concerned" (1967, 636).⁸ Cartographers know this well: one works with calipers, but their opening is not fixed by Nature. And if there is no fact of the matter answering the question "How long is the coastline?" one wonders whether it is even meaningful to think of the coast as of a natural, objectively determined boundary. One begins to wonder whether the island *itself* might not in some sense be the product of our subjective and approximate world-making. And what goes for islands goes for all prima facie natural entities: lakes, rivers, craters, glaciers, mountains, planets. "Even stars?" asked Israel Sheffler (1980, 205). Yes, even stars, answered Nelson Goodman: "As we make constellations by picking and putting together certain stars rather than others, so we make stars by drawing certain boundaries rather than others" (1983, 104).⁹

It is a short step, now, to extend such doubts to all those boundaries that intuitively belong to the *de re* category, hence all those entities that would seem to enjoy a bona fide, mind-independent reality. We have already seen that ordinary material objects tend to dissolve as soon as we acknowledge their microscopic structure: this apple is just a smudgy bunch of hadrons and leptons whose exact shape and properties are no more settled than those of a school of fish. Or perhaps they are more settled, for the causal connectedness of the apple appears to be stronger than that of a school of fish; yet this seems to be a matter of degree, not the sort of categorial difference involved in the *de re/de dicto* opposition. But never mind that line of reasoning; one need not resort to microscopes to realize that the notion of a natural object is far from being clear and distinct. Take this cat, Tibbles-a paradigm example of a living creature, hence a perfect candidate for the status of a bona fide individual. Regardless of its subatomic structure, the question of what counts as it has no clear answer. Tibbles is eating a chunk of tuna. When it was in the plate, that chunk was definitely not part of Tibbles. But now it is in Tibbles's mouth: is it part of Tibbles? Will it be part of Tibbles only after some chewing? Only when Tibbles swallows it? Only at the end of the digestive process? Surely, whatever mewing-and-purring portion of reality we mark off and elect to call "Tibbles" is something that exists in its own right; as with the North Sea, its objectivity is not affected by the fact that our stipulation is a matter of arbitrary choice. Yet surely the stipulation adds a fiat element to its individuality. Tibbles is not entirely a product of our own making, yet its identity and survival conditions will obtain only relative to us. And this would remain true even if our stipulation were not quite "arbitrary"-even if it fully returned what Jack Wilson (1999) calls the "biological individuality" of our cat. Alas, biology is a science—and as such it involves its own stipulations. A workshop held at the University of Utah in 2008, on the topic "Edges and Boundaries of Biological Objects," focused precisely on the thesis that "delimiting biological objects cannot be determined by empirical facts alone; which facts are salient, and what counts as evidence, often depend on theoretical and conceptual context" (University of Utah 2008).

What goes for objects goes for events, including biological processes. Earlier I mentioned a person's birth and death as obvious examples of *de re* temporal boundaries, yet the controversies on abortion and euthanasia seem to push for a different treatment: sometimes it is a matter of our deciding whether a person is still alive. We decide whether her "vital" functions are still in force, and the criteria for such decisions give expression to our beliefs, our principles, our theories. Similarly, on closer look the initial boundary of a person's life is hardly fixed by nature alone. Surely that boundary does not coincide with the person's *birth*—itself a rather intricate, messy, often protracted process (later cleaned up by the registry office)—but neither is there an earlier moment that fits the bill comfortably. When would it be: the moment of *fertilization*, when the membrane of the sperm cell fuses with that of the egg? Or upon formation of the zygote? When the zygote begins to undergo a process of genetic replication and cell division? Or upon formation of the morula? The beginning of the *implantation* process? The beginning of the *gastrulation* process, which gives structure to the embryo? The candidates are many. We can base our decision on as many factors we like, including up-to-the-minute scientific findings-but a decision it is. And if it is a matter of our (arbitrary or informed) decision, then the boundary is not genuinely de re after all and even a person's life becomes, to some extent at least, a *fiat* process.

What about natural properties? Natural kinds? Natural taxa? Here things get more complicated, of course, since the relevant notion of a boundary especially the notion of a *de re* boundary—is less clear. In fact, when it comes to such entities, the natural/artificial distinction intertwines with the whole realism/nominalism controversy and our geographic metaphor is bound to appear naïve and dismissive. Still, even here it's obvious and well known that, on closer look, our parochial concerns-historical and cultural circumstances, practical interests and limitations, theoretical priorities-tend to play a major role in the maps we draw of the world. Surely quadrupeds do not form a "natural" kind. But it would be quite remarkable, to use Catherine Elgin's example (1995, 297), if a taxonomy that draws the distinction between horses and zebras where we do aligned with categories fitting the cosmos as a whole but indifferent to our human faculties and ends (including our interest in domesticating animals). Surely emeroses do not form a natural kind—but neither do roses. Why settle on Rosa chinensis? Even in physics, our microscopic categories seem to suffer from a variety of human contingencies. If we construe different isotopes as variants of the same type of atom, it is because of certain reasonable interests that predominated in the development of our best theories. One could as well construe oxygen-17 and oxygen-18 as different types of atom altogether, hence as "natural" kinds of their own. Here the problem is not that there are no differences in the physical world; the problem is that

there are *too many* differences, and to privilege some over the others is to draw a *fiat* line—like the dotted line demarcating this apple from its exterior, or Tibbles from the rest of the world. Besides, it is a fact that even within specific domains of inquiry, our scientific practices are not uniform. The thought that the taxa countenanced by biology (for instance) are *fiat* entities seems to clash with "the certainty of biologists on the objective reality of evolution," as David Stamos puts it (2003, 131 n.35). Yet even today many taxonomists base their classifications more on phenetic than on phylogenetic criteria, regardless of the avowed principles under which they operate, and in phenetics a natural classification is simply one in which the members of each taxon are on the average more similar to one another than they are to members of other taxa "at the same level" (by itself a problematic notion). Maybe such taxonomists are being sloppy. Maybe the phylogenetic criterion is better. (It even fits the creationists' paradigm, as the God of Genesis supposedly created all living things to reproduce "according to their kinds.") But that's enough to cast the doubt: a criterion is a criterion. We should not shudder if, in the end, we read in the Annual Review of Ecology and Systematics that "taxa are human constructs" and "natural taxa are those that are natural to humans" (Sokal 1986, 424). And we should not complain if nonhuman animals have different tastes. After all, there are horses and zebras-but also zorses and hebras. Didn't Locke even see "the issue of a cat and a rat" (1690, III.vi.23)?

4 Conventionalism and Realism

It's pretty clear where all this is going. In the *Phaedrus*, Socrates famously recommends that we should carve the world along its natural joints, trying "not to splinter any part, as a bad butcher might do" (Plato 1995, 265d), and we know that science and common sense alike have taken this advice very seriously. If all boundaries were the product of some cognitive or social *fiat*, if the lines along which we "splinter" the world depended entirely on our *cognitive* joints and on the categories we employ in drawing up our maps, then our knowledge of the world would amount to neither more nor less than knowledge of those maps. The thesis according to which all boundaries—hence all entities—are of the *fiat* sort would take us straight to the brink of a precipice, to that extreme form of conventionalism according to which "there are no facts, just interpretations." On the other hand, to posit the existence of genuine, *bona fide* boundaries—to think that the world comes preorganized into natural objects and properties—reflects a form of naïve realism that does not seem to stand close scrutiny.

We know how the compromise solution goes. Perhaps all boundaries are, on closer look, de dicto boundaries; it doesn't follow that they must be utterly *arbitrary*, lacking all foundation in reality. It is like beef or veal, says Umberto Eco: "In different cultures the cuts vary, and so the names of certain dishes are not always easy to translate from one language to another, yet it would be very difficult to conceive of a cut that offered at the same moment the tip of the nose and the tail" (1997, 39). In other words, perhaps there are no obligatory paths, no one-way streets in the realm of Being; it still does not mean that anything goes. Some paths will still display a "no entry" sign, some constraints or "lines of resistance" may still be there, making it hard to cut the beast as we like. Out of metaphor, there may still be objective limits to our freedom to carve the world, and it is precisely in this spirit that the realism/conventionalism dichotomy is supposed to be handled. If it is presumptuous to think that the boundaries depicted on our physical maps are perfectly accurate, it is also implausible to think that they are completely off the mark. If it is implausible to think that biology can identify the exact moment at which a person's life begins, it is also implausible to suppose that life begins before fertilization, or at kindergarten. The very notion of "natural kind" to which scientists refer would not betray a commitment to naïve realism but, rather, a form of scientific realism whose cash value is first and foremost *pragmatic*. Just as the maps in our atlases have become more and more accurate, so will the maps drawn by the sciences. And just as cartographers are often forced to redraw their maps as a result of unexpected geopolitical changes-the artificial boundary of Israel, but also the prima facie natural boundary of sand between Lybia and Egypt, which keeps drifting under the wind—so biologists and other scientists will not refrain, if necessary, from updating their maps of nature in an effort to achieve greater accuracy and truthfulness. (The taxonomic misadventures of the platypus would be a good illustration of this fact. What sort of beast is that? Not a mammal, for it lays eggs. Not a reptile, for its blood is warm. Not a bird, for it's got four legs. For over eighty years, the naturalists were baffled: as René-Primevére Lesson observed in 1839, that double-damned beast had set itself "athwart the path of taxonomy to prove its fallaciousness."¹⁰ Yet it was there, and eventually the category "monotreme" was created ex novo.)

Now, I have no intention of denying the pragmatic reasonableness of this stance. But I do not share its fundamental optimism, as I do not recognize the distinction between metaphysical realism and scientific realism if not, indeed, on a deflated understanding of the term 'realism'. That nobody cuts the beef in funny ways does not mean that the laws of nature prevent it. It means that in spite of our cultural diversities, the culinary taste and aesthetic sense of human beings display surprisingly robust regularities-literally, as well as out of metaphor. Consider the debate on unrestricted composition. There is no question that we feel more at ease with certain mereological composites than with others. We feel at ease, for instance, with regard to such things as the fusion of Tibbles's parts (whatever they are), or even a platypus's parts; but when it comes to such unlovely and gerrymandered mixtures as Lewisean trout-turkeys, consisting of the front half of a trout and the back half of a turkey (Lewis 1991, 7), we feel uncomfortable. Such feelings may exhibit surprising regularities across contexts and cultures, yet arguably they rest on psychological biases and Gestalt factors that needn't have any bearing on how the world is actually structured. As James van Cleve has pointed out (1986, 145), even if we came up with a formula that jibed with all ordinary judgments about what counts as a natural fusion and what does not, it wouldn't follow that there may exist in nature only such objects as answer the formula. If anything, it would follow that the factors that guide our judgments of unity impose systematic constraints on our *fiat* articulations.

Besides, the controversies on biotechnology demonstrate that even such factors are less robust than one might think. We feel horror and disgust for such unlovely and gerrymandered mixtures as chimeras and genetically modified organisms, but we have long learned to feed on orange-mandarines, yogurt, peppermint, and seedless grapes-and we didn't have many scruples when it came to forcing zoological categories to make room for mules and poodles. According to the Royal Horticultural Society, in less than 150 years we have managed to fill the world with over 110,000 orchid hybrids. Either we are adamant that DNA is our model for an organism's individuality—and this is a metaphysical thesis that demands argument—or we must recognize that even the "no entry" signs on the way of Being are on closer look an expression of our contingent biases, however reasonable they might be. Of course, we are free to fight for or against such biases and to study their network in the spirit of honest descriptive metaphysics. After all, the world as we represent it is the only world we really care about, for it is the world on which we bet everything, including our happiness: that's why Husserl called it the Lebenswelt (1936, sec. 9h). Nonetheless, that would not be a way of solving the dilemma between realism and conventionalism; it would be a reasonable way of getting rid of it altogether.

For those who think that the dilemma ought to be addressed explicitly, however, I am going to conclude with three remarks aimed at deflating, at

least partly, the sense of collapse that usually comes with the conventionalist hypothesis, understood in the radical sense according to which there would be no *de re* boundaries whatsoever, whether "one way" or "no entry."

First, one should not mistake the conventionalist hypothesis with the ghost of Berkeleyan idealism. As I have described it here, the notion of a *de dicto* boundary is intelligible only to the extent that we acknowledge an appropriate real basis for the sorts of demarcation that are effected by our pencils, trenching tools, and cookie cutters. Even assuming that all boundaries are of this sort, and wholly arbitrary, it does not follow that *everything* is the product of a percipemus. This is why I stressed that the cognitive dependence of a *fiat* entity affects its individuality but not its *objectivity*. This may well be seen as a "last-ditch effort" to save realism, as Andrew Cortens dubbed it (2002, 54), but so be it. That the factual material onto which we project our categories should itself be a cognitive construct is a different, stronger thesis, which I do not even understand except in the figurative sense made popular by the thought experiments of rational skepticism (Cartesian demons, brains in a vat, the Matrix). From this perspective, neither should conventionalism be mistaken for Goodmanian irrealism. For Goodman-and for Richard Rorty-all we learn about the world is contained in right versions of it, "and while the underlying world, bereft of these, need not be denied to those who love it, it is perhaps on the whole a world well lost" (Goodman 1975, 60).¹¹ For a conventionalist, the world is boneless, impoverished, almost bankrupt-but our love for it is not at stake. For Goodman, a world-version need not be a version of the world, just as a Pegasus-picture need not be a picture of Pegasus. For a conventionalist, all the maps we draw are indeed maps of one and the same reality. Putnamian *relativism*, then? Even less. For Hilary Putnam, the "cookie cutter" metaphor founders on the question, "What are the 'parts' of the dough?" (1987, 19). No neutral description is available to compare a Leśniewskian world (with *x*, *y*, and the fusion of *x* and *y*) and a Carnapian world (with only x and y), and assigning a univocal meaning to 'exists' is already "wandering in Cloud Cockoo Land" (Putnam 2004, 85). For the conventionalist, the metaphor holds and 'exists' corresponds to the standard existential quantifier. The number of *fiat* entities is up for grabs; but the parts of the dough, which provide the appropriate real basis for our fiat acts, are whatever they are and the relevant mereology is a genuine piece of metaphysics. (As far as I am concerned, composition is unrestricted, and hence the mereology is Leśniewskian. So either Carnapians are not speaking with their quantifiers wide open—as when we say "There

is no beer," meaning "There is no beer in the refrigerator"—or else they are objectively wrong. Others might favor a different account. For instance, one may describe the conflict as stemming from the appeal to different "counting criteria," as Searle suggested [1995, 163], or from a disagreement about whether composition is "innocent," in Lewis's sense [1991, 85]. But surely the world could not care less about our criteria and our jury verdicts.)

Second, what's so bad with conventions being arbitrary? We have just seen that from a pragmatic perspective it is not the putative *de re* structure of the world that drives our ways of carving it up, in our folk life or in science, but rather the robustness and utility of certain ways against the ephemeralness and futility (if not the absurdity) of others. If we simply replaced Socrates' natural joints with Jefferson's pen strokes then it would be a disaster, and unfortunately that's exactly what happens in some cases. The thought of classifying people on the basis of their skin color or intelligence quotient isn't much better than the idea of drawing the Dutch-Belgian border through the houses of Baarle. Nonetheless, in most cases the arbitrariness of our conventions-those that govern our social interactions as well as those that are accorded scientific dignity-epitomizes a democratic reasonableness that treasures experience and cooperation. Conventionalism, just like pre-Kantian empiricism, does obliterate any substantive differences between the laws of nature and train timetables, as Maurizio Ferraris complains (2004, 17). But timetables are not drawn up at random. They ensue from the necessity to solve, in an arbitrary but efficient way, coordination problems that can be extremely complex and that can drastically impair our daily lives. If we come up with a timetable that works poorly, we change it. If a convention fails to measure up to our expectations, we replace it with a new, hopefully better one. The same can be said for the laws of nature. For a conventionalist, not all biological taxonomies (for instance) are on a par. Some are better than others because they better support the "laws" that govern biology's coordination game (laws of variation, selection, organic evolution, population growth, etc.). One may object that this sort of pragmatic efficiency calls for more than arbitrary, fiat demarcations-but the burden of proof is on the objector, not on the friend of conventionalism. Linnaeus's Systema Naturae (1735), the bible of all classical taxonomies, was soaked with essentialism-and the platypus didn't fit in. Darwin, by contrast, was adamant that the term 'species' is one "arbitrarily given for the sake of convenience to a set of individuals closely resembling each other" (1859, 52)-and his theory is much better.

Indeed, the arbitrariness of conventions does not even rule out that in some cases there is a single best way to go, a single best theory. On David Lewis's classic account, conventions are arbitrary insofar as they always admit of equally good "alternatives" (1969, 69). If a problem has a unique solution, then the solution is not conventional. By contrast, as I understand the term here, conventions are arbitrary insofar as they do not depend on the *bona fide* structure of the world. This need not imply that there always be at least two equally good choices that we could make; it simply implies that it is up to us—*in nostro arbitrio*—to make the choice. Hence I agree with Mark Heller: in some cases, "the adopted convention may be the single best choice, even the obvious best choice" (1990, 44 n.44). We may adopt a convention precisely because it is the best choice, but an arbitrary choice it remains. It follows, therefore, that when it comes to the monism/pluralism debate concerning the status of scientific taxonomies (hence: theories), a conventionalist may even side with the monist, at least relative to a certain domain of inquiry. A conventionalist may be a monist, not insofar as there is a unique *correct* way of carving up the world, but insofar as there may be a unique best way of doing that.

Third, for those who, like myself, believe in the significance of so-called prescriptive-or revisionary, revolutionary, hard-core-metaphysics, it is worth emphasizing that even a radical conventionalist stance need not yield the nihilist apocalypse heralded by postmodern propaganda. The pervasiveness of *de dicto* boundaries does not coincide, for instance, with the death of the individual. On a Putnamian metaphysics, there are no individuals except in a relative sense; on a Goodmanian metaphysics, we make individuals by drawing boundaries as we like, and this goes "all the way down" (1983, 107 n.6). Not so for my conventionalist-or not necessarily. For a conventionalist the identity of a cat, like the identity of a people, a nation, or a constellation, turns out to lack autonomous metaphysical thickness. But other individuals may present themselves. For instance, on a Quinean metaphysics, there is an individual corresponding to "the material content, however heterogeneous, of some portion of space-time, however disconnected and gerrymandered" (Quine 1960, 171). (What then distinguishes a material "substance" from other individuals is a detail, namely, "that there are relatively few atoms that lie partly in it (temporally) and partly outside.") That the content of some such portions of spacetime have *de re* boundaries is a possibility, but it is equally possible that the only boundaries are those warranted by geometry. Either way, the corresponding notion of an individual is perfectly intelligible. The relevant identity conditions are perfectly determinate, and one may suppose that

it is perfectly determinate, for any property, whether any given individual possesses it. Such individuals are perfectly nonconventional, yet the overall picture is one that a conventionalist is free to endorse. The conventionalist stance simply entails that *which* of them come to play a role in our life is up to us. From a God's-eye point of view, they are all equally real. It's just that only some are salient for us, only some make us "feel comfortable," and only some are selected by us through the imposition of more or less precise and official *de dicto* boundaries. (Heller is one philosopher who explicitly endorses this view.¹²)

Evidently, a metaphysics of this sort presupposes the existence of a large, all-embracing four-dimensional "dough." But we have already seen that such a presupposition is not incompatible with a conventionalist stance. Nor is reference to the boundaries warranted by geometry vetoed by that stance, as though it surreptitiously reintroduced *de re* demarcations. As we have outlined it here, the *de dicto/de re* distinction does not apply to the geometric boundaries of spacetime any more than it applies to the boundaries presupposed by set theory. Those are bare boundaries, so to speak: they are not "artificial," but neither are they "natural." As for the overall plausibility of the theory, this is not the place to embark on an articulated defense. Surely the intuitive plausibility is pretty low, and perhaps also its scientific tenability. But philosophically the deflationary purism of a theory of this sort would have some advantages, including the extermination of the essentialist cancer that besets those metaphysical theories which try to save common sense against the paradoxes of persistence, vagueness, and material constitution.

But as I said, this is not the place to embark on an articulated defense of such metaphysics. (And it is but an example. A Sidellean metaphysics of "pure stuff" would do just as well.¹³) The relevant point is that, also in relation to the third remark, the conventionalist hypothesis need not result in a delegitimization of all philosophical inquiry; only a redistribution of the relative tasks and concerns of the different fields of inquiry, including metaphysics along with physics, psychology, or sociology. The costs are obvious, for epistemology and also for ethics—but so are the advantages. There are, on this view, no obligatory or forbidden paths; conventionalism is as liberal as it gets, and it is up to us to erect the "one way" or "no entry" signposts that we find appropriate, just as it is up to us to remove them when things take a turn for the worse. As Michael Dummett put it, the picture of reality as an amorphous lump, not yet articulated into discrete objects, is a good one "so long as we make the right use of it" (1973, 577). Most importantly, given that even common sense is far from being an all-or-nothing affair, except for certain surprising regularities, it is up to us to acknowledge *our* parochial limits without baptizing them as "natural." At least then we would stop pretending that boundary wars can have a "just" solution. To conclude on a rhetorical note, if there is a solution it lies in the reciprocal and democratic agreement among all interested parties, hard as it might be to achieve.

Acknowledgments

Many thanks to Sara Bernstein for her commentary at the Inland Northwest Philosophy Conference and to Roberto Casati, Elena Casetta, Umberto Eco, Antony Galton, Allen Hazen, Hud Hudson, David Mark, Barry Smith, Roy Sorensen, and Dean Zimmerman for helpful discussion on various aspects of this work. I am also thankful to Luca Morena and Giuliano Torrengo for their comments and criticisms on an earlier version of the paper, parts of which appeared in Italian as "Teoria e pratica dei confini," *Sistemi intelligenti* 17 (2005): 399–418.

Notes

- 1. See Smith 1995, which appears in extended form as Smith 2001.
- 2. See Abelard 1994, 26; cf. Aristotle 1984a, XI, 1060b12ff.
- 3. Here I draw on Smith and Varzi 2000.

4. In the same spirit, D. Zimmerman speaks of "a part of the object which was first imbedded in the [object] and then disclosed wearing a new skin" (1996, 25).

- 5. Here I refer to Varzi 1997, esp. p. 42.
- 6. Here I concur with Galton (1994).
- 7. On all this, see Smith 1995, 1999, 2001.

8. Some philosophers take questions such as these, and the lack of a definite answer, to be a sign of ontological vagueness; see, e.g., Copeland 1995.

9. See also Goodman's original reply in 1980, 213.

10. I owe this quotation to Eco (1997, 216; English trans., 250), though I was not able to locate the source.

- 11. Cf. Rorty 1972.
- 12. See Heller 1990.
- 13. See Sidelle 1989.

References

Abelard, P. 1994. From the "Glosses on Porphyry." In *Five Texts on the Mediaeval Problem of Universals*, trans. P. V. Spade. Indianapolis: Hackett.

Adams, E. W. 1984. On the superficial. Pacific Philosophical Quarterly 65:386-407.

Aristotle. 1984a. Metaphysics. In *The Complete Works of Aristotle*, vol. 2. ed. J. Barnes. Princeton: Princeton University Press.

Aristotle. 1984b. Physics. In *The Complete Works of Aristotle*, vol. 1. ed. J. Barnes. Princeton: Princeton University Press.

Bolzano, B. 1851. *Paradoxien des Unendlichen*. Ed. F. Přhonský. Leipzig: Reclam. English translation: *Paradoxes of the Infinite*, trans. D. A. Steele. London: Routledge & Kegan Paul, 1950.

Brentano, F. 1906/1976. Nativistische, empiristische und anoetistische Theorie unserer Raumvorstellung. In *Philosophische Untersuchungen zu Raum, Zeit und Kontinuum*, ed. S. Körner and R. M. Chisholm. Hamburg: Meiner. English translation: Nativistic, empiricist, and anoetistic theories of our presentation of space, trans. B. Smith. In *Philosophical Investigations on Space, Time, and the Continuum*. London: Croom Helm, 1988.

Brentano, F. 1914/1976. Vom Kontinuierlichen. In *Philosophische Untersuchungen zu Raum, Zeit und Kontinuum*, ed. S. Körner and R. M. Chisholm. Hamburg: Meiner. English translation: On what is continuous, trans. B. Smith. In *Philosophical Investigations on Space, Time, and the Continuum*. London: Croom Helm, 1988.

Calmette, J. 1913. La frontière pyrénéenne entre la France et l'Aragon. *Revue des Pyrénées* 25:1–19.

Copeland, B. J. 1995. On vague objects, fuzzy logic, and fractal boundaries. *Southern Journal of Philosophy* 33:83–96.

Cortens, A. 2002. Dividing the world into objects. In *Realism and Antirealism*, ed. W. P. Alston. Ithaca: Cornell University Press.

Curzon, Lord of Kedleston. 1907. Frontiers: The Romanes Lecture. Oxford: Clarendon Press.

Darwin, C. 1859/1985. *The Origin of Species by Means of Natural Selection*. London: Murray. Reproduced by Penguin Classics.

da Vinci, L. 1938. *The Notebooks: Selected English Translations*. Trans. E. MacCurdy. London: Reynal and Hitchcock.

Dummett, M. 1973. Frege. Philosophy of Language. London: Duckworth.

Eco, U. 1997. *Kant e l'ornitorinco*. Milan: Bompiani. English translation: *Kant and the Platypus: Essays on Language and Cognition*, trans. A. McEwen. New York: Harcourt Brace, 1999.

Elgin, C. Z. 1995. Unnatural science. Journal of Philosophy 92:289-302.

Euclid. 1908. Elements. Trans. T. Heath. Cambridge: Cambridge University Press.

Febvre, L. 1922. La terre et l'évolution humaine: Introduction géographique à l'histoire.
Paris: Alben Michel. English translation: A Geographical Introduction to History, trans.
E. G. Mountford and J. H. Paxton. New York: Knopf, 1925.

Ferraris, M. 2004. Goodbye Kant! Milan: Bompiani.

Frege, G. 1884. *Die Grundlagen der Arithmetik*. Breslau: Köbner. English translation: *Foundations of Arithmetic*, trans. J. L. Austin. Oxford: Blackwell, 1950.

Galton, A. 1994. Instantaneous events. In *Temporal Logic: Proceedings of the ICTL Workshop*, ed. H. J. Ohlbach. Saarbrücken: Max-Planck-Institut für Informatik, Technical Report MPI-I-94-230.

Galton, A. 2003. On the ontological status of geographical boundaries. In *Foundations of Geographic Information Science*, ed. M. Duckham, M. F. Goodchild, and M. F. Worboys. London: Taylor and Francis.

Goodman, N. 1975. Words, works, worlds. Erkenntnis 9:57-73.

Goodman, N. 1980. On starmaking. Synthese 45:210-215.

Goodman, N. 1983. Notes on a well-made world. Erkenntnis 19:99-108.

Heller, M. 1990. *The Ontology of Physical Objects: Four-Dimensional Hunks of Matter*. Cambridge: Cambridge University Press.

Husserl, E. 1936. Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. *Philosophia* 1:77–176. English translation: *The Crisis of European Sciences and Transcendental Phenomenology: An Introduction to Phenomenological Philosophy*, trans. D. Carr. Evanston: Northwestern University Press, 1970.

Jackendoff, R. 1987. Consciousness and the Computational Mind. Cambridge, Mass.: MIT Press.

Lewis, D. K. 1969. *Convention: A Philosophical Study*. Cambridge: Cambridge University Press.

Lewis, D. K. 1986. On the Plurality of Worlds. Oxford: Blackwell.

Lewis, D. K. 1991. Part of Classes. Oxford: Blackwell.

Linnaeus, C. 1735. Systema naturae, sive, Regna tria Naturae systematice proposita per classes, ordines, genera & species. Leiden: de Groot.

Locke, J. 1690. An Essay Concerning Human Understanding. London: Bassett.

Mandelbrot, B. B. 1967. How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science* 156:636–638.

Monmonier, M. 1991. How to Lie with Maps. Chicago: University of Chicago Press.

Peirce, C. S. 1893/1933. The logic of quantity. In *Collected Papers of Charles Sanders Peirce*, vol. IV, ed. C. Hartshorne and P. Weiss. Cambridge, Mass.: Harvard University Press.

Plato. 1995. Phaedrus. Trans. A. Nehamas and P. Woodruff. Indianapolis: Hackett.

Putnam, H. 1987. The Many Faces of Realism. LaSalle: Open Court.

Putnam, H. 2004. *Ethics without Ontology*. Cambridge, Mass.: Harvard University Press.

Quine, W. V. O. 1960. Word and Object. Cambridge, Mass.: MIT Press.

Rorty, R. 1972. The world well lost. Journal of Philosophy 59:649-665.

Scheffler, I. 1980. The wonderful worlds of Goodman. Synthese 45:201–209.

Searle, J. R. 1995. The Construction of Social Reality. New York: Free Press.

Sidelle, A. 1989. *Necessity, Essence, and Individuation: A Defense of Conventionalism.* Ithaca: Cornell University Press.

Simons, P. M. 1991. Faces, boundaries, and thin layers. In *Certainty and Surface in Epistemology and Philosophical Method: Essays in Honor of Avrum Stroll*, ed. A. P. Martinich and M. J. White. Lewiston, Maine: Edwin Mellen Press.

Smith, B. 1995. On drawing lines on a map. In *Spatial Information Theory: A Theoretical Basis for GIS*, ed. A. U. Frank and W. Kuhn. Berlin: Springer Verlag.

Smith, B. 1999. Agglomerations. In *Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science*, ed. C. Freksa and D. M. Mark. Berlin: Springer Verlag.

Smith, B. 2001. Fiat objects. Topoi 20:131-148.

Smith, B., and A. C. Varzi. 2000. Fiat and bona fide boundaries. *Philosophy and Phenomenological Research* 60:401–420.

Sokal, R. R. 1986. Phenetic taxonomy: Theory and methods. *Annual Review of Ecology and Systematics* 17:423–442.

Stamos, D. N. 2003. *The Species Problem: Biological Species, Ontology, and the Metaphysics of Biology*. Lanham, Md.: Lexington Books.

Suarez, F. 1856–1866. Metaphysicae disputationes. In *Opera omnia*, vol. 26, ed. C. Berton. Paris: Vivés.

University of Utah. 2008. *Edges and Boundaries of Biological Objects*. Annual Colloquium, University of Utah, Department of Philosophy, March 13–15. http://www.phylosophy.org/eb.

van Cleve, J. 1986. Mereological essentialism, mereological conjunctivism, and identity through time. *Midwest Studies in Philosophy* 11:141–156.

Varzi, A. C. 1997. Boundaries, continuity, and contact. Noûs 31:26-58.

Wilson, J. 1999. *Biological Individuality: The Identity and Persistence of Living Entities*. Cambridge: Cambridge University Press.

Zimmerman, D. 1996. Could extended objects be made out of simple parts? *Philosophy and Phenomenological Research* 56:1–29.

8 Natural Kinds and Biological Realisms

Michael Devitt

The species category does not exist. —Ereshefsky (1998, 113)

Taxa of higher rank than species do not exist in the same sense as do species. —Eldredge and Cracraft (1980, 327)

To the cladist true believer, there is no such thing as a reptile. —Sterelny and Griffiths (1999, 197)

1 Introduction

There are a number of "realism" issues in biology, issues about what "exists," what is "real," what is "objective."¹ In general, realism issues tend to be confused and the biological ones are no exception. We shall see that the interesting realism issues in biology are best seen as ones over which kinds "carve nature at its joints"—which kinds are "natural kinds"—and that seeing them as "realism" issues has caused unclarity and confusion.²

I shall start with the issue that arises out of the debate between "species monists" who think that there is just one good "species concept"—one good account of what it is to be a species—and "species pluralists" who think that there are many. To that end, in section 2 I summarize the various species concepts, and in section 3 the motivation for pluralism. In section 4, I describe realism in general, particularly "realism about the external world." Against this background I turn, in section 5, to the issue in question, focusing on the apparent clash between Marc Ereshefsky's "pluralistic antirealism" (1998) and Philip Kitcher's "pluralistic realism" (1984). In section 6 I consider some "realism" issues about genera and higher categories in the Linnaean hierarchy.

2 Species Concepts

Species pluralism arises out of the controversy over which species concept is correct. "The species problem is one of the oldest controversies in natural history" (O'Hara 1993, 231); it is "one of the thorniest issues in theoretical biology" (Kitcher 2003, xii).³ There are around two dozen species concepts and "at least seven well-accepted ones" (Ereshefsky 1998, 103). Samir Okasha (2002) places them in "four broad categories":

1. *Phenetic* concepts On this sort of view, organisms are grouped into species on the basis of overall similarity of phenotypic traits. This is thought by its proponents to have the advantage of being fully "operational." Okasha says that phenetic concepts are "the least popular" (2002, 199), and this is hardly surprising since they arise from the "philosophical attitude . . . of empiricism" (Sokal and Crovello 1970, 29). "Phenetic taxonomists have often wanted to segregate taxonomy from theory" (StereIny and Griffiths 1999, 196).

2. *Biological Species* concepts (BSC) The most famous example of a BSC is due to Ernst Mayr. He defined species as "groups of interbreeding natural populations that are reproductively isolated from other such groups" (Mayr 1969, 26). Kim Sterelny and Paul Griffiths remark that "If the received view has a received species concept," it is BSC (1999, 188).

3. *Ecological Niche* concepts According to these concepts, a species occupies a certain ecological niche. "A species is a lineage . . . which occupies an adaptive zone minimally different from that of any other lineage in its range and evolves separately from all lineages outside its range" (van Valen 1976, 70). Okasha puts the view succinctly: species "exploit the same set of environmental resources and habitats" (2002, 200).

4. *Phylogenetic-Cladistic* concepts (P-CC) On this view we "identify species in terms of evolutionary history . . . [with] particular chunks of the genealogical nexus . . . Species come into existence when an existing lineage splits into two . . . and go extinct when the lineage divides, or when all members of the species die" (Okasha 2002, 200). Sterelny and Griffiths claim that "something like a consensus has emerged in favor of a *cladistic* conception of systematics" (1999, 194).

The various species concepts are answers to what Mayr (1982, 253–254) calls the "species category" problem: they are telling us what it is it for a kind to be a species rather than a subspecies (variety), genus or what have you. In claiming this I am not being controversial. However, my claim about the following questions is controversial: Do these species concepts

tell us anything about what Mayr calls the "species taxon" problem? Do they tell us about what it is for an organism to be a member of a kind that happens to be a species? The answer to the "first-level" taxon problem tells us why Fido is a dog; an answer to the "second-level" category problem tells us why dogs are a species but poodles are not. Now it is common to think that species concepts not only answer the category problem but also the taxon problem (Dupré 1981; Sterelny and Griffiths 1999; Wilson 1999; Okasha 2002). I have argued elsewhere (Devitt 2008) that this common thought is wrong. But the main point to make here is that theories about the species category are one thing, theories about species taxa are another. In particular, realism about the *category* is one thing and realism about *taxa* is another.⁴

3 Species Pluralism

As I have noted, controversy has raged over which of the many species concepts is right. In the face of this controversy some have argued that we should abandon the monist idea that just one concept *is* right, and hence that there is just one species category. Rather we should adopt the pluralist idea that *many* of the concepts are right and hence that there are many species categories. According to Kitcher, many concepts "can be motivated by their utility for pursuing a particular type of biological inquiry" (1984, 118). Kyle Stanford puts the point thus: "certain explanatory demands are *inextricably bound* to certain species concepts" (1995, 72). And there are many different, but equally legitimate, types of biological inquiry and explanatory demands: "we have independent and legitimate explanatory interests in biology which require distinct concepts of species" (ibid., 76).⁵

In brief, there are a range of cases that force our attention to "the diversity of biological interests" (Kitcher 1984, 125), motivating different ways to classify organisms into species. And these ways will often lead to different classifications. Thus Ereshefsky claims that the BSC interbreeding and the P-CC phylogenetic approaches to species "carve the tree of life in different ways. Many interbreeding species fail to be phylogenetic species, and many phylogenetic species fail to be interbreeding species" (1998, 105).

I have doubts about the extreme form of pluralism urged by Kitcher, doubts that I will air in section 5. Still, the case for some form of pluralism strikes me as strong. But what bearing does this monism-pluralism issue have on biological "realism"? To assess this, we need to be clear about what

biological realism is. And to get clear about that it will help to start by considering realism in general.

4 Realism in General

The background issue that is most relevant is often known as "realism about the external world," concerned initially with the observable entities of common sense but spreading to scientific entities, both observable and unobservable. Let us attend only to scientific entities. What is realism about these entities? It has two dimensions, one committed to the *existence* of entities, the other to their *mind-independence*. We can capture the doctrine well enough as follows:

Realism: Entities of most scientific kinds exist mind-independently.

I have argued the case for this doctrine at length in *Realism and Truth* (1997) and subsequent papers (1999, 2001, 2002, 2005).

This doctrine, Realism, needs to be kept quite distinct from another traditionally called "realism." This other realism is about "universals," abstract entities like kinds, properties or sets. Insofar as this realism arises from the "one-over-many problem," I follow Quine in thinking it arises from a pseudo-problem (Devitt 1980; Devitt and Sterelny 1999, 277–279). Furthermore, I lean toward the view that there are *no* good reasons for realism about universals—that is, I lean toward nominalism—but I shall not try to argue the matter. Indeed, it seems to me best to discuss the issues of realism in biology without any commitment on this vexed, millennia-old, metaphysical issue.⁶ I shall continue to talk of kinds as if they existed but remain neutral on whether this is a mere manner of speaking that can be paraphrased away or whether it amounts to a real commitment.

One further aspect of realism debates looms large in biology. Philosophers have been concerned not only with whether the posits of science exist mind-independently but also with whether these posits are "appropriately special" rather than somewhat arbitrary. In particular there has been a concern about whether our scientific posits "carve nature at its joints," about whether there is something in the nature of the world that, in some sense, *determines* our categorization of it. I take this to be a concern about whether the kind of entity posited by a theory plays a causally significant role, whether it is partly *because* an entity is of that kind that it has the characteristics and behavior that it has. Theories need to posit such so-called "natural" kinds if the theories are to be genuinely explanatory. And Realists are likely to take it for granted that their paradigm entities—for example, cats and planets—are indeed of natural kinds.⁷

This admittedly vague idea of naturalness is what Kitcher has in mind in saying that "natural kinds are the sets that one picks out in giving explanations" (1984, 132 n.16). Similarly, Richard Boyd says that "the naturalness (and the 'reality') of natural kinds consists solely in the contribution that reference to them makes to [the accommodation between conceptual and classificatory practices and causal structures]" (1999, 141). There are two things we should note about definitions along these lines. First, they are insufficient: being a motor car, a hammer or a paperweight are causal-explanatory kinds. We need somehow to capture the idea that natural kinds are causal-explanatory "in science." Second, explanatory significance, and hence naturalness, comes in degrees: positing some kinds may be very explanatory, positing others only somewhat explanatory, positing still others not explanatory at all.

This provides one reason for not using the term 'realism' to label the issue of naturalness: *existence, unlike naturalness, does not come in degrees*. But a more important reason is that the requirement that a kind be natural is *quite distinct from* the requirement that entities of that kind exist objectively and mind-independently. As we shall see in the next section, it is easy to name kinds of entities that are more arbitrary than causal-explanatory, and yet those entities still exist mind-independently. And if there were kinds of entities that were mind-dependent in, say, a Kantian way, those kinds could still be causal-explanatory. So I shall keep the issues distinct. But they have not been kept distinct in biology, with some unfortunate consequences that will soon become apparent.

I shall start with Ereshefsky's antirealism about species and his apparent disagreement with Kitcher.

5 Ereshefsky versus Kitcher

In section 2, we noted the importance of distinguishing realism issues at two levels to do with species: at the first level we are concerned with species taxa—for example, tigers—and at the second level we are concerned with the species category itself. Ereshefsky is concerned with the latter. He claims that "the species category does not exist," but he "does not call into question the reality of those lineages we call 'species'" (1998, 113, 104).

Ereshefsky's argument, the "Heterogeneity Argument," starts from species pluralism: "a number of species concepts should be accepted as legitimate" (1998, 103). He takes this pluralism to imply antirealism and thus sees himself as being at odds with pluralistic realists like Kitcher (1984) and John Dupré (1993). Ereshefsky rightly points out that "species pluralism implies that the world contains different types of species" (1998, 111). He goes on:

What do these different types of species have in common that renders them species? If species taxa lack a common unifying feature, then we have reason to doubt the existence of the species category. (Ibid.)

Ereshefsky finds a suggestion in the literature for what the common feature might be, a suggestion with two components: (i) a similarity in the "process that renders species taxa cohesive entities"; (ii) a similarity of "structure." He argues that neither component can do the job (ibid., 111–113), and concludes that "what is left as the common feature of species taxa is the term 'species'" (ibid., 113).⁸ Hence he draws his conclusion that the species category does not exist.⁹

What could he mean by this conclusion? He certainly does not mean that tigers and the like do not exist: we have noted already that he does not reject first-level taxa realism. And he seems to accept that there are "different types of species," each one captured by a different species concept; for example, "interbreeding species," "phylogenetic species," and "ecological species," or more briefly, "biospecies," "phylospecies," and "ecospecies" (ibid., 115, 117). So he accepts that there exist organisms that are members of a biospecies, organisms that are members of a phylospecies, organisms that are members of a phylospecies, organisms that are members of other types of species. But then, since biospecies, phylospecies, ecospecies, and perhaps others are all types of *species*, all of these organisms are members of a species. Indeed to be a member of a species simply *is* to be a member of a biospecies, phylospecies, ecospecies or perhaps other types of species. So what entity could Ereshefsky be denying the existence of?

His response to a "realist rejoinder" points to the answer. In this response, he denies that there is any "distinctive commonality" among biospecies, phylospecies, and ecospecies. He is not satisfied with a disjunctive species category of the sort just illustrated. "A disjunctive definition of the species category would not tell us why various taxa are species . . . disjunctive definitions lack ontological import" (1998, 115). What he finds lacking in the species category is explanatory significance or anything close to that; the category is too close to being arbitrary. But how could this lack be a lack of *existence*? I take it that Ereshefsky must be thinking of the species category as an abstract entity, a "universal," and

must be presupposing a sort of *selective* realism about such entities. Whereas an *unselective* realist is committed, roughly, to there being a universal for every predicate, a selective realist is committed to there being one for some but not all predicates.¹⁰ And Ereshefsky is committed to there being one for a predicate where the objects it applies to share a distinctive commonality, but not for any other predicates. According to his pluralism, species lack that commonality. So the species category, a universal, does not exist.

Now, it would be better if we could capture Ereshefsky's position on biology without commitment to a heavy-duty, highly controversial, metaphysical thesis of selective realism about universals. And we can. We can capture it as a rejection of the view that the species category is a "natural" kind. And this rejection is not based on denying the *existence* of anything but on denying the *explanatory significance* of kinds being species.

To see that non-naturalness does not arise from nonexistence, it is helpful to note that entities of many kinds exist, even exist mindindependently, where their being members of those kinds is of little or no explanatory significance. Consider cousins, for example: these include first cousins, second cousins, fifth cousins thrice removed, and so on. The explanatory significance of being a cousin is surely close to zero, and we get even closer to zero if we consider step-cousins. And there is nothing to stop us naming a *totally arbitrary and non-natural* kind: thus we could call anything that is either an acid, a river or a bachelor a "grugru." Yet despite non-naturalness and explanatory insignificance, cousins, stepcousins, and grugrus really exist, and do so mind-independently.

So the best metaphysical message to take from biological pluralism is not that the species category does not exist or that it is not "real." Nor is the best message that species do not correspond "to something in the *objective structure* of nature" (Kitcher 1984, 128), are not "an objective feature of the living world" (Sterelny and Griffiths 1999, 180). For, according to pluralism to be a species is to be a biospecies, phylospecies, ecospecies or perhaps other type of species—and there is nothing subjective or otherwise mind-dependent about that.

To think clearly about realism issues it is vital to distinguish sharply two sorts of freedom, a freedom we have and a freedom we do not (Devitt 1997, 245). The freedom we *do* have is to choose to name any kind we like, whether for explanatory reasons or frivolous reasons or no reason at all; naming kinds is a subjective matter. The freedom we *do not* have is to choose whether something is a member of a kind, whatever our reason for naming that kind in the first place; kind-membership is an objective

matter. We have chosen to name cats for very good explanatory reasons and grugrus for no good reason at all, but grugrus exist as objectively and mind-independently as cats—and they existed long before I named them "grugrus." My naming them "grugrus" didn't make them grugrus any more than people naming cats "cats" made them cats. It is common to talk as if, in doing science, we impose our concepts to "carve up reality." But this is not literally so: we choose our concepts in an attempt to discover the causally significant features of a nature that is already "carved up." The importance of distinguishing theory-making from world-making could hardly be exaggerated.¹¹

In sum, Ereshefsky is best seen as rejecting the view that the species category is a natural kind because that category is not explanatory. He is best seen as rejecting this because that is what his argument from pluralism supports without the baggage of selective realism about universals.

I noted that Ereshefsky sees himself in disagreement with Kitcher and Dupré who, he claims, "suggest a realistic interpretation of species pluralism: various species concepts provide equally real classifications of the organic world" (Ereshefsky 1998, 103). Is there an actual disagreement, and if so what is it? I shall only consider Kitcher and shall set aside any possible disagreement arising solely from Ereshefsky's selective realism about universals.

Kitcher does, of course, *call* his position "pluralistic realism," but what exactly is "realist" about his position? First of all he sets aside as "trivially true" a realism about the species category (and taxa) that is like the doctrine I have called simply Realism, requiring only mind-independent existence. He is interested in a stronger realism:

Pluralistic realism rests on the idea that our objective interests may be diverse, that we may be objectively correct in pursuing biological inquiries which demand different forms of explanation, so that the patterning of nature generated in different areas of biology may cross-classify the constituents of nature. (Kitcher 1984, 128)

Kitcher clearly thinks that *various species categories*—biospecies, phylospecies, ecospecies, and the like—are explanatory and hence natural kinds. Setting aside for a moment a disagreement about the variety of species categories, I assume that Ereshefsky would agree (1998, 117). But, of course, the view that the various species categories are natural kinds is not the view that *the species category itself* is natural. The latter view requires that it be explanatory significant that some kinds are *species*. Yet, according to Ereshefsky's argument, pluralism has the consequence that the species category is disjunctive and not explanatory. So if Kitcher's pluralistic

realism is about the species category itself, and not merely about various species categories, then Ereshefsky is indeed in disagreement with him. *Is* Kitcher's realism about the species category itself? He certainly does not *argue* that the species category is explanatory. And the message I take from "the moral for philosophy of science" that he draws at the end of his paper is that he thinks that the category is *not* explanatory: 'species' refers not to a natural kind but to a "heterogeneous collection" of natural kinds (1984, 129). If I am right about this then Kitcher's pluralistic "realism" is not, in this respect, different from Ereshefsky's pluralistic "antirealism," and so they seem not to be in disagreement after all.

My suggestion is that the appearance of disagreement here between Ereshefsky and Kitcher has arisen from confusingly describing an issue about natural kinds as an issue about realism.

However, Ereshefsky and Kitcher do differ in their pluralisms and hence in the variety of species categories that they are prepared to be "realist" about. Kitcher's pluralism is radical. He follows Mayr in drawing a distinction—a very important one in my view (Devitt 2008)—between two types of explanation in biology, ones that Kitcher calls "structural" and "historical." Structural explanations seek to "explain the properties of organisms by means of underlying structures and mechanisms"; the historical seek to "identify the evolutionary forces that have shaped the morphology, behavior, ecology, and distribution of past and present organisms" (1984, 121). Kitcher claims that these two types "generate different schemes for classifying organisms" (ibid., 122). He finds variations within the two types, which leads him to posit nine different taxonomies *and hence nine distinct species categories* (ibid., 124).

Ereshefsky's pluralism is more conservative. He is prepared to accept only species categories that are justified by historically motivated taxonomies, for example, the categories of biospecies, phylospecies, and ecospecies. His form of pluralism

assumes that all species taxa are genealogical entities. To assume otherwise places species outside of the domain of evolutionary biology. The explanatory backbone of evolutionary theory is the assumption that organisms are connected by geneal-ogy. (Ereshefsky 1998, 107)

However, this is not really an *argument* against Kitcher's radicalism. For Ereshefsky is attending only to the explanatory needs of evolutionary biology, whereas Kitcher is emphasizing that there are *other* explanatory concerns in biology: there are the concerns of *structural* explanations. And Kitcher thinks that these explanations motivate taxonomies, and hence

species categories, that are different from those motivated by historical explanations.

Is there an argument against Kitcher's radicalism? I think so. Suppose that we go along with him about the nine different taxonomies: what about the inference to nine different categories? If the taxa picked out by a certain taxonomy are to justify a certain species category, *S*, it has to be *explanatorily significant* that the taxa are *S*. Now given the role of species in theories of evolution it is plausible to think that a *historically* motivated taxonomy *will* justify a species category: as Ereshefsky puts it, "species taxa are the paradigmatic units in which descent with modification occurs" (1998, 107).¹² But why suppose that a *structurally* motivated taxonomy will justify a category? That is, even if structural explanations demand certain taxonomies, what significance is there for such explanations that some taxon is a *species*? Consider Kitcher's example:

A biologist may be concerned to understand how, in a particular group of bivalve mollusks, the hinge always comes to a particular form. The explanation that is sought will describe the developmental process of hinge formation, tracing the final morphology to a sequence of tissue or cellular interactions, perhaps even identifying the stages in ontogeny at which different genes are expressed. (Kitcher 1984, 121)

It is hard to see how it makes any difference to the structural explanation we seek whether that group of mollusks is a species, subspecies, genus or whatever. Our interest in structural explanations may demand that we group those mollusks together in our taxonomy but it does not seem to demand that we assign them to any particular category. So I am inclined to think that only historically motivated taxonomies can justify species categories and that Ereshefsky is right to be conservative.

I shall return to this disconnect between taxonomies and categories in discussing the higher categories in the next section.

6 The Higher Categories

The issues that come up for species can come up for any of the higher categories ("ranks") in the Linnaean hierarchy, "the tree of life": genera, families, orders, classes, phyla, and kingdoms. Thus, consider the kinds we call "genera." At the first level, there is the Realist issue of whether *Canis* and other genera exist mind-independently. Then there is the further issue of whether it is explanatorily significant that they are *Canis*: the issue of whether *Canis* is a natural kind. At the second level, we are concerned with the analogous pair of issues about the *category* of being a genus; in particular,

with whether it is explanatorily significant that *Canis* and the like are *genera*, and hence with whether being a genus is a natural kind.

Let us start with the first level. No issue came up at this level for species and that is surely appropriate.¹³ We should all hold the Realist view that dogs and the like exist mind-independently. And being a dog is surely a natural kind if anything is: it features in historical explanations because being a dog is part of the evolutionary story; it features in structural explanations because a lot of the morphology, physiology, and behavior of Fido is explained by his being a dog.¹⁴ And one would have thought that there should similarly be no first-level issue with the higher taxa. Once again we should be Realist and we should take the higher taxa as natural kinds. Thus, whether or not being a genus is a natural kind surely being a Canis is: it features in historical evolutionary explanations and structural explanations, just as being a dog (Canis familiaris) does. And so too, say, being a mammal. This is not to say that membership in a higher taxon is as explanatorily significant as species membership: explanatory significance comes in degrees, as we noted (sec. 4). Still, surely some of the morphology, physiology, and behavior of Fido are explained by his being a Canis.

In light of this, one wonders what to make of the view that "taxa of higher rank than species do not exist in the same sense as do species" (Eldredge and Cracraft 1980, 327). According to Ereshefsky this sort of view is part of "The Modern Synthesis." He describes that Synthesis as holding: "Higher taxa . . . are merely artifacts of evolution at the species level. So while species are real and the 'units of evolution,' higher taxa are merely aggregates and 'historical entities'" (2001, 229). Ereshefsky himself rejects this view (1991, 381), as does James Mallet: "Whether species do have a greater 'objective reality' than lower or higher taxa is either wrong or at least debatable" (1995, 296). Clearly I think that Ereshefsky and Mallet are right in their rejection. I suspect that the rejected view arises out of a confusion of a Realism issue with a naturalness issue and/or a confusion of a first-level issue with a second-level one.

Another first-level issue has arisen from the fierce controversy between those favoring "phenetic," "evolutionary," and "cladistic" taxonomies. Sterelny and Griffiths favor the cladistic taxonomy and take it to be the emerging consensus (1999, 194). Yet cladistics involves the following surprising "metaphysical" first-level claim:

real groups in nature are all, and only, *monophyletic* groups . . . groups that consist of a species and all, and only, its descendents. To the cladist true believer, there is no such thing as a reptile. 'Reptile' does not name a real group, for there is no species

that is ancestral to *all* the reptiles that is not also ancestor of the birds. (Sterelny and Griffiths 1999, 197)

Now, we should note first of all that the claim that "there is no such thing as a reptile" is a most unhappy way of putting the cladistic point. *Of course* there are reptiles because there are crocodiles, snakes, and lizards and to be a reptile is simply to be one of those. Reptiles exist, and do so mind-independently, just as much as cousins, step-cousins, and grugus. And it is not much more apt to say that reptiles are not a "real group" in nature: there is nothing interestingly *unreal* about crocodiles, snakes, and lizards. I take it that the best way to put the cladistic point is to say that being a reptile is *not an explanatory significant kind*. In general, the cladist thinks that non-monophyletic groups are not natural kinds: once again, the non-natural is being confusingly described as the non-real.

Is this restriction to monophyletic groups appropriate? Although Sterelny and Griffiths favor cladism, they are inclined to think that the restriction is too extreme:

there may well be sensible evolutionary hypotheses about all the nonmarine mammals. The group is not a monophyletic clade, because there is no species ancestral to all the land-breeding mammals that is not also ancestral to the whales. (1999, 198)¹⁵

This is an evolutionary reason for abandoning strict monophyly. And there are other explanatory concerns in biology that seem to demand this as well: the sorts of structural concerns that partly motivate the phenetic and evolutionary taxonomies. There seems no reason to think that only monophyletic classifications can serve those other concerns. Perhaps being a "reptile" is explanatorily significant in many cases.¹⁶ Doubtless *some* non-monophyletic groups that biologists have posited do not "carve nature at its joints," but the cladists case that *none* do seems inadequate at best, reflecting attention only to historical explanations.¹⁷

Turn now to the second level. We might expect taxonomic disputes to play a role here in two ways. (A) The pluralism about the species category that we have been discussing may have repercussions for higher categories (ranks). (B) There is the possibility that the just-mentioned dispute between phenetic, evolutionary, and cladistic classifications is relevant.

I shall be very brief about (A). If anything like Kitcher's or Ereshefsky's species pluralism is right then it seems that we would have to have a distinct tree of life for each type of species; the species of each type are the base taxa for a distinct tree. If so then the *best* we could hope for with the higher categories is that each type of a category—for example, each type

of genus based on a type of species—is a natural kind. But turning to (B), we find that this hope seems not to be realized.

Consider (B). In discussing Kitcher's pluralistic realism, I noted that if the taxa picked out by a certain taxonomy are to justify a certain species category, S, it has to be explanatorily significant that the taxa are S. And using Kitcher's mollusk example, it is hard to see how it makes any difference to structural explanations whether a group is a species, subspecies, genus or whatever. This carries over to the higher categories: taxonomies may not yield explanatory categories in general. So even if we could establish that a certain taxonomy-phenetic, evolutionary, cladistic or whatever-was theoretically sound, so that classifying organisms in that way served our explanatory purposes in biology, that alone would not justify the view that any category posited by the taxonomy was a natural kind. Just because a taxonomy is right to classify a group of organisms as Canis and a subgroup as Canis familiaris does not show that there is any explanatory significance in treating the former as a genus and the latter as a species. We still need to show that the category itself does explanatory work.

Do the higher categories do any explanatory work? Cladists think not:

taxonomic ranks make little sense. . . . they do not think there will be any robust answer to the questions when should we call a monophyletic group of species a genus? a family? an order? Only monophyletic groups should be called anything, for only they are well-defined chunks of the tree. But only silence greets the question are the chimps plus humans a genus? (Sterelny and Griffiths 1999, 201)

So, on the cladistic picture, all categories (ranks) above species must be abandoned. Ereshefsky agrees although, as we have noted, he abandons the species category as well. He rightly points out that if a certain category is to be acceptable, the taxa of that category must be "comparable" and draws attention to reasons for thinking that this condition is not met (1999, 299). Brent Mishler claims that "practicing systematists know that groups given the same rank across biology are not comparable in any way" (1999, 310–311). In a lengthier critique of the Linnaean hierarchy, Ereshefsky mentions the drive to introduce more ranks, leading to a hierarchy in flux and to disagreements about the rank of certain taxa (2001, 215, 226). The signs are that, although the higher categories may have some pragmatic value, they are doing no explanatory work: they are not natural kinds.

In sum, at the first level I have found nothing against the view that the higher taxa, even some that are not monophyletic, are natural kinds. The

second level is different: the signs are that the higher categories are not natural kinds and so the Linnaean hierarchy must be abandoned. Finally, we should note that abandoning the Linnaean hierarchy is not abandoning a hierarchy altogether, it is not abandoning a tree of life. It is abandoning the labeling of categorical ranks in that tree (Ereshefsky 1999, 299; Mishler 1999, 311). But in light of (A) and (B), we may have to accept that there is more than one correct *uncategorized* tree of life, each reflecting legitimate explanatory concerns.

7 Conclusions

My aim in this essay has been to examine certain issues in biology to show that they are really issues over which kinds are causal-explanatory and hence "carve nature at its joints." The issues are over which kinds are "natural," not over the mind-independent existence of anything. So describing the issues as being about "realism"—as about what "exists," is "real," or is "objective"—has led to unclarity and confusion. Throughout my discussion I have emphasized the importance of distinguishing between first-level issues about *taxa* and second-level issues about *categories*.

Species pluralism is the view that there are several equally good accounts of what it is to be a species. Ereshefsky presents his argument from species pluralism to antirealism as an argument against the existence of the species category. I have argued that it is better seen as an argument against the explanatory significance of that category, hence an argument against that category being a natural kind. Not surprisingly, Ereshefsky sees his "pluralistic antirealism" as opposed to Kitcher's "pluralistic realism." Yet on close inspection, the two positions are similar: they agree that the species category itself is not explanatory but that various types of that category are explanatory. However, they differ in that Kitcher's pluralism is more radical: Kitcher thinks that structural explanations in biology justify some species categories whereas Ereshefsky thinks that only historical evolutionary explanations can do so. I presented an argument that Ereshefsky is right: even if structural explanations motivate taxonomies they do not seem to show that a species category plays an explanatory role.

Finally, I considered antirealism about the higher categories. At the first level, despite the urging of cladists, there seems to be no good reason to suppose that only monophyletic higher taxa are natural kinds. However, at the second level, the signs are that the higher categories are not natural kinds: they seem to do no explanatory work. If they don't, then the

Linnaean hierarchy must be abandoned. Furthermore, the case for various taxonomies suggests that we may have to accept more than one uncategorized tree of life.

Notes

1. This essay draws heavily on Devitt 2009.

2. Elliott Sober (1980, 203) struggles with what Ernst Mayr means by *real*. David Hull draws attention to the unclarity of these "realism" issues (1999, 25–26).

3. Although, interestingly enough, an issue that Darwin himself was skeptical about: he talks of "the vain search for the undiscovered and undiscoverable essence of the term species" (1859/2004, 381).

4. Mayr's distinction is established but often overlooked—for example, in Dupré 1981; Griffiths 1999; and Sterelny 1999. In his brief discussion of Stanford and Ereshefsky on realism, Kevin De Queiroz (1999, 74–75) does distinguish the taxa issue from the category issue but he does so against a background of a definition of realism that applies only to taxa. That definition also takes realism to be a commitment only to mind-independent existence, ignoring the far greater importance of explanatory significance in this debate.

5. As Hull emphasizes, many biologists require that a species concept be not only theoretically explanatory but also easily applicable: it must be "operational." Hull rightly points out that "the philosophical arguments against operationism are decisive," but then goes on to be surprisingly tolerant of an "operational criteria for theoretical concepts" (1997, 371).

6. And even more from commitment on the equally vexed issue of the nature of properties (supposing there are any).

7. We should acknowledge the popular view that species are *individuals* rather than kinds (Ghiselin 1974; Hull 1978). On that view, the species of tigers is an individual constituted by the fusion of all tigers and any particular tiger is a *part of* that individual rather than a member of the tiger kind. Adjusting my discussion of biological realisms to take account of this view would have no significant effect, so I shall continue to write as if species are kinds.

8. Richard Boyd claims that the types of species are unified by "an especially close homeostatic relation between the classificatory practices" (1999, 171); see also Wilson 1999, 203–204.

9. In another paper, Ereshefsky gives a related reason for doubting the existence of the species category: the failure of attempts to distinguish species from the higher taxa, in particular the failure to do so in terms of the processes of speciation and interbreeding (1999, 269).
10. Thus, David Armstrong (1978) is a selective realist, holding that empty predicates, negative predicates, and most pertinently, disjunctive predicates have no corresponding universal. He thinks that some predicates apply to the world in virtue of many universals. Most importantly, he looks to science to tell us which properties there are.

11. It is easy even for staunch realists to slip into loose ways of talking that suggest world-making. Thus Kornblith says that when we "group objects together under a single heading on the basis of a number of easily observable characteristics . . . we thereby create a nominal kind" (1993, 41). But we don't! We create a *concept* that picks out a kind that may or may not be "real" in Locke's terms, but which has its members independently of our creation. And Boyd, talking of kinds with nominal essences, says that their "boundaries" are "purely matters of convention" (1999, 142). But they aren't! Our naming a kind picked out by a certain set of descriptions is conventional, but the boundary of the kind thus picked out is not.

12. Despite this, I do have my doubts that *any* species category is really explanatorily significant.

13. Stanford (1995) seems to disagree; see Ereshefsky 1998 and Devitt 2009 for criticisms.

14. It is perhaps worth mentioning that being a member of a certain subspecies or variety is also explanatorily significant: Fido's being a pit bull explains a lot about his morphology, physiology, and behavior. Cf. Joel Cracraft's description of concern about "the ontological status of subspecies" (1983, 100).

15. Boyd is also skeptical of monophyly (1999, 182).

16. It is also surely explanatorily significant that something is a predator or a parasite. Consider this, for example: "The Lotke-Volterra equations . . . describe the interactions of predator and prey populations" (Sober 1980, 202). But these are not the sort of classifications that concern the cladists. (Thanks to Marc Ereshefsky.)

17. Mishler concludes his summary of the argument for monophyly with this remarkably inadequate claim: "Because the most effective and natural classification systems are those that 'capture' the entities resulting from processes that generate the things being classified, the general biological classification system should be used to reflect the tree of life" (1999, 309–310). It is probably the case that the classification systems in *all* sciences "capture" entities resulting from processes—entities don't come from nothing!—but it doesn't follow that they should be classified to reflect those processes.

References

Armstrong, D. M. 1978. *A Theory of Universals: Universals and Scientific Realism*, vol. II. New York: Cambridge University Press.

Boyd, R. 1999. Homeostasis, species, and higher taxa. In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge, Mass.: MIT Press.

Claridge, M. F., H. A. Dawah, and M. R. Wilson, eds. 1997. *The Units of Biodiversity*. London: Chapman and Hall.

Cracraft, J. 1983. Species concepts and speciation analysis. In *Current Ornithology*, vol. 1, ed. R. Johnston. New York: Plenum Press. Reprinted in Ereshefsky 1992; citations are to Ereshefsky.

Darwin, C. 1859/2004. *The Origin of Species by Means of Natural Selection*. New York: Barnes & Noble.

De Queiroz, K. 1999. The general lineage concept of species and the defining properties of the species category. In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge, Mass.: MIT Press.

Devitt, M. 1980. "Ostrich nominalism" or "Mirage realism"? *Pacific Philosophical Quarterly* 61: 433–439. Reprinted in *Properties*, ed. D. H. Mellor and A. Oliver. Oxford: Oxford University Press, 1997. Reprinted in Devitt 2010.

Devitt, M. 1997. *Realism and Truth*, 2nd ed. with Afterword. (1st ed. 1984, 2nd ed. 1991). Princeton: Princeton University Press.

Devitt, M. 1999. A naturalistic defense of realism. In *Metaphysics: Contemporary Readings*, ed. Steven D. Hales. Belmont, Cal.: Wadsworth.

Devitt, M. 2001. Incommensurability and the priority of metaphysics. In *Incommensurability and Related Matters*, ed. P. Hoyningen-Huene and H. Sankey. Dordrecht: Kluwer Academic. Reprinted in Devitt 2010.

Devitt, M. 2002. Underdetermination and realism. Philosophical Issues 12:26-50.

Devitt, M. 2005. Scientific realism. In *The Oxford Handbook of Contemporary Philosophy*, ed. F. Jackson and M. Smith. Oxford: Oxford University Press. Reprinted in *Truth and Realism*, eds. P. Greenough and M. Lynch. Oxford: Oxford University Press. Reprinted in Devitt 2010.

Devitt, M. 2008. Resurrecting biological essentialism. *Philosophy of Science* 75: 344–382. Reprinted in Devitt 2010.

Devitt, M. 2009. Biological realisms. In *From Truth to Reality: New Essays in Logic and Metaphysics*, ed. H. Dyke. London: Routledge & Kegan Paul.

Devitt, M. 2010. *Putting Metaphysics First: Essays on Metaphysics and Epistemology*. Oxford: Oxford University Press.

Devitt, M., and K. Sterelny. 1999. *Language and Reality: An Introduction to the Philoso-phy of Language*, 2nd ed. (1st ed. 1987). Cambridge, Mass.: MIT Press.

Dupré, J. 1981. Natural kinds and biological taxa. Philosophical Review 90:66–90.

Dupré, J. 1993. The Disorder of Things. Cambridge, Mass.: Harvard University Press.

Dupré, J. 1999. On the impossibility of a monistic account of species. In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge, Mass.: MIT Press.

Eldredge, N., and J. Cracraft. 1980. *Phylogenetic Patterns and the Evolutionary Process*. New York: Columbia University Press.

Ereshefsky, M. 1991. Species, higher taxa, and the units of evolution. *Philosophy of Science* 58: 84–101. Reprinted in Ereshefsky 1992; citations are to Ereshefsky 1992.

Ereshefsky, M., ed. 1992. *The Units of Evolution: Essays on the Nature of Species*. Cambridge, Mass.: MIT Press.

Ereshefsky, M. 1998. Species pluralism and anti-realism. *Philosophy of Science* 65:103–120.

Ereshefsky, M. 1999. Species and the Linnaean hierarchy. In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson, 285–305. Cambridge, Mass.: MIT Press.

Ereshefsky, M. 2001. *The Poverty of the Linnaean Hierarchy: A Philosophical Study of Biological Taxonomy*. Cambridge: Cambridge University Press.

Ghiselin, M. T. 1974. A radical solution to the species problem. *Systematic Zoology* 47: 350–383. Reprinted in Ereshefsky 1992; citations are to Ereshefsky 1992.

Griffiths, P. 1999. Squaring the circle: Natural kinds with historical essences. In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge, Mass.: MIT Press.

Hull, D. L. 1978. A matter of individuality. *Philosophy of Science* 45: 335–360. Reprinted in Ereshefsky 1992; citations are to Ereshefsky 1992.

Hull, D. L. 1997. The ideal species concept—and why we can't get it. In *The Units of Biodiversity*, ed. M. F. Claridge, H. A. Dawah, and M. R. Wilson. London: Chapman and Hall.

Hull, D. L. 1999. On the plurality of species: Questioning the party line. In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge, Mass.: MIT Press.

Kitcher, P. 1984. Species. *Philosophy of Science* 51:308–333. Reprinted in Kitcher 2003 and Ereshefsky 1992a; citations are to Kitcher 2003.

Kitcher, P. 2003. In Mendel's Mirror: Philosophical Reflections on Biology. New York: Oxford University Press.

Kornblith, H. 1993. *Inductive Inference and Its Inductive Ground: An Essay in Naturalistic Epistemology*. Cambridge, Mass.: MIT Press.

Kuhn, T. 1970. *The Structure of Scientific Revolutions*, 2nd ed. (1st ed. 1962). Chicago: University of Chicago Press.

Mallet, J. 1995. A species definition for the modern synthesis. *Trends in Ecology & Evolution* 10:294–299.

Mayr, E. 1969. Principles of Systematic Zoology. New York: McGraw Hill.

Mayr, E. 1982. *The Growth of Biological Thought*. Cambridge, Mass.: Harvard University Press.

Mishler, B. 1999. Getting rid of species. In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge, Mass.: MIT Press.

O'Hara, R. J. 1993. Systematic generalization, historical fate, and the species problem. *Systematic Biology* 42:231–246.

Okasha, S. 2002. Darwinian metaphysics: Species and the question of essentialism. *Synthèse* 131:191–213.

Sober, E. 1980. Evolution, population thinking, and essentialism. *Philosophy of Science* 47:350–383. Reprinted in Sober 1994 and Ereshefsky 1992; citations are to Sober 1994.

Sober, E., ed. 1994. *Conceptual Issues in Evolutionary Biology*, 2nd ed. Cambridge, Mass.: MIT Press.

Sokal, R., and T. Crovello. 1970. The biological species concept: A critical evaluation. *American Naturalist* 104:127–153. Reprinted in Ereshefsky 1992; citations are to Ereshefsky 1992.

Stanford, P. K. 1995. For pluralism and against realism about species. *Philosophy of Science* 62:70–91.

Sterelny, K. 1999. Species as ecological mosaics. In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge, Mass.: MIT Press.

Sterelny, K., and P. Griffiths. 1999. Sex and Death. Chicago: University of Chicago Press.

van Valen, L. 1976. Ecological species, multi-species, and oaks. *Taxon* 25:233–239. Reprinted in Ereshefsky 1992; citations are to Ereshefsky 1992.

Wilson, R. A. 1999. Realism, essence, and kind: Resuscitating species essentialism? In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge, Mass.: MIT Press.

9 Three Ways of Resisting Essentialism about Natural Kinds

Bence Nanay

1 Essentialism about Natural Kinds

Essentialism about natural kinds has three tenets. The first tenet is that all and only members of a natural kind have some essential properties. The second tenet is that these essential properties play a causal role. The third tenet is that they are explanatorily relevant. I examine the prospects of questioning these tenets and point out that arguing against the first and the second tenets of kind-essentialism would involve taking part in some of the grand debates of philosophy. But, at least if we restrict the scope of the discussion to the biological domain, the third tenet of kindessentialism could be questioned more successfully.

It is not an easy task to pin down what is meant by essentialism about natural kinds (Putnam 1975; Kripke 1980). First, one can be essentialist about individuals and about kinds. I will not say anything here about essentialism regarding individuals. Maybe, as Kripke claims, specific individuals have essential properties, maybe not (on this important and complex question, see, for example, Robertson 1998; Hawthorne and Gendler 2000; Matthen 2003). Essentialism about individuals is logically independent from essentialism about kinds (see also Okasha 2002, 192). The question I am interested in is whether natural kinds have essential properties.

Second, there are a number of potential definitions for essentialism about kinds. As I intend to argue against essentialism, I will use the most general of these. Richard Boyd identified a widespread and fairly strong version of essentialism, according to which natural kinds "must possess definitional essences that define them in terms of necessary and sufficient, intrinsic, unchanging, ahistorical properties" (Boyd 1999, 146). Essential properties in, say, chemistry may all be intrinsic, unchanging, and ahistorical. But it is not clear that all essential properties need to satisfy any of these three requirements. In fact, a rather easy way of arguing against essentialism about at least some natural kinds—namely, biological kinds—is to point out that biological properties are extrinsic, historical, and change over time since biological entities are evolving over time. But it is unlikely that arguments of this kind will defeat essentialism about biological, natural kinds. A new wave of biological essentialists all seek to specify essential properties of biological kinds that are extrinsic, and yet are neither unchanging nor ahistorical.¹ The simple argument from the observation that biological entities are evolving over time cannot be used to argue against these versions of biological essentialism.

Thus, if we want a target that is worth arguing against, we need to weaken this strong definition of essentialism. As most of the new essentialists, I am also happy to go along with David Hull's characterization, according to which "each species is distinguished by one set of essential characteristics. The possession of each essential character is necessary for membership in the species, and the possession of all the essential characterist (1994, 313). I will use Hull's definition as my starting point for characterizing kind-essentialism in what follows.²

Third, essentialism about kinds is a complex thesis that goes beyond the simple claim that there are some properties that all and only members of a natural kind have in all possible worlds. Marc Ereshefsky specified three tenets of any version of essentialism about kinds:

One tenet is that all and only the members of a kind have a common essence. A second tenet is that the essence of a kind is responsible for the traits typically associated with the members of that kind. For example, gold's atomic structure is responsible for gold's disposition to melt at certain temperatures. Third, knowing a kind's essence helps us explain and predict those properties typically associated with a kind. (Ereshefsky 2007, sec. 2.1)

Most philosophers who, like Ereshefsky, argue against essentialism, only consider the first tenet. Proponents of essentialism also tend to be concerned only with this first tenet. In contrast, I would like to focus on the second and especially the third tenet.

My claim is that questioning the second or third tenets may be a more promising way of resisting kind-essentialism. In short, a promising and so far almost completely unexplored anti-essentialist strategy would be to say that even if it turns out that "all and only the members of a kind have a common essence," this essence is unlikely to play any significant causal or explanatory role. The plan for the essay is simple: I go through the three tenets of essentialism and explore which of them would be the easiest to question. My final response will be that arguing against the first and the second tenets of kind-essentialism would involve taking part in some of the grand debates of philosophy. But at least if we restrict the scope of the discussion to the biological domain, the third tenet of kind-essentialism could be questioned more successfully.

2 Questioning the First Tenet: All and Only the Members of a Kind Have a Common Essence

I am not sure that the question whether "all and only the members of a kind have a common essence" can be settled. A rather straightforward way of arguing against the first tenet would be to use some general metaphysical considerations. The first tenet of kind-essentialism states that all and only members of a kind have a certain essential property: an essence. The crucial point is that 'property' here means "property-type." The set of essential properties that defines natural kinds is a set of essential property-types. The *instantiation* of each essential property-type is necessary for membership in the natural kind and the *instantiation* of all the essential property-types and property-instances explicit.

The term 'property' is ambiguous. It can mean universals: properties that can be present in two (or more) distinct individuals at the same time. But it can also mean tropes: abstract particulars that are logically incapable of being present in two (or more) distinct individuals at the same time (Williams 1953; Campbell 1981, 1990; Schaffer 2001; Simons 1994; Sanford manuscript).

Suppose that the color of my neighbor's black car and my black car are indistinguishable. They still have different tropes. The blackness trope of my car is different from the blackness trope of my neighbor's car. These two tropes are similar but numerically distinct. Thus, the blackness of my car and the blackness of my neighbor's car are different properties.

If, in contrast, we interpret properties as universals—or as I will refer to them, as property-types—then the two cars instantiate the very same property-type: blackness. Thus, depending on which notion of property we talk about, we have to give different answers to the question about whether the color-property of the two cars is the same or different. If by 'property' we mean "trope," then my car has a different (but similar) color-property—that is, color-trope—from my neighbor's. If, however, by 'property' we mean "property-type," then my car has the very same property—that is, property-type—as my neighbor's.

In the light of this distinction, the first tenet of essentialism about natural kinds can be broken down to the conjunction of two claims: (a) property-types exist, and (b) some property-types are essential property-types. Many anti-essentialist arguments question (b) while accepting (a) (e.g., Hull 1986, as far as the biological domain is concerned). I will focus on the more radical strategy of questioning (a) because, in sections 5 and 6, I will argue that (at least in the domain of biology) a strategy quite similar to this may be used in order to argue not against the first, but against the third tenet of essentialism about natural kinds.

As the first tenet of kind-essentialism states that all and only members of a kind have a certain essential property-type, a straightforward way of arguing against this tenet is to question (a): to show that property-types do not exist and hence, *a fortiori*, essential property-types do not exist either. In other words, if we accept a version of nominalism, then there is a simple way of resisting any version of essentialism about natural kinds. If only particulars exist and property-types do not, then how could we even formulate essentialism about natural kinds?

This would be a simple and straightforward argument against essentialism about natural kinds, but it is not clear that we have any reason to accept its main premise: that property-types do not exist. The grand debate between nominalism and realism is one of the oldest in philosophy and it has definitely not been resolved. Taking for granted the premise that there are no property-types would significantly weaken an anti-essentialist argument in the eyes of those who are not fully convinced by nominalist considerations.

Further, we need to be a bit more careful about what version of trope nominalism the anti-essentialist strategy I outlined above needs to endorse. Many trope-nominalist accounts define property-types as sets or resemblance-classes of tropes. Thus, according to these accounts, although the existence of property-types in some sense reduces to the existence of tropes (and, as a result, in some sense they are not "real"), they do have mindindependent existence: there is a fact of the matter about whether a trope subsumes under a certain set or resemblance-class.

The version of trope nominalism that would be needed to question (a) needs to be more radical than this: it cannot allow for there being a fact of the matter about whether a trope subsumes under a certain set or resemblance-class. If it did allow for this, then the essentialist view could be rephrased in terms of "essential" sets or "essential" resemblance-classes

of tropes. In order to block (a), the trope nominalist needs to claim that property-types are just our ways of grouping tropes: they do not have mind-independent existence (Nanay 2009, 2010a, 2011). But this is an even more controversial assumption for an anti-essentialist argument to rely on.

But even if such a version of trope nominalism is a dubious premise to build our argument on, perhaps if we restrict the scope of our argument to the biological domain then the first tenet could be questioned successfully.

3 Questioning the First Tenet of Kind-Essentialism in the Biological Domain

Biology has always been considered to be a problem case for essentialism or at least a potential exemption. According to the traditional "antiessentialist consensus" (Okasha 2002, 195; Walsh 2006, 325) among biologists and philosophers of biology, at least regarding biological kinds, essentialism is false (Dupré 1993, 2002; Hull 1965; Ghiselin 1974; Hacking 2007). Putnam and Kripke may be right about chemical kinds, but biological kinds do not have (and cannot have) any essential properties (Wilkerson 1995; Ellis 2001).

But over the last several years, more and more philosophers have argued for a version of essentialism about biological kinds. Paul Griffiths, for example, argues that biological kinds have "essential relational properties" not essential intrinsic properties—and claims that if we accept that essential properties can be relational, then all the traditional considerations against essentialism about biological kinds lose their appeal (Griffiths 1999). (For a similar claim, see Okasha 2002. The idea of using relational properties for defining biological kinds, not necessarily in an essentialist manner, comes from Matthen 1998; Millikan 1999; and Elder 1995.) Denis Walsh goes even further and claims that "recent evolutionary developmental biology provides compelling evidence" for essentialism (2006, 425).

In order to assess the merits of this new wave of essentialism about biological kinds, the traditional anti-essentialist arguments need to be reevaluated. There are anti-essentialist arguments that prove to be inconclusive. One such argument concerning essentialism about biological kinds, is the following (Hull 1965; for objections, see Sober 1980, 356; Okasha 2002, 195–196; Walsh 2006, 431). According to evolutionary theory, the present species have evolved from ancestral ones. Thus, species cannot have essences, as they are clearly capable of changing. This

argument may be taken to jeopardize a version of essentialism, one that takes kinds to be unchangeable, but it does not apply in the case of other versions of essentialism.

Some other anti-essentialist considerations are not obviously inconclusive. The most important of these are based on the concept of "population thinking." Since using population thinking to argue against essentialism shares some important features with my own argument against the third tenet of essentialism in sections 5 and 6, I will spend some time trying to understand what population thinking is supposed to mean, and why it is assumed to be an effective weapon in the fight against essentialism.

Population thinking has traditionally been *the* main consideration against essentialism about biological kinds. The strategy of philosophers with anti-essentialist convictions about biology has been to point out that since population thinking implies anti-essentialism and population thinking is the right way of thinking about biology, we have to be anti-essentialist about biology.

This strategy can be attacked at two points. First, the essentialist could argue that population thinking is not the right attitude to take toward the biological domain. No philosophers or biologists seem to take this route. The second way of attacking the anti-essentialist strategy would be to deny that population thinking implies anti-essentialism. And in fact, with the rise of contemporary attempts to bring essentialism back in the domain of biology, it has been repeatedly argued that population thinking does not exclude essentialism. Thus, one can endorse population thinking and still agree with Putnam and Kripke about the essential properties of biological kinds.

All of the recent attempts to resurrect essentialism about biological kinds find it important to show that their version of essentialism is consistent with population thinking (Walsh 2006, 432–433; Okasha 2002, 195–196). Paul Griffiths says explicitly that "it would be quite consistent to be a Darwinian [population-thinking] essentialist, given the right choice of essential properties" (Griffiths 1999, 210). Or, more explicitly: "Population thinking excludes essential intrinsic properties, but it does not exclude essential relational properties" (ibid.; cf. Okasha 2002). Whether these attempts to carve out an essentialist way of construing population thinking succeed depends on the way we interpret population thinking. Conversely, whether population thinking really gives us some reason to have doubts about essentialism about biological kinds also depends on the way we interpret population thinking in

their arguments, it is important to examine what population thinking is and what it implies.

My claim is that Ernst Mayr's influential idea of what makes the biological domain special, the idea of "population thinking," could, and should, be interpreted as a version of trope nominalism (Nanay 2010a). Here is Mayr's characterization of population thinking from 1959: "Individuals, or any kind of organic entities, form populations of which we can determine only the arithmetic mean and the statistics of variation. Averages are merely statistical abstractions; only the individuals of which the populations are composed have reality" (Mayr 1959, 326). Mayr contrasts population thinking with typological thinking, according to which "there are a limited number of fixed, unchangeable 'ideas' underlying the observed variability, with the *eidos* (idea) being the only thing that is fixed and real, while the observed variability has no . . . reality" (ibid.). The contrast Mayr makes is a very sharp one: population thinking and typological thinking are exclusive of each other (ibid., 326–327).

Mayr's distinction between typological and population thinking may appear straightforward, but in fact it could be (and has been) interpreted in at least two ways. First, population thinking could be interpreted as an ontological claim about *entities*: only the individual is real, everything else is abstraction. There are various problems with this reading. If only the individual is real, then populations and species should be thought of as groups of individuals which, *as* groups, lack reality themselves. This would make much of post-Darwinian biology nonsensical from the population thinker's point of view. As Elliott Sober says:

If [as Mayr claims] "only the individuals of which the populations are composed have reality," it would appear that much of population biology has its head in the clouds. The Lotka-Volterra equations, for example, describe the interactions of predator and prey *populations*. Presumably, population thinking, properly so called, must allow that there is something real over and above individual organisms. [It does not] embody a resolute and ontologically austere focus on individual organisms alone. (Sober 1980, 352)

Even more problematic for this reading is the fact that Mayr himself is certainly not nominalist about populations and species (Mayr 1942, 120; Mayr 1963, 19). His dictum that "only the individuals . . . have reality" seems to flatly contradict his famous "biological species concept" which indeed attributes reality to populations and species. It is tempting to resolve this seeming contradiction by dismissing Mayr's claim about the importance of the individual in evolution as an exaggeration or even as "rather silly metaphysics" (Ariew 2008, 2).³

Elliott Sober chooses this route when he says that "describing a single individual is as theoretically peripheral to a populationist as describing the motion of a single molecule is to the kinetic theory of gases. In this important sense, population thinking involves *ignoring individuals*" (Sober 1980, 370). The conclusion he draws is that "population thinking endows individual organisms with more reality *and* with less reality than typological thinking attributes to them" (Sober 1980, 371).

This conclusion prompted some to be "a little confused about which one, individuals or populations, are real" (Ariew 2008, 8). It also opened up the concept of population thinking to many diverging interpretations, some of which seems to contradict Mayr's original claims (Walsh 2006, 432–433; Griffiths 1999, 209–210).

I argue that population thinking is an ontological claim about *properties* and not about entities. It is indeed a version of nominalism. However, it is not nominalism about entities, but about properties. In other words, Mayr advocated a version of trope nominalism: for the population thinker, only the property-instances, that is, *tropes*, are real. Property-*types* are not real.

We have to be careful when formulating this claim. The population thinker presumably would not deny that groups of individual organisms do have properties and that these properties are real. A population of 431 geese has the property of having the population size of 431, for example, and this property seems very real indeed. The distinction I am making (and the distinction I believe Mayr was making) is not one between the properties of individuals and the properties of populations. Rather, it is between individual property-instances (or tropes) and property-types (or universals) that can be instantiated in many different entities. In short, the population thinker can acknowledge the existence of populations and species. These entities are real in the same way as individuals are real. And all of these entities have very real property-instances or tropes. What the population thinker denies is that there are property-types.

My claim is that Mayr's provocative statement, according to which "averages are merely statistical abstractions; only the individuals of which the populations are composed have reality" should be read as "*property-types* are merely statistical abstractions; only the *tropes* of individuals (or of populations) have reality." Mayr's population thinking is a version of trope nominalism (see esp. Mayr 1959, 326, where he talks about the uniqueness of features, i.e., properties, and not the uniqueness of individual entities).⁴

We can now put together an argument against the first tenet of kindessentialism in the biological domain: (i) Population thinking is the right way of thinking about the biological domain. (ii) Population thinking implies trope nominalism. (iii) Trope nominalism implies anti-essentialism about biological kinds. I have presented some considerations in favor of (ii) and (iii), but as yet have said nothing that would make us accept (i): that population thinking, that is, trope nominalism, would be the right way of thinking about the biological domain. And I am not sure what argument could be given in favor of (i), besides appealing to the authority of Ernst Mayr—without this premise, the argument collapses.

In section 6, I will give an argument for the claim that biological property-types play no explanatory role in evolutionary explanations. Could we use this argument to establish (i)? Mayr's claim is much stronger than mine, as he denies the reality (not merely the explanatory relevance) of biological property-types: he claims that they are merely our statistical abstractions.

Yet, depending on one's meta-metaphysical convictions, there may not be such a huge difference between these two versions of trope nominalism: Mayr's stronger version and the weaker "explanatory trope nominalism" I will argue for below. One could, after all, use the weaker claim that biological property-types are explanatorily superfluous and, with the help of the principle of parsimony, conclude that we have no reason to postulate their existence.

But not everyone will find this last step unproblematic, and I do not want to argue that it is unproblematic. If someone believes that we can infer from the fact that something is explanatorily superfluous that it does not exist (as Mayr may have believed), then she will not find the distinction between my "explanatory trope nominalism" and Mayr's population thinking a very interesting one. She will probably not find the distinction between the first and the third tenet of kind-essentialism a meaningful one either. But if someone does not believe that explanatory irrelevance implies non-existence, then she will still have no reason to accept (i) and hence to accept the population-thinking inspired rejection of the first tenet of essentialism about biological kinds.

The argument I considered in this section was unsuccessful in the end. But I will use its conceptual framework in the hopefully more successful argument against the third tenet of kind-essentialism, in sections 5 and 6.

4 Questioning the Second Tenet: Causal Responsibility

The second tenet of kind-essentialism is about causal responsibility. As Ereshefsky says, "a kind's essence causes the other properties associated with that kind. The essence of the natural kind gold, for example, is gold's

atomic structure... the atomic structure of gold causes pieces of gold to have the properties associated with that kind, such as dissolving in certain acids and conducting electricity" (forthcoming, 1).

What we say about this second tenet depends on what we say about the *relata* and properties of (singular) causation. If one holds that propertytypes play a role in (singular) causation, then the second tenet of essentialism remains unscratched. But if one holds that property-types play no causal role—maybe because no properties play any causal role (Davidson 1967, 1970), or maybe because only property-instances (or tropes) play any causal role (Ehring 1997; Nanay 2009)—then we have reason to doubt the second tenet. But let us go through these considerations more slowly.

The question is whether property-types play any role in causation. The first clarification we need to make when answering this question is whether it is about general or singular causation. Property-types play a clear and important role in *general* causal claims. If Fs cause Gs, then there is a property, in virtue of which Fs cause Gs: a property all Fs have in common (that is, a property-type all Fs have an instantiation of) and that, presumably, can account for why Fs cause Gs. But it is unclear whether property-types play a role in *singular* causation, and it could be argued that questions about causal relevance are questions about singular, not general, causation.

There are two ways in which property-types can play a causal role. First, they may be part of what specifies the *relata* of causation. It is not clear what the *relata* of causation are. They may be events (Davidson 1967), facts (Mellor 1995), states of affairs (Armstrong 1997) or maybe tropes (Ehring 1997). Facts (or Kimian events) are specified in terms of property-types. Thus, if we accept that these are the *relata* of causation, then property-types will be causally relevant. But if we hold that the *relata* of causation are Davidsonian events or tropes, then property-types play no role in the specification of the *relata* of causation. Thus, we have no *prima facie* reason to accept that they are causally relevant.

But even if we take (Davidsonian) events to be the *relata* of causation, a further question arises: what is it in virtue of which one event causes another? First, the obvious answer would be that an event causes another event in virtue of having an instantiation of a property-type: the sleeping pill I took last night made me fall asleep in virtue of some chemical property-type it had an instantiation of. If we accept this answer, then property-types will be very relevant causally.

But there are other ways of answering this "in virtue" question. Donald Davidson famously denied that events cause other events in virtue of *any*

properties. As he memorably said: "If causal relations and causal powers inhere in particular events and objects, then the way those events and objects are described, and the properties we happen to employ to pick them out or characterize them, cannot affect what they cause. Naming the American invasion of Panama 'Operation Just Cause' does not alter the consequences of the event" (Davidson 1993, 8). In other words, properties play no role in causal relations—although they may be very important in causal explanations. Another, less radical, way of denying that events cause other events in virtue of having instantiations of certain property-types would be to say that events cause other events in virtue of having tropes. Properties, if what we mean by that is tropes, *do* play an important role in causation: events cause other events in virtue of *them*. But property-types do not play any role (Robb 1997; Nanay 2009).

If we accept the Davidsonian or the trope answer to the "in virtue" question, then we can conclude that property-types are not causally relevant. But if property-types are not causally relevant, then essential property-types are not causally relevant either. But the second tenet of essentialism was that essences—that is, essential property-types—must be causally relevant. So, as in the case of the first tenet, we have a simple argument against kind-essentialism.

Note, however, that this argument rests on three heavily contested premises about the nature of causation; in order to run this argument, we need to make three important assumptions. First, we need to take singular, and not general, causation to be where causal relevance lies. Second, we need to take the *relata* of causation to be (Davidsonian) events (or tropes). And finally, we need to hold that events cause other events in virtue of having tropes (and not in virtue of having instantiations of propertytypes), or we need to endorse a Davidsonian view on the properties of causation. Few people hold all of these premises. Hence, as in the case of the argument against the first tenet of kind-essentialism, this argument is also based on premises concerning the nature of causal relations that many would question from the start.

5 Questioning the Third Tenet: Explanatory Relevance

The third tenet of kind-essentialism was the following: "Knowing the essence of a kind . . . allows us to predict and explain the properties associated with the members of a kind. For instance, the atomic structure of gold provides the basis for explaining why gold conducts electricity, and it allows us to predict that a particular chunk of gold will conduct electricity"

(Ereshefsky, forthcoming, 1). Or as Philip Kitcher puts it, "natural kinds are distinguished by some special underlying feature that explains the behavior of members of this kind—like atomic number, for example, in the case of the elements" (Kitcher 2007, 294; Dupré 2002, 176–181; Wilson et al. forthcoming). (See Platts 1983 for a classic summary and Okasha 2002, 203 for some critical remarks.) If we can show that this third tenet of essentialism is unjustified, then we have a good way of arguing against essentialism per se.

As in the case of the first and second tenets, I find it unlikely that we can give a general argument against the third tenet—any such argument would need to presuppose a rather specific and, as a result, heavily contested theory of explanation. But I do think that at least as far as the biological domain is concerned, we are in a good position to question the third tenet.

We have seen that it is unlikely that we would find some strong reason to reject the very idea that biological property-types are "real." But we may be able to find an argument for the claim that biological property-types play no explanatory role in biology. They may be "real" and exist independently of us, but if they are explanatorily superfluous, this is enough to undermine the third tenet of essentialism about biological kinds. I will give an argument in favor of this claim in the next section.

6 Questioning the Third Tenet of Kind-Essentialism in the Biological Domain

My claim is that property-types are explanatorily superfluous in evolutionary explanations. All the explanatory work is done by property-instances. This claim needs to be clarified and qualified at a number of points. First, I want to remain silent about whether property-types play any explanatory role in non-biological, non-evolutionary explanations. Maybe they do. When we are trying to explain why a certain gold sample melts at 1,948° F, we can explain this by referring to a property-type all gold samples have an instantiation of (maybe the property-type of having a certain atomic structure). In this explanation, we have a property-type as part of the *explanans*. The property-type may or may not be causally relevant, but it is explanatorily relevant. My claim is that this is not the case in evolutionary explanations, where the *explanans* refers only to property-instances.

Second, it is important to note that I do not claim that using tropes instead of property-types in the metaphysical framework *increases* the explanatory power of evolutionary theory. After all, it has been argued, convincingly, that statements about tropes and statements about instantiations of property-types are notional variants: one can always be rephrased in terms of the other (Daly 1997). All I claim is that adding biological property-types to a trope nominalist metaphysical framework *does not increase* the explanatory power of evolutionary theory. Hence, property-types are explanatorily superfluous.

I said that in the domain of biology, property-types do not do any explanatory work and property-instances do all the work. An important clarification about this claim: I talked about the biological domain, biological property-types and biological tropes. But it is not clear where the boundaries of the biological domain lie. Is DNA part of the biological domain or is it already part of the domain of chemistry? Also, there are many different kinds of explanation (Van Fraassen 1980). Saying that property-types play no role in any of them would be a difficult claim to argue for. So I will restrict the scope of my claim in the following manner: property-types play no role in evolutionary explanations. When I talk about the explanatory role (or lack thereof) biological property-types play, what I mean is explanatory role in an evolutionary explanation. Biological kinds are evolved kinds and biological entities are evolved entities. Thus, if a property-type is supposed to play some explanatory role in biology, like the atomic structure of gold explains why it melts at certain temperature, then, as the explanation of the properties of evolved entities is an evolutionary explanation, this means that this property-type is supposed to play at least some role in evolutionary explanations. I will attempt to show that this is not so: no property-type plays any role in evolutionary explanations.5

Let us go back to what the third tenet of essentialism entails. Ereshefsky's example is that "the atomic structure of gold provides the basis for explaining why gold conducts electricity, and it allows us to predict that a particular chunk of gold will conduct electricity" (Ereshefsky, forthcoming, p. 1). Philip Kitcher uses a similar example: "natural kinds are distinguished by some special underlying feature that explains the behavior of members of this kind—like atomic number, for example, in the case of the elements" (Kitcher 2007, p. 294). They both take their examples from chemistry. But what would be the equivalent of these claims in biology?

What is important from our point of view is that both Ereshefsky and Kitcher talks about the explanation of the behavior of a particular token member of the kind (or a token chunk of gold). Hence, in the domain of biology, an essential property would need to be able to explain why specific token organisms have the traits they do. And, luckily, a lot has been written about exactly this kind of explanation.

In fact, one of the most important recent debates in philosophy of biology is about whether natural selection can explain why specific organisms have the traits they have. The view that selection can play a role in explaining why organisms have the traits they have, has been defended by Karen Neander (1995a,b; see also Millikan 1990; Nanay 2005, 2010b; Matthen 1999). On the other side of the trench the central figure is Elliott Sober (1984a, 1995; see also Walsh 1998; Dretske 1988, 1990; Pust 2001; Lewens 2001; Cummins 1975; and Stegmann 2010).

Sober claims that selection is a negative force: it does not create; it only destroys (Sober 1984a, chapter 5). Random mutations create a variety of traits (or genetic plans) and selection eliminates some of these, but the explanation of the traits of one of these individuals is provided by random mutation and inheritance (and some developmental factors), not by the elimination process. Selection can explain why certain individuals were eliminated and it may also explain why a trait is present (or widespread) in a population, but it cannot explain the traits of specific individuals that were *not* eliminated.

Karen Neander argues against the validity of this argument, at least as far as cumulative selection is concerned (Neander 1995a). After a couple of rounds of exchanges without any sign of rapprochement, one gets the sense that there is some sort of miscommunication between Neander and Sober. One gets the sense that the opponents and the advocates of this argument may not mean the same by the term 'selection'.

My aim here is not to decide who is right in this debate (I attempted to do this in Nanay 2005 and Nanay 2010b). My aim is to show that regardless of which of these two views about the explanatory power of selection we accept, we can conclude that property-types do not play any role in explaining why token organisms have the token traits they have.

Take Sober's position first. He has argued repeatedly that the theory of natural selection can only be formulated with the help of property-types, that is, property types play a very important role in the theory of natural selection (most explicitly in Sober 1981, but also in Sober 1980 and Sober 1984). As he says in a paper co-authored with Richard Lewontin: "selection theory is about genotypes not genotokens" (Sober-Lewontin 1982, p. 172; see also Sober-Lewontin 1983, p. 649). And even more explicitly: "to understand what it means to talk about the selection of genes, organisms, or groups, one must quantify over properties" (Sober 1981, p. 162).

This may sound like bad news for my claim that property-types do not play a role in explaining why token organisms have the traits they have, but remember that Sober maintains that what explains why token organisms have the traits they have is not selection. What does this explanatory work, according to Sober is random mutation and inheritance (and, presumably, some developmental factors) (Sober 1995). And mutation, inheritance and development should be taken to be token phenomena here: the mutation in a token ancestor of the organism, inheritance from one token organism to another and the developmental processes of the organism are all processes that operate on token traits. In short, if we take Sober's side in the grand debate about the explanatory power of selection, we can conclude that property-types do not play any role in explaining why specific token organisms have the traits they have.

But what if we take Neander's side, who claims that cumulative selection does explain why specific token organisms have the traits they have? Here, the answer is more complicated. First, it is important to note that Neander's claim is that *cumulative* selection can explain why specific token organisms have the traits they have. And here is what she means by cumulative selection: what makes cumulative selection *cumulative*: that "the probable outcome of future [rounds of selection] depends on the results of previous [rounds of selection] (Neander 1995b, p. 584).

It could be pointed out that this conception is much stronger than what Sober means by cumulative selection (most famously in his discussion of the 'selection toy' [Sober 1984, p. 99])-and this may explain the miscommunication between Sober and Neander. But what is important from the point of view of the present argument is that Neander's way of interpreting cumulative selection does not presuppose any talk of property-types. Take the (uniparental) organism, a, whose neck is 12 cm long. It has two offspring, b and c, with 14 and 10 cm long neck, respectively. As organism b gets to reach branches with leaves that organism c cannot, it gets to survive, whereas organism c starves to death. Organism b also has two offspring, d and e, with 16 and 12 cm long neck and d survives, whereas e starves to death. This is cumulative selection in Neander's sense: the traits of the organisms change from generation to generation, not just the trait frequencies. Further, the fact that *b* survives and *c* dies influences neck size of the next generation. As it is b (and not c) who gets to survive, the starting point for the variation in neck size for the next generation is 14 cm (and not 10 cm).

This way of thinking about cumulative selection explains the token traits of specific organisms in terms of past selection of token traits. It explains why organism d has such a long neck by pointing out that d's parent, b had longer neck than c and got to survive as a result. The difference between traits that are responsible for the death of some organisms and the survival of others is a difference between token traits of specific organisms (i.e., the difference between b's 14 cm long neck and c's 10 cm long neck) and not a difference between abstract trait types.

A possible objection: Couldn't we refer to the trait tokens of specific organisms as instantiations of a trait type, say, the trait type of being 12 cm long? We could, but this would have very problematic consequences. An instantiation of the trait type of being 12 cm long was responsible for a's survival, but another instantiation of the same trait type is responsible for e's death (whose neck is also 12 cm long). Hence, it would be problematic to take what is responsible for the death or survival of organisms to be trait types, because instantiations of the very same trait type are responsible for the death in one generation and survival in another. What is responsible for the death of e and the survival of a are not trait types, but trait tokens (see Nanay 2010a for a more detailed version of this argument as well as for some clarifications that I could not include here).

But then this way of thinking about selection does not appeal to any property-types. Hence, if selection, interpreted in this way, can explain why specific token organisms are the way they are, then it is still true that property-types do not explain why specific token organisms are the way they are. Regardless of which side of the debate over the explanatory power of selection we choose, it remains true that property-types do not play any explanatory role in explaining why specific token organisms are the way they are.

Of course not all evolutionary explanations are *selective* explanations. I was focusing on the explanatory role of selection above, but similar arguments could be given with regards to non-selective evolutionary processes, such as the founding effect. Like selection, founding effect can also be fully accounted for by individual level processes, such as the specific evolutionary history of specific organism, without appealing to any trait types. Thus, we have no reason to attribute any explanatory role to any property-type.⁶

My conclusion is then that even if we do not have any reason to deny that there may be some properties, that is, property-types, that all and only members of a biological kind possess in all possible worlds, if the argument I presented in this section is correct, we need to conclude that these properties, that is, property-types, although they may exist, play no explanatory role for the simple reason that no property-type plays any explanatory role in the biological domain. Thus, essentialism about biological kinds fails not because of the first (or the second), but because of the third tenet. The problem is not with the existence of essential property-types, but with their explanatory role.⁷

The conclusion, then, is that we should be skeptical about the third tenet of kind-essentialism in the biological domain. Putnam and Kripke may be right about chemical kinds, but essentialism is unlikely to be the correct view about biological kinds.

Acknowledgments

I am grateful for comments by Peter Godfrey-Smith, Jason Rhein, Karen Neander, David Sanford, Marc Lange, David Papineau, Roberta Balarin, James Young, and Jeff Foss on an earlier version of this paper. I presented this paper at the Inland Northwest Philosophy Conference and at the British Columbia Philosophy Conference. I am grateful for my commentators as well as the audiences. This paper grew out of the discussion at my PhD seminar on scientific realism at Syracuse University in fall 2007. I'm also grateful for the feedback of my students in my seminar on natural kinds in fall 2008, and for the tough criticism from my colleagues at the mysterious SPDMBABWS Society.

Notes

1. Griffiths 1999; Boyd 1999; Okasha 2002. See Walsh 2006 and Devitt 2008 for a different way of resurrecting essentialism about natural kinds, and Ereshefsky (Manuscript) for objections to both projects.

2. There may be ways of weakening essentialism even more by denying that essentialism implies that all and only the members of a kind must have a kind-specific essence (Boyd 1999). I will say a bit more, in sections 5 and 6, about the relevance to such accounts of the argument I present in this paper.

3. It is worth noting that one way of defending Mayr's position from worries of this kind would be to embrace the recently popular view that populations are individuals and members of populations are the parts of this individual (Ghiselin 1974; Hull 1978).

4. It is also worth noting that as Mayr considers property-types as "merely statistical abstractions," he denies that they have mind-independent existence—they do not have "reality," as he puts it. So Mayr's version of trope nominalism is of the radical kind I considered at the end of section 2.

5. It is not an easy task to give an exact definition for what counts as an evolutionary explanation, but what I mean by this concept is quite broad: an evolutionary explanation is an explanation where the *explanandum* can be pretty much anything (although most of the time it is the apparent teleology of a property), but the *explanans* is an evolutionary process. Evolutionary processes, in turn, are defined conjunctively to include selection, founding effect, etc.

6. To give an example of a property-type that is considered to be a good candidate for a property that all and only members of a species have in all possible worlds: the property of being a member of a population with such and such distinctive evolutionary history (Griffiths 1999; Okasha 2002; the idea, again, comes from the anti-essentialist Matthen 1998). But everything this property-type can explain can be explained by individual-level processes, such as the specific evolutionary history of a specific organism, without appealing to any trait-types.

7. Richard Boyd argued that we can give an even weaker formulation of essentialism than the one we have considered so far. More precisely, essentialism does not necessarily imply that all and only the members of a natural kind must have a kind-specific essence. According to his "homeostatic property cluster theory," the members of a kind share a cluster of similar properties, but no property is necessary for membership in this kind (Boyd 1999). Boyd's "homeostatic property cluster theory" is quite complex and I do not intend to give a definitive argument against it. But we may be able to use the considerations above to make the following conditional claim. Boyd explicitly states that "the homeostatic clustering of properties . . . is causally important" (1999, 143). If this is to be understood in such a way that it is also *explanatorily* relevant, and if the "homeostatic clustering of properties" is supposed to be understood as a *type* that can have a number of different token instantiations, then Boyd's view contradicts the considerations I presented above in favor of the claim that property-types play no role in biological explanations.

References

Ariew, A. 2008. Population thinking. In *Oxford Handbook of Philosophy of Biology*, ed. M. Ruse. Oxford: Oxford University Press.

Armstrong, D. M. 1997. A World of States of Affairs. Cambridge: Cambridge University Press.

Boyd, R. 1999. Homeostasis, species, and higher taxa. In *Species: New Interdisciplinary Essays*, ed. R. Wilson. Cambridge, Mass.: MIT Press.

Campbell, K. 1981. The metaphysics of abstract particulars. *Midwest Studies in Philosophy* 6:477–488. Reprinted in *Properties*, ed. D. H. Mellor and A. Oliver. Oxford: Oxford University Press, 1997.

Campbell, K. 1990. Abstract Particulars. Oxford: Blackwell.

Cummins, R. 1975. Functional analysis. Journal of Philosophy 72:741–765.

Daly, C. 1997. Tropes. In *Properties*, ed. D. H. Mellor and A. Oliver. Oxford: Oxford University Press.

Darwin, C. 1859. On the Origin of Species. Cambridge, Mass.: Harvard University Press.

Davidson, D. 1967. Causal relation. *Journal of Philosophy* 64: 691–703. Reprinted in D. Davidson, *Essays on Actions and Events*. Oxford: Clarendon Press, 1980.

Davidson, D. 1970. Mental events. In *Experience and Theory*, ed. L. Foster and J. Swanson. Amherst: University of Massachusetts Press. Reprinted in D. Davidson, *Essays on Actions and Events*. Oxford: Clarendon Press, 1980.

Davidson, D. 1993. Thinking causes. In *Mental Causation*, ed. J. Heil and A. Mele. Oxford: Oxford University Press.

Dawkins, R. 1982. The Extended Phenotype. Oxford: W. H. Freeman.

Devitt, M. 2008. Resurrecting biological essentialism. *Philosophy of Science* 75: 344–382.

Dretske, F. 1988. Explaining Behavior. Cambridge, Mass.: MIT Press.

Dretske, F. 1990. Reply to reviewers. *Philosophy and Phenomenological Research* 50:819–839.

Dupré, J. 1993. The Disunity of Science: Metaphysical Foundation of the Disunity of Science. Cambridge, Mass.: Harvard University Press.

Dupré, J. 2002. Humans and Other Animals. Oxford: Oxford University Press.

Ehring, D. 1997. *Causation and Persistence: A Theory of Causation*. Oxford: Oxford University Press.

Elder, C. L. 1995. A different kind of natural kind. *Australasian Journal of Philosophy* 73:516–531.

Ellis, B. 2001. Scientific Essentialism. Cambridge: Cambridge University Press.

Ereshefsky, M. 1998. Species pluralism and anti-realism. *Philosophy of Science* 65:103–120.

Ereshefsky, M. 2007. Species. In *The Stanford Encyclopedia of Philosophy*, ed. E. Zalta. Fall 2007 edition. http://plato.stanford.edu/.

Ereshefsky, M. Forthcoming. Natural kinds in biology. *Routledge Encyclopedia of Philosophy.*

Ereshefsky, M. Manuscript. What's wrong with the new biological essentialism?

Ghiselin, M. T. 1974. A radical solution to the species problem. *Systematic Zoology* 23:536–544.

Godfrey-Smith, P. 2009. *Darwinian Populations and Natural Selection*. Oxford: Oxford University Press.

Griffiths, P. 1999. Squaring the circle: Natural kinds with historical essences. In *Species: New Interdisciplinary Essays*, ed. R. Wilson. Cambridge, Mass.: MIT Press.

Hacking, I. 2007. Natural kinds: Rosy dawn, scholastic twilight. In *The Philosophy of Science, Royal Institute of Philosophy Supplements*, vol. 61, ed. A. O'Hear. Cambridge: Cambridge University Press.

Hawthorne, J., and T. S. Gendler. 2000. Origin essentialism: The arguments reconsidered. *Mind* 109:285–298.

Hull, D. 1965. The effect of essentialism on taxonomy: 2000 years of stasis. *British Journal for the Philosophy of Science* 15:314–326; 16:1–18.

Hull, D. 1978. A matter of individuality. Philosophy of Science 45:335-360.

Hull, D. 1981. Units of evolution: A metaphysical essay. In *The Philosophy of Evolution*, ed. U. J. Jensen and R. Harré. Brighton: Harvester Press.

Hull, D. 1986. On human nature. *PSA: Proceedings of the Biannual Meetings of the Philosophy of Science Association* 2:3–13.

Hull, D. 1994. Contemporary systematic philosophies. In *Conceptual Issues in Evolutionary Biology*, 2nd ed., ed. E. Sober. Cambridge, Mass.: MIT Press.

Hull, D. L. 2001. Science and Selection. Cambridge: Cambridge University Press.

Hull, D. L., R. E. Langman, and S. S. Glenn. 2001. A general account of selection: Biology, immunology, and behavior. *Behavioral and Brain Sciences* 24:511–528. Reprinted in Hull 2001.

Kitcher, P. 2007. Does 'race' have a future? *Philosophy & Public Affairs* 35:293-317.

Kripke, S. 1980. Naming and Necessity. Cambridge, Mass.: Harvard University Press.

Lewens, T. 2001. Sex and selection: Reply to Matthen. *British Journal for the Philosophy of Science* 52:589–598.

Matthen, M. 1998. Biological universals and the nature of fear. *Journal of Philosophy* 95:105–132.

Matthen, M. 1999. Evolution, Wisconsin style: Selection and the explanation of individual traits. *British Journal for the Philosophy of Science* 50:143–150.

Matthen, M. 2003. Is sex really necessary? And other questions for Lewens. *British Journal for the Philosophy of Science* 54:297–308.

Mayr, E. 1942. *Systematics and the Origin of Species*. New York: Columbia University Press.

Mayr, E. 1959. Typological versus population thinking. In *Evolution and Anthropology*, ed. B. J. Meggers. Washington: The Anthropological Society of America. Reprinted in *Conceptual Issues in Evolutionary Biology*, ed. E. Sober. Cambridge, Mass.: MIT Press, 1994.

Mayr, E. 1963. *Animal Species and Evolution*. Cambridge, Mass.: Harvard University Press.

Mayr, E. 1969. The biological meaning of species. *Biological Journal of the Linnean Society. Linnean Society of London* 1:311–320. Reprinted in E. Mayr, *Evolution and the Diversity of Life.* Cambridge, Mass.: Harvard University Press, 1976.

Mayr, E. 1982. *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge, Mass.: Harvard University Press.

Mayr, E. 1988. The ontology of the species taxon. In *Towards a New Philosophy of Biology*. Cambridge, Mass.: Harvard University Press.

Mayr, E. 1996. What is a species, and what is not? *Philosophy of Science* 63:262–277.

Mellor, D. H. 1995. The Facts of Causation. London: Routledge.

Millikan, R. G. 1990. Seismograph readings for "explaining behavior." *Philosophy* and *Phenomenological Research* 50:819–839.

Millikan, R. G. 1999. Historical kinds and the "special sciences." *Philosophical Studies* 95:45–65.

Nanay, B. 2005. Can cumulative selection explain adaptation? *Philosophy of Science* 72:1099–1112.

Nanay, B. 2009. The properties of singular causation. Monist 92:113-135.

Nanay, B. 2010a. Population thinking as trope nominalism. Synthese 177:91–109.

Nanay, B. 2010b. Natural selection and the limitations of environmental resources. *Studies in History and Philosophy of Biological and Biomedical Sciences* 41:418–419.

Nanay, B. 2011. What if reality has no architecture? Monist 94:181–197.

Neander, K. 1995a. Pruning the tree of life. *British Journal for the Philosophy of Science* 46:59–80.

Neander, K. 1995b. Explaining complex adaptations. A reply to Sober's "Reply to Neander." *British Journal for the Philosophy of Science* 46:583–587.

Okasha, S. 2002. Darwinian metaphysics: Species and the question of essentialism. *Synthese* 131:191–213.

Oliver, A. 1996. The metaphysics of properties. Mind 105:1-80.

Platts, M. 1983. Explanatory kinds. British Journal for the Philosophy of Science 34:133–148.

Pust, J. 2001. Natural selection explanation and origin essentialism. *Canadian Journal of Philosophy* 31:201–220.

Putnam, H. 1975. *Mind, Language, and Reality*. Cambridge: Cambridge University Press.

Robb, D. 1997. The properties of mental causation. *Philosophical Quarterly* 47:178–194.

Robertson, T. 1998. Possibilities and the arguments for origin essentialism. *Mind* 107:729-749.

Rosenberg, A. 1983. Coefficients, effects, and genic selection. *Philosophy of Science* 50:332–338.

Sanford, D. H. Manuscript. Tropiary for beginners: Commentary on Peter Simons's "Particulars in Particular Clothing."

Schaffer, J. 2001. The individuation of tropes. *Australasian Journal of Philosophy* 79:247–257.

Simons, P. 1994. Particulars in particular clothing. *Philosophy and Phenomenological Research* 54:553–575.

Smith, B., and A. C. Varzi. 2001. Environmental metaphysics. In *Metaphysics in the Post-Metaphysical Age*, ed. U. Meixner. Vienna: Hölder-Pichler-Tempsky.

Smith, B., and A. C. Varzi. 2002. Surrounding space: On the ontology of organismenvironment relations. *Theory in Biosciences* 120:139–162.

Sober, E. 1980. Evolution, population thinking, and essentialism. *Philosophy of Science* 47:350–383.

Sober, E. 1981. Evolutionary theory and the ontological status of properties. *Philosophical Studies* 40:147–176.

Sober, E. 1984. The Nature of Selection. Cambridge, Mass.: MIT Press.

Sober, E. 1995. Natural selection and distributive explanation. *British Journal for the Philosophy of Science* 46:384–397.

Sober, E., and R. Lewontin. 1982. Artifact, cause, and genic selection. *Philosophy of Science* 49:157–180.

Sober, E., and R. Lewontin. 1983. Reply to Rosenberg on genic selection. *Philosophy* of Science 50:648–650.

Stegmann, U. 2010. What can natural selection explain? *Studies in History and Philosophy of Biological and Biomedical Sciences* 41:61–66.

van Fraassen, B. 1980. The Scientific Image. Oxford: Oxford University Press.

Walsh, D. M. 1998. The scope of selection: Sober and Neander on what natural selection explains. *Australasian Journal of Philosophy* 76:250–264.

Walsh, D. M. 2006. Evolutionary essentialism. British Journal for the Philosophy of Science 57:425-448.

Walsh, D. M., T. Lewens, and A. Ariew. 2002. The trials of life: Natural selection and random drift. *Philosophy of Science* 69:429–446.

Wilkerson, T. E. 1995. Natural Kinds. Aldershot: Avebury.

Williams, D. C. 1953. On the elements of being. *Review of Metaphysics* 7:3–18, 171–192.

Wilson, R. A. 1999. Realism, essence, and kind: Resuscitating species essentialism? In *Species: New Interdisciplinary Essays*, ed. R. A. Wilson. Cambridge, Mass.: MIT Press.

Wilson, R. A., M. J. Barker, and I. Brigandt. Forthcoming. When traditional essentialism fails: Biological natural kinds. *Philosophical Topics*.

Wimsatt, W. C. 1980. Units of selection and the structure of the multi-level genome. *Philosophy of Science Association* 2:122–193.

Wolterstorff, N. 1973. On Universals. Chicago: University of Chicago Press.

10 Arthritis and Nature's Joints

Neil E. Williams

1 Disease Kinds and Essences

It should come as a surprise to almost no-one that the thought that diseases constitute natural kinds does not generally sit well with the essentialist picture of natural kinds as championed by Kripke and Putnam in the 1970s.¹ According to that essentialist picture, in order for a class of entities to be a natural kind it is required that all and only members of the class instantiate some very specific property or properties, and that these properties explain the presence of any other properties typically associated with being a member of the kind. These privileged properties constitute the *essence* of the kind. In the case of diseases, the essentialist claims that the properties in question are etiological: the essence of a disease kind is whatever underlying physical condition causes the instances of the disease and the associated symptoms.²

There are well known examples of diseases for which the essentialist picture appears to work just fine, such as tuberculosis: originally identified and classified on the basis of the name-giving tubercle, the stereotyped "nominal" essence classification of tuberculosis gave way to "real" essence classification with Koch's 1882 discovery that the *Mycobacterium tuberculosis* bacterium was the cause of the disease. Alongside tuberculosis, the essentialist can list cholera, meningitis, plague, botulism, malaria, syphilis, and a number of other diseases; but despite this list of disease types that fit the essentialist model, the "success" stories for the essentialist treatment of disease kinds have been rather limited.³ For every disease type that appears to satisfy the essentialist desiderata there are a dozen or more that do not. For instance, contrast the case of tuberculosis with that of rheumatoid arthritis (RA). The American Rheumatism Association (ARA) presently defines an instance of rheumatoid arthritis as one displaying at least four of the seven stipulated diagnostic criteria.⁴ The current classification

of the disease is entirely in terms of its clinical picture. Not only is there no known cause, even the diagnostic criteria are only rough guides. (Recent studies report that more than 70 percent of those who test positive for rheumatoid factor may not have the disease [Newman et al. 1996, 8].)

Optimistic essentialists might reply that diseases like RA are only pseudo-counterexamples to essentialism: essentialism is a *metaphysical* thesis concerning the constitution of natural kinds, and nothing about the case of RA shows that there is not some underlying physical cause which is present in all and only instances of RA. It would be fallacious to assume on the basis of our not *knowing* the cause of RA, at present, that there is no cause to be known. In fact, the ARA even stipulate that their understanding of RA is subject to revision in light of future understanding, and anticipate that future classification will be based directly on improved understanding of the underlying disease pathology. Perhaps all is well in the essentialist camp.

But what happens if RA fails to have a tidy causal structure, or has multiple causes? Even as we learn more about the biological details of the disease, the possibility of multiple causes does not get ruled out. Nor can we rule out RA having a cause that is also the cause of other types of disease.⁵ Should it turn out that RA has many causes, no cause, or no cause unique to RA, then the essentialist is forced to tell us that RA is not a natural kind. Should RA turn out to have no cause at all, then it is not a disease type but a syndrome (i.e., a collection or pattern of symptoms lacking a joint cause). In fact, even in the heralded case of tuberculosis, it turns out that the disease has at least two causes (viz., infection by Mycobacterium tuberculosis or infection by Mycobacterium bovis), so neither can count as the essence.⁶ But the problems for essentialism do not end there. In the purportedly successful cases where the causal agent has been identified, disease-kind membership depends crucially on there being clearly delineated bacterial and viral kinds, and they too face problems of classification (Franklin 2007). And even if those difficulties should be surmountable, it is now 130 years after Koch's discovery of the cause of tuberculosis and there remains bitter disagreement among medical practitioners about how the disease name 'tuberculosis' should be used, indicating widespread concern with the essentialist's picture of natural kinds.⁷ Rather than continuing to consider ways that the essentialist picture of the metaphysics of natural kinds lets us down when it comes to disease kinds, or how our present knowledge and handling of diseases suggests something other than the essentialist's metaphysical picture of disease kinds, it should suffice to say that the essentialist's account of natural kinds looks ill suited for dealing with disease kinds.

The problem is that the essentialist picture provides us with only two options, and neither is desirable: (i) satisfy the essentialist desiderata by locating some etiological feature both necessary and sufficient for kind membership, or (ii) give up on the thought that similarity between disease instances can be understood in terms of natural kinds.

The facts about disease largely rule out the first option; at best only a small subset of the recognized disease types would qualify as natural kinds. This is unsatisfactory as it conflicts with the strong intuition that most of the disease types we identify pick out genuinely natural groupings of disease instances, an intuition that plays a pivotal role in determining the methodologies we adopt in investigating, treating, and preventing diseases, not to mention in disease classification. Moreover, if only a very small class of disease types satisfies the essentialist desiderata, we must forego "naturalness" as even a rough means of distinguishing genuine diseases from fabricated ones. That is, we run the risk of seeing all non-essentialist disease groupings as non-natural nominalizations, and so lose a relatively easy way of explaining that many of the groups of symptoms we treat together we do so because they form a natural unit, whereas other groupings are of our own making.⁸

The second option is clearly no better; whereas the first option severely restricts the number of disease kinds, the second option rules them out altogether. As I have suggested, the thought that diseases form natural kinds is the theoretical foundation on which our methods of treating, investigating, and preventing disease are predicated. Without this foundation, our inductive and explanatory practices concerning disease are without grounding.⁹ If we are to continue our practice of treating disease types as natural kinds—without unduly restricting the class of disease types—then I suggest that we have no choice but to abandon the traditional essentialist picture of natural kinds when dealing with disease kinds and replace it with a notion that is more disease-kind friendly.¹⁰

In what follows that is exactly what I attempt to do. That is, I will propose a treatment of natural kinds suited for dealing with disease kinds. What I will propose is that disease kinds be interpreted as a specialized case of *homeostatic property cluster kinds* (HPC) of the sort proposed by Richard Boyd.¹¹ In the HPC account of kinds, natural causal processes produce stable groups of properties that contingently "cluster" together, and these clusters determine kind-membership. Consequently, part of the present task will be to specify the sorts of properties that are included in the property clusters. However, as diseases are not substances, treating disease kinds as HPC natural kinds will require modifying the HPC theory in what should be a welcome direction.

Before proceeding, I offer two words of caution: (1) in order to make the range of diseases slightly more manageable, the central focus will be on *human* diseases, and will exclude mental illnesses. I am optimistic that the account will be applicable to the latter group, but considering mental illnesses brings in a range of additional worries that are best left out of the present discussion. Likewise I expect the present treatment to apply to diseases in multi-cellular organisms in general, but I will nevertheless ignore nonhuman diseases.¹² (2) This is intended as a *metaphysical* account of what constitute natural kinds of diseases, and should not be interpreted as a guideline for diagnosis.¹³ It is an interesting question how disease kinds relate to our attempts to diagnose what type of disease a particular patient might have, but this is (mostly) a distinct issue about the application of disease kinds to a specific clinical picture and is largely an epistemic matter, and one I lack the space to go into here.

The procedure will be as follows. In section 2, I continue my defense of the claim that we should treat disease kinds as natural kinds (and so the bad fit with the essentialist picture of natural kinds means we need a replacement account). Section 3 contrasts essentialism with the HPC theory, providing a brief overview of each and the HPC framework that the present treatment adopts for disease kinds. As no treatment of disease kinds can be fully divorced from an account of what a disease is, section 4 provides an account of what diseases are. Section 5 contains the proposed treatment of disease kinds.

2 Why Treat Disease Kinds as Natural Kinds?

If natural kinds are as the essentialist claims, then few or none of the disease types we recognize constitute natural kinds. I suggest that we resolve this incongruence by adopting an alternative picture of natural kinds. But why not just drop the notion of natural kinds and make do with something else? Hacking, for instance, states that the student of kinds will want a theory of kinds within which natural kinds "take their proper, rather limited place" (Hacking 1991, 109). Most kinds, he thinks, are social rather than natural, and have more to do with their roles as tools than as classifications. Perhaps we should think of disease kinds in this way: membership in a disease kind is more a matter of our efforts to control our environment than of natural grouping. Something strikes me as remotely correct with this picture, but I cannot help but think that our disease-oriented practices strongly favor the naturalness of kinds over the picture of disease kinds as mere tools. In fact, I am confident that

naturalness is required for the ways we employ natural kinds in medical thinking. Let us consider then why we ought to treat disease kinds as *natural* kinds.

The first reason is simple: essentialism is but one way of approaching natural kinds. Essentialists clearly agree that disease kinds are natural kinds; rejecting the essentialist treatment of natural kinds does not require rejecting the shared intuition that disease kinds are natural kinds. That is, we can maintain the intuition that diseases form natural kinds without requiring that each disease kind has some property possessed by all and only instances of the kind, and which explains the presence of any other property associated with that natural kind. One gets a sense from much of the kinds literature (especially as it concerns diseases) that the essentialist picture of natural kinds is the *only* picture of natural kinds; but this impression is mistaken. The thought that there are kinds in nature is a basic and useful one, but nothing about this thought requires that to *be* natural the kinds must satisfy the essentialist's desiderata.

A second reason for treating disease kinds as natural kinds is that this thought plays a significant role in determining and justifying how and why we go about our medical investigations in the ways we do. When we find a group of similar symptoms arising in a statistically significant number of instances, we treat that statistical significance as an indication that the group of symptoms is *unified*—that is, that they are all indicators of the same disease. We then engage in various activities that only make sense if we are treating that group as forming a *natural* group—that is, they only make sense if we are taking the collections of symptoms to hang together in virtue of mechanisms outside of our conceptual organization. For starters, we try to find out what causes the symptoms. For each instance in which the symptoms arise we assume that there is some single salient cause of these symptoms. If we discover a cause in one instance, we then look for that cause in others. We might discover that the same group of symptoms has different causes in different instances, in which case we think of that group as having multiple potential causes. Or we might discover that some subset of the symptoms is highly correlated with some other cause, and come to treat that as a distinct disease. These are the actions of people treating diseases as natural kinds.

Nor do these actions end there. Prevention is about preventing the underlying process that produces a particular group of symptoms from arising; cure is a matter of locating the cause of these symptoms and attempting to arrest or reverse its effects; diagnosis is a matter of trying to match a patient's clinical picture with those clinical pictures associated with specific types of disease, despite what is typically impoverished data about which symptoms are present in a given case.¹⁴ If we learn that a remedy is effective in a significant number of instances, then we continue to apply that remedy to further instances, and track its success. For instance, one of the SAARDs (slow-acting anti-rheumatic drugs) used in the treatment of RA is the injection of gold. That we would attempt a treatment in one instance because it has worked in others is a clear indication that we believe the instances are members of a naturally occurring group, and so our inductions are justified; that we are successful as often as we are is a reasonable indictor that we are correct.¹⁵

What this boils down to is what Boyd has identified as a largely realist commitment to the causal structure of the world, and of a "deference to nature" with regards to which symptoms to take to be connected (Boyd 1991, 140). "Kinds useful for induction or explanation must always 'cut the world at its joints' in this sense: successful induction and explanation always require that we accommodate our categories to the causal structure of the world" (ibid., 139). It is in virtue of this causal structure that we have developed the understanding of disease that we have, and that we seek explanations in the way we do. We take treatments to be repeatable, and information gathered from one instance of a disease to be relevant to further instances of that disease, because we take similarity of disease instances within disease types to be a naturally occurring feature of our world. In short, we treat medical information as *projectible*, and we do so on the grounds that disease kinds are natural kinds.

By way of contrast, consider what would be the case were we not treating disease kinds as natural kinds. Though we would certainly recognize symptoms as arising naturally, we would not be annexing those symptoms to a single type. Remedial practices of clusters of symptoms would no longer be a matter of treating those symptoms *jointly*, but rather *severally*. This might still have some success, not because some single cause or process had been stopped or reversed, but rather because numerous independent and unconnected processes had been stopped or reversed. This sort of difference parallels that which we find in uses of broad-spectrum antibiotics as opposed to narrow-spectrum antibiotics. In those cases where we have identified some specific bacterial infection as a (typical) cause of a disease, treatment proceeds via narrow-spectrum antibiotics that target just that specific type of bacterial infection. But where the cause is unknown (and the effects typically quick and fatal), broad-spectrum antibiotics are applied as a sort of anti-bacterial carpet bombing. If symptoms were treated severally and not jointly, then there would be fewer differences in how we

understand the application of narrow-spectrum versus broad-spectrum antibiotics. To be clear, this might (unfortunately¹⁶) be all we are doing when we treat diseases, but it is not what we *think* we are doing, and that attitude is sufficient evidence of our treating disease kinds as forming natural kinds.

A final reason in support of offering an alternative picture of natural kinds for disease rather than just giving up on the thought that disease kinds are natural kinds comes from the directly parallel response to essentialism with regards to species. Having argued that the essentialist picture of natural kinds did not sit well with any of the various biological classifications, biologists and philosophers of biology had originally backed away from treating biological kinds as natural kinds, and initiated a retreat from natural kinds in general.¹⁷ But not long afterwards it was recognized that the natural-kind concept was a fruitful and central concept in the biological sciences, and that the problem was with the essentialist picture of natural kinds, not with treating biological kinds as natural kinds. Consequently some biologists and philosophers of biology have started to "take back" the notion of natural kind, offering alternative conceptions which are better suited to biological classifications.¹⁸ After all, biological kinds species in particular-are paradigmatic natural kinds; if the essentialist picture of natural kinds was ruling them out, then it was this picture that needed to change.

My recommendation is that a similar revolution be initiated within the medical sciences. As the essentialist picture of natural kinds is no more appropriate for disease kinds than it is for other biological kinds, the 'natural kind' concept must be taken back and a new version devised that is better suited to medical science and disease classification.

3 Two Approaches to Natural Kinds

Like its essentialist predecessor, the proposed treatment of disease kinds here is realist and continues to be a metaphysical treatment of natural kinds, but it is otherwise quite different. As I have said, the new approach is a specialized version of the HPC theory of kinds: in this section I will give a brief overview of the HPC theory.¹⁹ This overview is presented against a backdrop of the essentialist view. I do this partly because it makes it easier to see how the two accounts differ, but also because there are features of essentialism we have not yet considered (concerning the larger model of science in which essentialism is embedded) that are similarly inappropriate for dealing with disease kinds.
We have already seen the central theses of essentialism: all members of a natural kind instantiate some very restricted set of properties (typically a single property), and the having of this property-set (or property) is both necessary and sufficient for kind-membership. This privileged set of properties constitutes the "essence" of the kind. On the assumption that substances can be sorted into clearly demarcated classes in virtue of their instantiating (or failing to instantiate) these essential properties, the essentialist's kinds are likewise clearly demarcated, and are disjoint. Where a substance appears to be a member of two kinds, this can only be so if the one kind is a subset of the other. Hence, according to the essentialist, all kinds are parts of a single, immutable, hierarchical structure.

We have also seen that the essential properties will explain the other properties typically had in common by the members of the kind (referred to at times as the "nominal" essence, "stereotype," or "surface properties" of the kind). The basic picture is this: members of a natural kind will tend to have numerous properties in common, but only a small subset of properties will be had by all, and fewer still make up the essence of the kind. Whereas these essential properties must be had by all members of the kind, the other properties will only *tend* to be had by all members—they will be had in most cases, but need not be had by all. In fact, essentialism is compatible with its being that case that no member of the kind has any of the surface properties. Consequently, the surface properties play no role in determining kind-membership. What will be the case, however, is that the presence of the surface properties is *explained* by the presence of the essential properties. For instance, in the case of gold, common surface properties such as instances of gold being shiny, malleable, heavy, and so on, are to be explained via the essence of the kind-in this case being composed of atoms with the atomic number 79. How exactly this explaining takes place is not clear, but the general response has to do with the ways in which essential properties slot into the laws of nature. Presumably the having of essential properties in the right sort of lawlike environment provides an account of why the surface properties are what they are. From here we begin to get a fuller sense of the essentialist's metaphysical picture: because the essential properties explain the surface properties through their interaction with the laws of nature, it follows that the essential properties will be categorical properties, and that they will tend to be intrinsic rather than relational. Consequently any dispositional properties we might find among the surface properties of the kind are to be similarly explained by the categorical properties in the essence; likewise for any relational properties.²⁰

Moving on to the wider framework surrounding the essentialist picture of natural kinds, we find frequent reference to the roles of natural kinds within the system of natural laws. We have already seen that they connect via the essential properties, but it is also part of the wider framework that the laws of nature range over natural kinds. This gives the essentialist an additional criterion by which to assess whether a kind is natural or not: only natural kinds are appropriate subjects of the laws of nature.²¹ Like the hierarchical structure of natural kinds, the natural laws are unchanging and exceptionless: the kinds and the laws mesh perfectly to form a seamless closed system.

The homeostatic property cluster (HPC) theory of natural kinds diverges in a number of significant ways from its essentialist counterpart. Where the essentialist demands determinateness and structure, the HPC theorist permits indeterminateness. Where the essentialist insists on sharp boundaries, the HPC theorist will sometimes draw sharp boundaries but at others draws rough ones. The HPC theorist likewise tolerates change, and greatly inflates the number of properties involved in determining the kind. Let us start with this final difference. In place of the essentialist's restricted set of essential properties that are necessary and sufficient for kind-membership, the HPC theorist suggests that there is a "cluster" of key properties. Like the essentialist, the HPC theorist distinguishes purely accidental properties arising in the kind from those important to kind-determination, including only the latter in the cluster, but this set of important properties is much larger than the very restricted set the essentialist uses for the same purpose. In a very rough sense, we might think of the HPC theorist's cluster as incorporating both the essentialist's essential properties as well as a number of the surface properties.²² However, within the cluster no particular property is either necessary or sufficient for membership within the kind. What matters is the *extent* to which the properties a substance instantiates overlap with the properties in the cluster, which can be satisfied to varying degrees. The extent to which a substance is a member of a class or its complement will vary accordingly. It follows that HPC natural kinds are not required to have sharp boundaries or determinate cut-off points between them, but rather can have ranges of peaks and valleys, where peaks indicate the highest degree of property-overlap with the cluster, and valleys the least.²³ According to Boyd, this indeterminacy is incapable of revision: to insist on anything sharper would be to render kinds so "unnatural" that they would no longer correspond to their causal engagement with the world. In effect, a sharpening of "natural" kinds would make them artificial: they would be sharper than the world itself.²⁴

As well as dulling the essentialist's sharp boundaries, the HPC theorist makes the metaphysical privilege the essentialist affords the essential properties a largely epistemic matter. For our own purposes of determining whether a given substance is a member of this or that kind, we treat certain properties as more important than others—but this importance need not be reflected in the kinds themselves. The privileged properties are reliable indicators of kind-membership for the HPC theorist, but they do not determine kind-membership. The range of the kind reflects the full cluster of properties, not this or that property. In removing the metaphysical force of an essential-versus-surface-property distinction, the HPC theorist also removes the explanatory relation the essentialist takes these disparate properties to share. In its place we are told that the co-occurrence of the cluster as a whole is explained by the causal mechanisms of the world: it is those properties that group together naturally, where 'natural' is understood causally. It is the cluster-producing effects of these homeostatic mechanisms that lend their name to the HPC account: the clusters are formed through a natural balance of environmental processes that act on members of the kind.²⁵ The property clusters are as they are because they reflect the causal structure of the world.

The next difference between essentialism and the HPC theory concerns the types of properties within the cluster. The cluster can include the intrinsic categorical properties the essentialist includes, but also dispositional and relational properties. The HPC theory adopts a permissive approach to cluster properties in general, not endorsing any particular metaphysics concerning the laws of nature. This is in direct contrast to the essentialist, whose account brings with it strong metaphysical commitments concerning the laws of nature and causal powers.²⁶ This demonstrates yet another point of departure: though the HPC theory is realist about there being a causal nature to reality, it remains neutral concerning its particular features. This allows it to be compatible with a range of different approaches to the laws of nature, including those that reject such laws outright.

A final departure concerns the ahistoricity of the essentialist's natural kinds. The essentialist imagines an immutable hierarchical structure; the HPC theorist allows that kinds can, and do, change with time. On the assumption that species are among the natural kinds in biology and are subject to evolutionary changes, it stands to reason that the natural kinds in biology will change with time. The properties associated with a particular species of organism reflect that species' place in the world and are, in part, a reflection of its interactions with the world. Internal mutations and

changing environmental factors will produce changes within the species of organism, but these need not be changes that constitute the appearance of a new species. As species change under environmental pressures, so too must natural kinds; this variability within the species is accommodated through the flexibility of the property clusters.

Despite the many differences between the essentialist and HPC treatments of natural kinds, it is important to note that they are similar to the extent that each offers an *a posteriori* treatment of the knowledge of natural kinds. According to both theories, natural-kind classifications are a matter of discovery: the natural kinds exist prior to—and independently of—our knowledge of them, and it is through our interactions with their members that we may come to learn about the kinds. Essentialism is, at its most basic, a metaphysically realist and robust account of kinds, and as such assumes that we start out in the dark; in this respect, the HPC account is no different.²⁷

The metaphysical picture the HPC theorist paints of natural kinds lacks the demand for sharp edges that the essentialist endorses: it permits rough edges, it allows colors to blend or fade into one another, and its features are often obscured—but it is realist nonetheless. Whereas the essentialist offers a rigid and eternal structure of natural kinds integrated with the laws of nature, the HPC theorist's realism is centered on the more modest assertion that the world has some sort of causal basis, and that this will produce kinds. The causal features of the world are not always perfectly tidy, so it is no surprise to the HPC theorist that the natural kinds which emerge may be similarly dappled.

4 What Is a Disease?

To be able to talk about what makes two disease instances members of the same disease natural kind, and therefore what constitutes a disease natural kind, it is imperative we have a working concept of disease. But deciding what a disease is—that is, defending a particular disease concept—is a substantial undertaking, and one I lack the space for here. Consequently, I will rely partly on the concept of disease I have defended elsewhere and present a brief and slightly revised version of it here.²⁸

One important preliminary aspect of the disease concept I offer is that it is value-free.²⁹ According to this account, calling something a disease does not depend on our having any presuppositions about what is good or bad for an organism; hence, the present account is a version of what is sometimes referred to in the literature as a "naturalist" account of disease.

It is naturalist in that the standards against which diseases are understood are statistically determined, where the negative effects typically associated with a disease type are not effects that we deem to be bad, but rather those that reduce the organism's ability to deal with environmental pressures. Why is a value-free treatment of disease important? It matters because some objectors are bound to insist that a value-laden notion of disease undermines the suggestion that diseases can be understood as natural kinds. The argument might be something like this: if what counts as a disease depends on our judgments about what is a good or bad effect (a normal or irregular process, etc.) then what counts as a disease is up to us, and therefore we cannot claim that what constitutes a disease kind can be given by *nature* independently of our evaluative judgments. As it happens, the argument is fallacious: even if the extension of 'disease' depends on our value judgments it would not follow that specific diseases did not form natural kinds.³⁰ But since there are those who remain suspicious, it is simpler to make use of a value-free concept of disease, as I do, thus avoiding the worry altogether.

For our present concerns, the central feature of the account is that it considers diseases in contrast with naturally occurring processes, primarily at the cellular level. Groups of cells form causally interactive groups-"networks"—that collaboratively manifest various homeostatic dispositions. The cells themselves possess a wide range of dispositions, some small number of which are routinely manifested in response to the cells' environment, which is comprised of other cells within the network and various chemicals. When this environment is typical, the cellular dispositions that are manifested are similarly typical, and they give rise to standard cellular processes.³¹ That is, they result in the ebb and flow of molecules and energy, in and out of the cells, in the cycles of excess and shortage that we recognize as typical homeostasis. But not all the cells' dispositions are for these typical manifestations. Either as a result of problems within individual cells (e.g., cellular misinformation, DNA problems), or owing to environmental changes (e.g., chemical changes, increased or decreased energy, invasion of foreign bacterial or viral agents, chemical absences), cells will manifest other dispositions. In sufficient numbers, or if repeatedly manifested for a sufficiently prolonged period of time, the manifestations of these other dispositions give rise to processes that deviate significantly from the standard homeostatic processes of the cellular network. These distorted processes are disease processes.³²

The previous paragraph suggests two responses to the question of what a disease is. The first is that a disease is a process; specifically, a process of cellular network interaction that deviates significantly from the standard, where that process—like the standard processes from which it deviates manifests various cellular dispositions. According to this response, a disease type is to be identified with a type of disease process. The second response suggested by the above paragraph is that a disease is a set of dispositions, namely those dispositions whose manifestations result in a disease process. So which is it to be: are disease sets of dispositions or processes?

My answer is that they are both. Or, more correctly, that some diseases are best thought of as sets of dispositions, whereas others are best thought of as processes.³³ The difference, I suggest, turns on the sorts of dispositions involved and the circumstances of their manifestation. In the typical case—that which results in standard homeostasis—we have a subset of the cell's dispositions manifesting in response to a typical cellular environment. Here we can isolate two key components: the set of dispositions involved, and the conditions (circumstances) in which those dispositions are manifested. Call the set of *all* dispositions possessed by a typical cell S_A , and the typical circumstances in which the cell finds itself *c*. There is some subset of S_A , call it S_C , comprised of just those dispositions that are manifested in circumstances *c*, where the manifestation of the dispositions in S_C result in standard homeostasis.

I submit that a disease is best thought of as a process if it arises from a change in the conditions, such that some *other* subset of the cell's dispositions is manifested. Instead of the typical circumstances c, we have a different cellular environment—circumstances x—meaning that the subset of dispositions that are manifested is no longer S_C , but S_X , where S_X does not result in homeostasis.³⁴ The cell has not necessarily changed (these are all dispositions typical cells possess; they are members of S_A), but changes to the cell's environment have changed which dispositions are manifested. In cases where the triggering of these typical-for-the-cell-to-have-but-not-typical-for-the-cell-to-manifest dispositions results in a disease process, we have a case in which the disease in question is best considered a process.

On the other hand, a disease is best thought of as a set of dispositions in those cases where a disease process arises (if it ever does) from dispositions that are not typical for the cell to have.³⁵ That is, where the set of dispositions the cell possesses is not S_A . This can either be due to the addition of *novel* dispositions or because of the *absence* of typical ones.³⁶ In neither case does it matter whether the conditions (i.e., cellular environment) are standard or not, as it is the presence (or absence) of dispositions that matters for having the disease. However, as either change is bound to be more apparent if the dispositions in question are ones that manifest in typical conditions (circumstances *c*), we will only look at those cases.³⁷ If we change S_C by *adding* to it novel dispositions (those not typically found in S_A) whose manifestations arise in circumstances *c*, then we can expect that with sufficient change to S_C (one or more dispositions added) the resultant manifestations will no longer support homeostasis. In this case the disease is a matter of having certain atypical cellular dispositions. On the other hand, we can get the same effect if we *remove* a sufficient number of dispositions (one or more) from S_C , in which case the disease is a matter of lacking certain typical dispositions.

Some examples are bound to help. Classical phenylketonuria (PKU) is a genetic disease characterized by the inability to metabolize phenylalanine (phe), an amino acid found in breast milk, bread, meats, potatoes, and numerous other common foods. This inability comes from a mutation of the gene for phenylalanine hydroxylase (PAH), an enzyme required for metabolizing phe.³⁸ Because of the enzyme deficiency, the ingestion of phe results in increased levels of phe in the bloodstream, leading to such symptoms as impaired cognition and microcephaly. However, by maintaining a lifelong phe-free diet, persons with the PAH enzyme deficiency can avoid this process and the associated symptoms. Nevertheless, despite avoiding the process and the symptoms, people with the enzyme deficiency have the disease.

PKU is a disease type best thought of as a set of dispositions. Most notable within that set are the inability to produce sufficient PAH and the inability to metabolize phe. Combined with the conditions arising from normal diet, these inabilities would lead to a disease process and various symptoms, but it is the absence of these two dispositions that is significant for having PKU. As Lange rightly asserts, phenlyketonurics maintain phe-free diets because they have the disease, not to avoid it (Lange forthcoming). For those who need further convincing, consider a hypothetical therapy involving the use of symptom-suppressing drugs. Though a successful drug regimen can halt a disease process and prevent the production of symptoms, no-one would consider such a treatment a cure of the disease. The organism is still diseased, even if the drugs successfully block or counter the symptoms, as would become apparent if the drugs were no longer administered. In the case of PKU the disease process is prevented through a strictly controlled diet, but with a regular diet the problems would surface. A phe-free diet is thus not a cure for PKU: no disease process is required for the disease to be present.

Contrast the situation with PKU with that of scurvy. In humans, ascorbic acid (vitamin C) is required for the enzymes that synthesize collagen

to operate. As collagen is the most abundant protein in the human body it is found in tendons, artery walls, skin, bones, and so on—this is obviously a vital process. Nonetheless, humans do not produce ascorbic acid, and so require diets rich in vitamin C. As foods high in vitamin C are part of most typical diets, cases of persons significantly deficient in vitamin C are rare. But if and when this deficiency occurs, it results in a disease process that can involve such symptoms as spots on the skin, teeth loss, and bleeding from mucous membranes—and if untreated, it is fatal in all cases.

Despite the similarities between PKU and scurvy (viz., deficiency, control through diet), scurvy is a disease type best thought of as a process, while PKU is best thought of as a set of dispositions. This is the case even though scurvy results from a vitamin deficiency and various cellular dispositions. And the reason is simple: persons that do not produce ascorbic acid but maintain diets rich in vitamin C *do not have scurvy*. Only if one's diet is atypical—perhaps due to a lengthy trip at sea without adequate means for storing fruits and vegetables, when the lack of ascorbic acid means that dispositions typically manifested (e.g., collagen synthesis) are no longer manifested—does one have scurvy. If the process has not been initiated, then scurvy is not present.³⁹ Likewise, unlike PKU, scurvy is *cured* if the process is stopped (which can be achieved by consuming a diet rich in vitamin C). Scurvy is thus a disease type best thought of as a process.

This gives us the main picture regarding the ontology of disease, but two lesser details are worth pointing out. The first is that even in cases where the disease type is identified with a process, *that* the process occurs does not require or guarantee that the associated symptoms will result. Symptoms, though part of a disease process, are not to be identified with the process. The relationship between symptom and disease process is that of part to whole; the appearance of a symptom may be one terminus of a disease, or it can be one of the stages of the disease process. Consequently it is possible to have certain parts or stages of the disease process that occur in the absence of symptoms. Additionally, as most of us would expect, even if a process is allowed to continue, different conditions within the diseased individual can mean that only certain symptoms appear.

The second detail is that we need to distinguish diseases from other medical conditions that are sometimes spoken of as if they were diseases. These conditions are what I call "disorders." Disorders are purely *structural* deviations from the standard: broken arms, cataracts, hernias, and so forth. These are all medical conditions that may require treatment, sometimes even immediate and extensive treatment, but they are not diseases. They

are states of the organism that differ significantly from the canonical structure in terms of topography, shape, proportion, and number. Disorders will not figure directly in the account of disease kinds (it is their exclusion that is relevant); I mention them because disorders can sometimes be partial causes of diseases, and can often be symptoms, and to make it clear that to be a disease is not simply a matter of being something that gets or requires medical attention or treatment.⁴⁰

Though I have offered two accounts of disease types in terms of sets of dispositions and processes, it is not the case that what determines disease kinds will differ much between the two. In fact, we shall see that there is a good deal in common between disease kinds, both in terms of their nature and their constitution. With a basic picture of what it is to *be* a disease in mind, we can turn to the proposed account of disease natural kinds.

5 Natural Disease Kinds

I have suggested that essentialism runs into problems when dealing with disease kinds and, because it is desirable to treat disease types as natural kinds, we should look for an alternative account of natural kinds. The proposed alternative is a version of the HPC theory of natural kinds according to which kind-membership is a matter of instantiating some or all of some set of properties. Having provided a basic understanding of what a disease is, I now want to put forward a recommendation as to which properties matter for the determination of disease kinds—that is, what sort of properties make up the clusters for disease kinds.⁴¹

Before we can do this, there is a prior issue that must be considered. Both disease concepts I have offered make extensive use of the notion of a disease process. However, the natural-kinds literature focuses almost exclusively on natural kinds of substances, so extending talk of natural kinds to processes marks something of a departure. Consequently, something must be said about this treatment of processes as natural kinds. In doing so not only will it become apparent that treating processes as natural kinds is a welcome development of natural kinds theories; it will also help to highlight which properties are relevant to the determination of disease kinds (which properties make up the clusters), and why the two concepts of disease sketched above do not require separate treatments when looking at disease kinds.

Despite the prevalence of substance-based examples in treatments of natural kinds (essentialist and HPC accounts in particular), applying the natural-kind concept to processes is less of a departure than it may initially seem. Ellis and Lierse agree with this.⁴² In their brief discussion, Ellis and Lierse (and later just Ellis) not only claim that processes form natural kinds, but that this is in accord with the natural-kind tradition. Ellis provides the examples of physical processes like chemical reactions, radioactive decay, and osmosis, and biological processes like meiosis and mitosis (Ellis 2001, 162). It is also clear that the ways in which we study such processes as well as track them, learn about them, and attempt to control and manipulate them, demonstrate that we treat them as natural kinds. Even Mackie, in his discussion of Kripke and natural kinds, speaks straightforwardly of natural kinds of processes, and suggests that he thinks of disease types as natural kinds of processes.⁴³ I thus submit that treating processes as natural kinds is not only compatible with standard natural-kind notions but part of them.

Nevertheless, it might be objected that processes "are really just the ways in which things of different kinds are bound to behave, given their circumstances and the laws of nature," such that kinds of processes just supervene on substances and the laws of nature, and should not be treated as kinds in their own right (Ellis 2001, 163). For instance, Sulmasy adopts just this sort of approach in his understanding of disease types, arguing that diseases are not themselves natural kinds, but are had by and sometimes caused by natural kinds.⁴⁴ Ellis's response to this line of objection is to claim that it would only stand if (1) it could be shown that the processes are independent of the intrinsic behavioral dispositions of the substances, and (2) it could be explained why the laws of nature discriminate between substances as they do. I do not think we need go nearly that far. I cannot see why, even if a process is nothing over and above a substance's obeying certain laws of nature, this should preclude its being a natural kind. That processes depend ontologically on substances, and perhaps only arise in accordance with the laws of nature, is orthogonal to the question of whether they form natural kinds: nothing about being a natural kind demands the members be fundamental existents. Consequently I cannot take seriously the thought that there cannot be natural kinds of processes.

What is it then for two processes to be of the same kind? An obvious suggestion is that they follow the same pattern: they should have a similar initiating cause, follow a similar progression, and result in some similar outcome. We would have few problems treating two such processes as processes of the same kind. Yet it would be asking too much to demand that any two instances of the same kind of process must run their full course. Two processes can be members of the same process kind even if one should fail to run to completion. Consider a simple chemical example: a small sample of gold is submerged in an adequately large quantity of aqua regia and begins to dissolve. Given sufficient time, the gold will go into solution. Call the process that would occur if the gold and solution were left until the gold had fully dissolved the "completed" process. Now imagine a similar case in which the procedure is interrupted, and the (now slightly smaller) sample of gold is removed while there is still something left of it. This is surely the same kind of process as the completed one, only now it has been cut short. So what makes it the same? First, that it has a similar initial cause (viz., immersion in aqua regia); second, that it has at the outset a similar set of dispositions (concerning the gold's solubility and the reciprocal dispositions of the solvent) regardless of whether they are manifested or not; and third, it has a similar terminus (viz., no more solid gold). In the completed process the dispositions in question are all manifested; in the truncated process only some are, but the rest are present and would have been manifested were it not for the interruption.⁴⁵ Likewise, the end result would have been the same had the process not been interrupted.

Another example, far more familiar than the last, is the biological process of human pregnancy. It has a well known set of initial causes (you can do it the old fashioned way or enlist the aid of a fertility specialist); it manifests certain dispositions and produces new ones; it has a typical course that runs around nine months, and ends with the appearance of a tiny human. That is the "completed" process that we typically desire—but shorter processes with other outcomes nevertheless constitute processes of the same natural kind. Notice how easily we recognize instances of this natural kind, despite the different ways in which the process can (now) be initiated and the variety of ways in which the process can unfold.

As biological processes, disease processes follow a similar basic pattern to what we find in the case of pregnancy. A disease process has some sort of initiating cause (either the triggering of atypical dispositions by atypical circumstances, or altering the subset of dispositions which is triggered by typical circumstances), it has some course that takes place within the organism, and if uninterrupted it will result in some set of symptoms. And just as we recognize that pregnancies can unfold in a variety of ways without failing to be members of the same process kind, disease processes can develop in a variety of ways—including abrupt interruptions and successful efforts to halt them—and still be members of the same diseaseprocess kind. And though there are many ways that a particular kind of process might unfold, experience teaches us that the processes we treat as natural kinds tend to proceed along inductively reliable lines. This speaks to the homeostatic unity of the properties involved, and to the naturalness of the ensuing cluster of properties.

To repeat, two processes do not fail to be of the same kind just because one is cut short. There has to be some limit to just *how* short an instance of a process kind can be and still be counted as a process at all. I lack any good sense of just where that line should be drawn, but that hardly matters here. What does matter is the set of dispositions that are manifested and will continue to manifest if a disease process is untreated. In short, when it comes to comparing disease instances, what matters is the set of dispositions involved. And this is why, when thinking of disease kinds, it ultimately makes little difference if the kind falls under the process concept or the set-of-dispositions concept: in both cases what matters is the set of dispositions involved.

Considering potential disease processes brings to light three ways in which instances of disease processes can be treated as similar: the first concerns the cause, the second the set of dispositions that would manifest themselves in the disease process, and the third involves the symptoms arising during (and at the end of) the disease process. All three ways lend themselves to specific property-types and will be applied in the determination of disease kinds. That is, all three types of property will be included (one way or another) in the cluster of properties that determine disease-kind membership. Let us take a closer look at this, starting with the causes.

Central to the essentialist's account of disease kinds is the thought that causes of diseases are involved in determining disease kinds. To this extent at least, the essentialist is surely correct: differences in causes can often mark the boundary between different disease natural kinds. As we have seen, no specific cause is necessary for a given disease kind, as many diseases have more than one cause (recall that tuberculosis can result from infection by *Mycobacterium tuberculosis* or infection by *Mycobacterium bovis*); conversely, many causes are not sufficient either (e.g., infection of an immunized organism will not result in a disease process). Therefore, though we should follow the essentialist in treating etiological features of diseases as important to determining disease kinds, we must be more liberal in their inclusion. One way to do this is to include within the cluster of properties all of the potential causes of a disease (note that this will still be a very limited set of causes): this cluster would capture all of the right instances.

Hence etiological features will figure importantly in the property clusters that define kinds. These properties—perhaps best construed as relational properties—will vary slightly between the two disease concepts. Where the disease type is identified with a type of process, the salient aspects of the cause are the conditions (circumstances) that trigger the atypical dispositions. Recall that the dispositions in question are not unusual for the cells and cellular networks to possess, but that it takes atypical conditions for these dispositions to be manifested. These will tend to involve changes in the cellular environment, either because other cells in the network are acting differently, or because there is an increase or decrease in this or that chemical in the cellular environment.

For those types of disease that are understood as sets of dispositions, the casual properties are those responsible for the presence or absence of the typical dispositions. Unlike the process diseases, what matters for dispositional disease types are deviations from the set of properties typically possessed by the cell and cellular networks; whatever affords this difference is the causal property we are interested in. In many cases this will be a genetic issue: various cells may have an atypical complement of dispositions owing to genetic inheritance. But it can also be the case that cells that start out normal are changed somehow, perhaps via environmental insult or lesion. These are the causes we are interested in.

The next group of properties included in the cluster are the dispositions for the disease process. We have seen that the essentialist eschews the thought that surface or stereotype properties could be used to determine kind-membership. The essentialist argument against the inclusion of such properties relies on the possibility of members of a kind failing to exhibit one or more of the surface properties associated with that kind. Samples of gold, for instance, can fail to be shiny or yellowy: hence surface properties are not necessary for membership in the kind. Nor are they sufficient: fool's gold exhibits surface properties of gold but is not a member of the kind. In the case of disease kinds, essentialism requires that the set of signs and symptoms associated with the disease type not be included among the set of properties that determine the natural kind. For instance, Putnam writes: "we are prepared to classify sicknesses as cases of multiple sclerosis, even if the symptoms are rather deviant, if it turns out that the *underlying condition* was the virus that causes multiple sclerosis" (Putnam 1975, 311).

Contrast this picture with that we find we find in contemporary disease classifications. If one inspects the World Health Organization's most recently adopted version of the International Classification of Diseases (ICD-10, formally endorsed by the WHO in 1990), one finds that one of the primary classificatory criteria is symptom-based.⁴⁶ When it comes to practice, then, and for practical purposes, the essentialist criteria are not adhered to. As one might expect, displaying most or all of the symptoms associated with a particular disease kind is strong evidence that the disease in question *is* a member of the kind. After all, let us not forget that the symptoms we associate with a disease tend to be found together as a result of natural processes; they do not arise individually and are not randomly collected or accidentally grouped. Distinguishing diseases according to their associated symptoms is clearly both beneficial and desirable, so what should be done?

Fortunately, in rejecting essentialism we also reject the essentialist's insistence that the only properties that determine kind-membership must be instantiated by every member of the kind. In the essentialist's thought experiment we are asked to imagine a member of the kind that fails to exhibit this or that surface property. The essentialist takes this as evidence that the property in question is not important for determining membership in the kind. But all it actually shows is that the property in question is not *essential* for being a member of that kind. It does not show that it is not a frequent or important indicator, nor does it show that only essential properties should be used to pick out the kind. In the debate between the "surface" properties and the "essential" properties, the HPC kinds theorist is not forced to choose—all such properties can, and should, be included in kind-determination, at least to some extent.⁴⁷

However, consider a disease process that is cut short, perhaps one that is cut off at a very early stage indeed, and well before any of the diagnostic symptoms has arisen. For instance, imagine that we have an instance of RA, but that before the disease can lead to any joint discomfort and so on, a wonder cure is ingested. We have already said that a disease process does not need to be very long to count as a disease process, so this case will count.⁴⁸ But there are no symptoms. Hence, if symptom-type categorical properties were *required* for kind-membership, then their absence in this case would mean that this was not an instance of RA, contrary to what we have imagined. But I submit that this is yet an instance of RA, and a central one: one wonders what we were doing administering the wonder cure if it were not (assume that the wonder cure is incredibly expensive and in short supply).

Consider a second case, wherein the RA disease process has gone much farther and symptoms start to arise. Lacking the wonder cure, a slightly less wonderful symptom-blocker is administered. The dispositions remain, and perhaps that portion of the disease process that does not include the symptoms continues (if there is any such part of the process), but the patient suffers no obvious ill effects. But beware, if her insurance runs out the symptoms will return—this is *not* a cure! This case has a few more symptom-type categorical properties we can point to, but still too few if all the symptoms are required to be a member of the kind. And yet it seems as central an instance of the disease process as any (or what were we preventing?).

As a final example consider two different instances of RA, where each instance displays exactly four of the seven ARA diagnostic symptoms such that the two instances have only one symptom in common. These ought to be two "peak" instances of RA, but the reduced number of symptoms and the lack of overlap suggest that at most one can be a peak instance, and that one or both are at least partway down the hill. This is yet another version of the same problem.

The solution is to reject the thought that symptoms contribute to the determination of disease kinds, opting instead for properties that even the nonsymptom cases generally possess: the *dispositions to produce the symptoms*. These are present—at least in one form or another—in even the shortest instances of the disease process, and in the set of disease types that fall under the set of dispositions concept.

That is not to suggest that the dispositions in the cluster are all dispositions that could immediately manifest themselves in symptoms, as some of the immediately-for-symptoms dispositions will require other dispositions to be manifested before they can then be manifested (these are "higher order dispositions"⁴⁹), but through the more basic dispositions they will help to determine kind-membership. If the disease process is initiated, and left to run its course, the symptoms will arise. Of course, other factors concerning the general health of the host organism will dictate the extent and sometimes severity with which the dispositions for symptoms manifest, but these are differences that the account of disease kinds is right to ignore. I submit that most of the differences between instances of the same disease kind, such as those in the four-out-of-seven RA case, can be explained by differences specific to the patient. In sum, by having dispositional properties in the cluster, we are not forced to treat the above problem cases as non-central instances of the kind, and we can maintain the role of symptoms in determining disease types (even if only indirectly⁵⁰).

As was suggested above, I suspect that dispositional properties play an important role in any natural process kinds, diseases or otherwise. Consider the pregnancy case: a woman who has recently conceived shows almost no signs of the process that has begun. But the process has begun, and a series of dispositions are present that may be manifested in due time. Of course, there is no guarantee that these dispositions will be manifested, as they can only arise if the correct conditions for their manifestation obtains—but the potential is there. As the process unfolds, some of these dispositions will be manifested (more correctly, the unfolding of the process *just is* the manifesting of these dispositions), and new dispositions will be generated.

The moral here is that most of the properties in the cluster will be dispositional in nature. They are dispositions for disease processes, and therefore include dispositions for the symptoms associated with the disease. These are the most important properties for the determination of disease kinds.

The final class of properties within the cluster are relations that pertain to the primary location of the dispositional properties that give rise to the disease process (if or when it is initiated), where location is understood in two different ways. The first of these concerns the particular site within the organism where the dispositions are located, and where the disease process tends to occur. That is, we get distinct disease kinds roughly on the basis of which organ or organ system the process occurs in. No disease process will ever be entirely contained within one area—disease processes engage with the organism's other processes (e.g., passing infection through the bloodstream), and some disease types have no specific location at all—but for many types there tends to be a primary site. In cases where a disease process spreads, distinctions will tend to be a feature of where the process begins; that is, those cells or networks that have the dispositions in question. This could be part or the whole of a single organ or organ system, or even the organism as a whole. Hence, though they are similar in a number of respects, we make a distinction in kinds between such diseases as bladder cancer and lung cancer.

For certain disease types it might turn out that the inclusion of locational properties is redundant. This will be the case when the set of dispositions already reflects the locale where they are found. For instance, it may be that the disposition to x is only found at some specific location l, removing the need for additional properties that spell out the location. Hence we might be able to make do without this third category of properties in the cluster. However, as there is at least the logical possibility of disease kinds that perfectly overlap in terms of their initiating causal relations and dispositions, but differ only in location, the locational properties are best included, even if they might duplicate information already contained within the cluster.

Like the first, the second locational aspect is something that might already be contained with the dispositions in question, but which we shall include regardless. This second notion of location concerns the class against which the standards are determined. This will include coarse-grained distinctions about the ages of the comparison class of the diseased organism, as well as the environment of the comparison class. As might be expected, the latter plays some role in determining such things as diet, pollutants, stress, and so on, that affect the class and therefore have some effect on the standard processes in which the body engages. The former concerns a comparison with regards to general age groups-something to the effect of infants, children, adolescents, and so on. The point here is similar to that made above: because diseases are dependent entities, partly constituted by the cellular processes of the organism that bears them, significant differences in the nature of the host organism make for similarly marked differences in the diseases themselves. And as the healthy cellular processes differ across these age groups and regions, so too will the disease kinds.

6 Conclusion

I suggest that this is how we should think of natural disease kinds. Following Boyd, it is important that we allow room for imprecise boundaries, and not attempt to sharpen them artificially, even if most of the time they are naturally disjoint. The human body is a vastly complex system, often lacking in sharp cut-offs and ceaselessly changing over time. Consequently it would be foolish not to permit some variation and indeterminacy in speaking of disease kinds: in this way we can preserve our intuition that diseases constitute natural kinds.

Notes

1. The question here and throughout this essay is whether the instances of a *specific* disease type form a natural kind. That is, do all the instances of x (where x is some type of disease, such as rheumatoid arthritis or tuberculosis) form a natural kind? This is not to ask whether all the various types of disease together constitute a single natural kind. For more on the latter discussion see D'Amico 1995 and Reznek 1987, 1995.

2. See Putnam 1975, 241; 1975, 311, and Mackie 1976, 99, for statements to this effect. When essentialists speak of "the" cause of a disease instance, what they really mean is "that causal factor which stands out against the causal field"; any effect will have numerous causes, but we might think of "the" cause as perhaps the most

salient among them. For convenience I too will speak in terms of "the" cause of a disease instance, with the understanding that it is the most salient among the causes. See Mackie 1965.

3. A common feature among the stock of success cases is that they tend to be diseases caused by an identifiable disease agent such as a virus, parasite or bacterium; diseases whose causes are immunological, intra-cellular, genetic or deficiency-based tend not to make the essentialist's list. The difference can be explained, in part, by the tendency of those who use these cases to focus on disease treatment and prevention, combined with the relative ease with which we single out a disease agent as *the* cause in the success cases (from among the various causes at work in any given instance), in contrast to the difficulties experienced in the latter cases.

4. These are: (i) morning stiffness; (ii) arthritis of three or more joint areas; (iii) arthritis of hand joints; (iv) symmetric arthritis; (v) rheumatoid nodules; (vi) serum rheumatoid factor; (vii) typical radiographic changes (Arnett et al. 1988).

5. Given the huge variability in human genetic constitution, level of nutrition, state of immune system, environment, and general health, it would not be at all surprising if similar causes are capable of producing instances of different disease types in different people.

6. I am assuming here that: (i) disjunctive essences are eschewed by essentialism, and (ii) all instances of tuberculosis, regardless of cause, admit of such overwhelming similarity that having two causes does not constitute an adequate reason for thinking that there are two kinds of tuberculosis. As I interpret it, essentialism requires that the causal role must be uniquely satisfied, and so cannot continue to be indicated vaguely along the lines of "whatever happens to cause the disease" when we know exactly what *they* are. In this way disease kinds follow what Putnam says of "jade"-type cases: the discovery that instances of jade have *two* microstructures (jadeite and nephrite) producing the same unique textural qualities makes for two kinds, not one. And by extension, "if H_2O and XYZ had both been plentiful on Earth, then we would have had a case similar to the jadeite/nephrite case: it would have been correct to say that there were *two kinds of 'water'*" (1975, 241).

7. In practice its use remains ambiguous between definitions that treat the presence of tubercles as contingent or not, and that similarly treat having-been-caused by the tubercle bacillus as contingent or not (Flier and De Vries Robbé 1999). It is worth noting that disputes over the use of 'tuberculosis' do not directly undermine the essentialist's metaphysical treatment of kinds, but something about the essentialist picture is clearly unsatisfactory to a significant number of medical practitioners.

8. There are also certain social issues that arise. Though largely fallacious, it is not uncommon for various institutions (legal, governmental or financial) to confer on those diseases that are less obviously "natural" a diminished status, and use this as a basis for allocating similarly diminished resources.

9. Could we not just rely on artificial or conventional groupings of disease instances to support our practices (in the absence of natural groupings)? No. It is the "naturalness" of natural kinds that justifies what Goodman (1954) dubbed 'projectability': the inference that the properties of a subset of the members of a kind would apply to the remainder. Though this inference is not infallible, if a grouping is entirely up to us we have no basis for the inference at all. See Mill 1884, bk. 4.

10. Why replace the essentialist picture—why not just have *many kinds* of natural kinds? For the most part this would be sage advice: a single notion of natural kind applied without revision to all the various disciplines and subdisciplines that utilize natural kinds would contradict most of what each discipline would otherwise have taken to be natural kinds. However, when the domain is restricted to disease, I can see no benefit to having multiple notions of natural kinds: this would undermine most of our medical practices and make the sharing of medical information a near impossibility.

11. Boyd 1989, 1991, 1999. See also Griffiths 1999, Millikan 1999, and Wilson 1999.

12. I am also naively optimistic that the present treatment will be applicable to *botanical diseases*, but it would take someone with greater botanical knowledge than I possess to tell me if that is remotely possible.

13. A number of prima facie objections can be avoided by keeping this distinction in mind while considering what follows.

14. This picture of prevention, cure, and diagnosis is indicative of the ideal cases that is, those in which we have a strong understanding of the disease type in question. In cases where we have limited understanding of a disease, or very poor information about a patient's condition, we are forced to take other forms of action, though I submit that even these other actions suggest that we are thinking of instances of diseases as forming natural kinds.

15. This is not to claim that we get it right even close to most of the time, nor that our diagnoses are frequently on target. It is simply a claim about what is going on when we approach treatment, prevention, diagnosis, etc., in this way, and what seems to be the case when we are successful.

16. This would be unfortunate because it would mean that even our successes in treating diseases would not provide as much information as they seem to do.

17. This is the response given in Dupré 1981.

18. Boyd is leading this reclamation within the biological sciences; see his 1989, 1991, 1999. Though I endorse the HPC account of natural kinds for some sorts of natural kinds, I remain neutral regarding the question of whether species should be understood in terms of HPC kinds.

19. The account of the HPC theory that follows is a modification of that found in Boyd 1989, 1991, 1999. One notable difference is that the present version omits the conventionalist and constructivist undertones suggested in Boyd's formulation— undertones that tend to compromise the extent to which his HPC account is a realist account of kinds.

20. The form of essentialism defended in Ellis and Lierse 1994 and Ellis 2001 is constructed directly in opposition to this wider account of laws, but they would be the first to admit that the common features of "traditional" essentialism are as I have presented them, which is all I am interested in.

21. This is not to suggest that a functional kind-member like a table is not subject to the laws of nature, it simply does not obey those laws *qua* 'table'. (Some other natural kinds, perhaps those that constitute the table, are presumably at work: electrons or molecules or something of the sort.)

22. This is very rough indeed. HPC theorists generally, and Boyd particularly, tend to give us very little indication of what exactly belongs in the cluster. I aim for greater clarity with regards to HPC disease kinds; see sec. 5 below.

23. This is not to suggest to that HPC kinds are always fuzzy, or that sharp cut-offs cannot exist in nature, but only that sharp cut-offs are not required for establishing kinds. More often than not HPC natural kinds *will* have fairly sharp boundaries; what matters—and what distinguishes them from the essentialist's kinds—is that sharp boundaries are not required, and that the kinds are shaped through natural causal processes.

24. Boyd also suggests that the indeterminacy of HPC kinds is partly a product of our own inductive and explanatory practices through which *we* engage with the causal nature of the world. As these practices are themselves somewhat inexact (relying on generalizations from test scenarios, laboratory conditions, sampling, etc.), and themselves involve causal processes, Boyd claims that the kinds that emerge will be similarly imprecise.

I think that Boyd's suggestion here is overly constructivist and threatens the naturalness of kinds that justifies our inductive and explanatory practices. It is therefore important to keep separate the disjointness that arises in our *picture* of the kinds from any disjointness that naturally arises from the *interaction* of substances in the world. The former is just good old healthy fallibilism but, if not separated from the latter, compromises the realism of the HPC kinds.

25. Boyd writes: "the presence of some of the properties in F tends (under appropriate conditions) to favour the presence of the others, or there are underlying mechanisms or processes that tend to maintain the properties in F, or both" (1999, 143).

26. I think it is fair to say that (traditional) essentialism carries with it a number of neo-Humean metaphysical commitments. The essentialism favored by Ellis and Lierse (1994) rejects this aspect of essentialism.

27. The form of realism presented here for HPC kinds marks a departure from the realism Boyd depicts them as having. Boyd's understanding of "realism" is captured by the following: "When we ask about the 'reality' of a kind or of the members of a family of kinds—or when we address the question of 'realism about' them—what we are addressing is the question of what contribution, if any, reference to the kind or kinds in question makes to the ways in which the classificatory and inferential practices in which they are implicated contribute to the satisfaction of the accommodation demands of the relevant disciplinary matrix" (1999, 159).

28. See Williams 2007. A number of the concepts I use here in presenting the account of disease (e.g., homeostasis, cellular dispositionality, standard conditions) get much a fuller treatment there. Mention of "homeostasis" as applied to cellular processes is not to be confused with the homeostasis that forms HPC kinds, though there is bound to be occasional overlap.

29. That is, as value-free as any physical or biological science happens to be, in the sense of being independent of our evaluative judgments.

30. Reznek (1987, 1995) argues that 'disease' does not name a natural kind, because he believes the concept to be value-laden; however, he nevertheless admits that individual diseases could turn out to be natural kinds.

31. What counts as "standard" is a statistical matter, determined within the relevant comparison classes. Here I follow Boorse 1975, 1977, 1997.

32. For the sake of brevity I have passed over two other important conditions that are necessary for the distorted process to count as a disease process: the first is that the cellular network is incapable of remedying itself (without producing further distortions in other networks; and where those further networks are either incapable of self-remedy without distorting some further network, etc.), and the second is that the process tends to reduce the organism's ability to cope with environmental pressures.

33. My current understanding of 'disease' differs slightly from that I defended in 2007. Whereas I previously argued that the dispositions for the nonstandard manifestations had to be manifested in order for the disease to be present in a person (and so diseases were to be identified with certain kinds of processes), I now believe that *some* diseases are such that the mere presence of certain dispositions, even if they are never manifested, suffices for the having of the disease, and so *some* diseases are dispositions. My revised view has been influenced by discussions with Barry Smith and Marc Lange. (Thanks to Lange for making unpublished material available to me; see Lange forthcoming.)

34. It is an empirical question how much S_x must differ from S_c before homeostasis is no longer maintained, the answer to which may plausibly vary for different cell types. For ease of exposition we shall assume the extreme case in which S_x and S_c

are disjoint, and no disposition in $S_{\rm C}$ is such that it will be manifested in circumstances x.

35. It is of course not required that the disease process *in fact* arise, for in these cases having the relevant dispositions is sufficient for having the disease; but it must be the case that the process that *would result* be a disease process, as outlined above.

36. The absence of a disposition for m in some circumstances e is just to have a disposition for *some manifestation other than* m (and that does not include m as a part) in circumstances e; hence the absence of a disposition is just the having of some other disposition (for those same circumstances). In what follows I will take this as understood and speak of absences of dispositions as dispositions.

37. Having novel dispositions for disease processes that only manifest themselves in atypical circumstances is still to have a disease, as is lacking dispositions that manifest in atypical circumstances, but on the assumption that things are typical most of the time, these diseases are far less likely to get noticed.

38. Lange (forthcoming) discusses PKU at length, arguing that PKU is an incapacity (a type of disposition), not a process. Lange considers only PKU, asking what sort of disease concept suits it best, but makes no claim about other disease types. As far as I can tell, the dispositional concept does not apply to all disease types: some are best thought of as dispositions, others are best thought of as processes. That said, all disease types will involve dispositions and have associated processes and symptoms; the difference is whether or not the process must be present to be an instance of the disease.

39. What of the reply, on behalf of someone who takes all diseases to be dispositions, that we all have scurvy all the time? The answer is that this is a desperate move, and the kind of thing one would only suggest if in the grips of a theory. The rest of us know better; I do not have scurvy, nor do I suffer from numerous other diseases (and death) that would arise were I to stop eating.

40. Consequently it is clear that definitions of 'health' that treat it as merely the absence of disease are way off the mark.

41. I will say little about the homeostasis that produces and maintains the property clusters. The basic picture is one of cellular processing and of interactive networks of cells, in addition to various environmental conditions; the ways these operate within the body provide the conditions that support disease. For more on these processes and their connection to disease see Williams 2007.

42. Ellis and Lierse 1994; Ellis 2001. Despite sharing their enthusiasm for natural kinds of processes, I reject their essentialist treatment of them.

43. He says first that "[i]f our archetypes or typical specimens of rusting are in fact oxidation of iron, then if any process, however superficially like rusting it was, were not the oxidation of iron it would not be rusting," and then following his claim

that the essence of malaria is the malarial parasite and the essence of measles is the measles virus, that "[w]hat such examples show is that it is not the difference between substances and non-substances that matters here, in the sense of the distinction between items which are supposed to 'subsist by themselves' and items which are not, but rather the difference between cases where it is useful or fruitful to think and speak preferentially of a possibly unknown or inadequately known 'nature'" (Mackie 1976, 99–100).

44. Sulmasy 2005. For the most part we can ignore the details of Sulmasy's account. He claims that diseases are states of affairs—which I reject. But more importantly, his argument against treating diseases as natural kinds is that they lack essences. Naturally I agree, but endorse an altogether different response.

45. For the sake of convenience I am overlooking all of the difficulties that arise when speaking counterfactually about dispositions and their manifestations. I should not, however, be interpreted as suggesting a counterfactual analysis of dispositions; I am merely pointing out what, *ceteris paribus*, would tend to occur.

46. The ICD-10 can be accessed at <http://www.who.int/classifications/icd/en>. As Kendell (2001) points out, the ICD-10 classifications are steeped in convention, to the extent that the same disease type is classified as a "different" disease type if more than one specialist treats it within his or her specialty. However, the inclusion of symptoms in determining disease kinds does not demand such conventionalism, and I suggest it be avoided.

Incidentally, among the "diagnostic issues to consider" in the workgroups that developed the ICD-11, it is stated that for the "[c]lustering of signs, symptoms, and operational characteristics," workgroups were to try to "identify the features that are necessary and sufficient to define the disease/disorder" (Üstün et al. 2007). I chalk this up to residual essentialist thinking, and conflating the privileging of certain criteria for diagnostic purposes when they are more reliable indicators—an epistemic matter—with that of metaphysical kind determination.

47. This is particularly beneficial in the case of syndromes, where the underlying cause is unknown, as is the case with RA. Assuming the syndrome does in fact have some as of yet unknown etiology, what we have is a cluster of symptoms that arise together as a result of the causal nature of the world. It is precisely for this reason that we group the symptoms together as a syndrome and treat them jointly, rather than treating them severally.

48. I am assuming that RA is a disease type of the process variety. If this is false, then the relevant dispositions need not be manifested at all.

49. For more on higher-order dispositions and their role in possibility see Williams and Borghini 2008.

50. This assumes—uncontroversially, I should think—that the presence of a symptom indicates that the disease previously had (and may continue to have) the

disposition to produce that symptom. Hence symptoms serve as guides—very good guides—for the dispositions that are (or were) present.

References

Arnett, F., S. Edworthy, D. Bloch, et al. 1988. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis and Rheumatism* 31:315–324.

Boorse, C. 1975. On the distinction between disease and illness. *Philosophy & Public Affairs* 5:49–68.

Boorse, C. 1977. Health as a theoretical concept. Philosophy of Science 44:542-573.

Boorse, C. 1997. A rebuttal on health. in *What Is Disease?* ed. J. M. Humber and R. F. Almeder. Totowa, NJ: Humana Press.

Boyd, R. 1989. What realism implies and what it does not. Dialectica 43:5-29.

Boyd, R. 1991. Realism, Anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies* 61:127–148.

Boyd, R. 1999. Homeostasis, species, and higher taxa. In *Species—New Interdisciplinary Essays*, ed. R. Wilson. Cambridge, Mass.: MIT Press.

Caplan, A. L., H. T. Engelhardt, Jr., and J. J. McCartney, eds. 1981. *Concepts of Health and Disease—Interdisciplinary Perspectives*. London: Addison-Wesley.

D'Amico, R. 1995. Is disease a natural kind? *Journal of Medicine and Philosophy* 20:551–569.

Dupré, J. 1981. Natural kinds and biological taxa. Philosophical Review 90:66-90.

Dupré, J. 2002. Humans and Other Animals. New York: Oxford University Press.

Ellis, B. 2001. Scientific Essentialism. New York: Cambridge University Press.

Ellis, B., and C. Lierse. 1994. Dispositional essentialism. *Australasian Journal of Philosophy* 72:27–45.

Flier, F., and P. De Vries Robbé. 1999. Nosology and causal necessity: The relation between defining disease and discovering its necessary cause. *Theoretical Medicine and Bioethics* 20:577–588.

Franklin, L. 2007. Bacteria, sex, and systematics. Philosophy of Science 74:69–95.

Goodman, N. 1954. *Fact, Fiction, and Forecast*. Cambridge, Mass.: Harvard University Press.

Griffiths, P. 1999. Squaring the circle: Natural kinds with historical essences. In *Species—New Interdisciplinary Studies*, ed. R. Wilson. Cambridge, Mass.: MIT Press.

Hacking, I. 1991. A tradition of natural kinds. Philosophical Studies 61:109–126.

Kendell, R. E. 2001. The distinction between mental and physical illness. *British Journal of Psychiatry* 178:490–493.

Kripke, S. 1980. Naming and Necessity. Cambridge, Mass.: Harvard University Press.

Lange, M. Forthcoming. The end of diseases. Philosophical Topics.

Mackie, J. L. 1965. Causes and conditions. *American Philosophical Quarterly* 2:245–264.

Mackie, J. L. 1976. Problems from Locke. New York: Oxford University Press.

Mill, J. S. 1884. A System of Logic. London: Longmans.

Millikan, R. 1999. Historical kinds and the special sciences. *Philosophical Studies* 95:45–65.

Newman, S., R. Fitzpatrick, T. Revenson, et al. 1996. *Understanding Rheumatoid Arthritis*. New York: Routledge.

Putnam, H. 1975. Mind, Language and Reality. New York: Cambridge University Press.

Reznek, L. 1987. The Nature of Disease. New York: Routledge.

Reznek, L. 1995. Dis-ease about kinds: Reply to D'Amico. *Journal of Medicine and Philosophy* 20:571–584.

Sulmasy, D. 2005. Diseases and natural kinds. *Theoretical Medicine and Bioethics* 26:487–513.

Üstün, T., R. Jakob, C. Çelik, et al. 2007. Production of ICD-11: The Overall Revision Process. Accessed through the World Health Organization website. http://www.who.int/classifications/icd/ICDRevision.pdf.

Williams, N. E., and A. Borghini. 2008. A dispositional theory of possibility. *Dialectica* 62:21–41.

Williams, N. E. 2007. The factory model of disease. Monist 90:555-584.

Wilson, R. 1999. Realism, essence, and kind: Resuscitating species essentialism? In *Species—New Interdisciplinary Studies*, ed. R. Wilson. Cambridge, Mass.: MIT Press.

11 Predicting Populations by Modeling Individuals

Bruce Glymour

1 Introduction

Despite their many differences, the so-called *dynamic* and *statistical* interpretations of evolutionary theory (ET) share a common understanding of that theory. ET is, among other things, a theory about how frequencies of types change in populations, and these changes are recorded as changes in the values of variables measured on populations. ET explains such changes, at least in part, by appeal to natural selection. Both dynamical (Bouchard and Rosenberg 2004; Stephens 2004) and statistical (Walsh, Lewens and Ariew, 2002) interpretations of ET take selection to be represented by equations whose state-variables are measured on populations. The equations are, up to changes in parameter values, more or less constant across populations, both in the particular state-variables employed and in the functional form of the mathematical dependencies between them. In this sense, then, both interpretations take ET to be a population-level theory: the theory explains properties of populations by appeal to other properties of populations, and the dependencies, nomic or otherwise, between these properties are more or less invariant over different populations. The interpretations differ essentially only in their understanding of the fitness parameter and its putative causal role.

That shared vision is inherited from an earlier generation of interpretations, advanced by the likes of Sober (1984), Rosenberg (1983), Brandon (1978), Mills and Beatty (1979), and Sterelny and Kitcher (1988). These philosophers took evolutionary outcomes to be explained by one set of equations, largely invariant over distinct populations, which equations express formally the way in which natural selection influences genic and genotypic frequencies. Then, as now, the equations in question are those of population genetics. Then, as now, selection was thought to be measured by fitness differences. Then, as now, the crucial interpretive differences among philosophers lay in their understanding of fitness. And then, as now, the assumption that ET is and ought to be a populationlevel theory went unchallenged.

This shared understanding of ET has both an historical, and therefore accidental, motivation and a principled motivation. My concern here is with the latter. It is and was common knowledge that the range of causes of survival and reproductive success is legion. Worse, these causes differ from population to population. Of the many causes of survival and reproductive success in *Homo sapiens*, for example, very many are *not* causes of survival or reproductive success in *Pan troglodytes*, and even fewer in *Drosophila melanogaster*. To all appearances, there are so many potential and varied causes of survival and reproductive success that there is simply no hope of cataloging them all. These bits of common knowledge are, to a first approximation, correct.

But common knowledge goes further. According to the received wisdom it follows from the foregoing that any interpretation of ET on which it is a general theory, applicable to any biological population you like, will have to abstract from the particulars, the vicissitudes, by which survival and reproductive success are generated. That kind of abstraction requires that the relevant state-variables in any formal representation of the theory be measured on populations, for while the causes of *individual survival and* reproductive success in any two populations may differ, at the appropriate level of abstraction the causes of changes in genotypic frequencies might nonetheless be the same. Hence, ET must be a population-level theory: since the causes of individual behavior are so varied, any hope for a general theory requires that it be formulated at the population-level. Or so the story goes. I claim the received wisdom is, in these latter respects, mistaken. In this essay I will explain why, and adumbrate two consequences that follow from the errors. Those who take the consequences seriously will find in them ground for a general rule for determining when and why reductive theories are necessary.

2 The Scientists' Problem, and Two Strategies for Solving It

Population biologists face the following problem. They observe a collection of organisms, a sample of a population, over some period of time. At any given time, organisms in the population differ in their phenotypic properties, that is, one or more phenotypic variables exhibits non-zero variance. What is more, the frequency distribution over these variables changes over time. The problem is to predict the future frequency distributions from current and past observations. The population geneticist faces essentially the same problem, except that the variables of interest track allelic or genotypic properties. Succinctly, the scientist wants predictions about populations but has evidence only about individual members of the population.

One standard procedure for solving this kind of problem, which we will call the population-level strategy, goes like this. Find an equation-type which writes the next generation (or time-step) frequency distribution as a function of the current generation frequency distribution. The finding is a bit of model selection—one has to choose the state-variables and the functional form of the mathematical dependencies between current and future values of the state-variables so that reliable prediction is possible. The hope is that the resulting model will be applicable to most or all populations. Once a model has been chosen, one identifies (estimates populationspecific parameter values for) the model for the particular population whose behavior one cares to forecast.

What is distinctive about these models is that the state-variables are measured on populations, and the model is held to be generally applicable to any population in the domain, at least to a first approximation.¹ That is, while parameters may change from population to population, the same state-variables appear, and are related by the same equations, up to a change in parameter values. This can be seen in both diffusion approximations such as:

$$P(x,t) = -\frac{1}{2} \frac{\partial}{\partial x} (V_{\delta x} \varphi(x,t)) + M_{\delta x} \varphi(x,t)$$
(1)

and in the more standard formalisms of textbook population genetics:

$$p' = \frac{p^2 w_1 + pq w_2}{\overline{w}} \tag{2}$$

In the diffusion approximation, $M_{\delta x} \varphi(x,t)$ and $V_{\delta x} \varphi(x,t)$ (respectively the mean and variance in the rate at which genic frequencies change) are population specific, but the basic equation changes only marginally from population to population. Similarly, the state-variables in the population genetics equation are *p* and *q*, representing allelic frequencies, and only the fitness parameters w_1 , w_2 and *w* change over populations.²

This modeling procedure is systematically adopted in a number of domains in and beyond biology. It is nearly universal in population genetics (cf. Ewens, 2004). It underwrites most of the early demographic models (e.g., Lotka-Voltera models, see Kingsland (1985) for a historical overview),

as well as much of the current work in population regulation (for example predator-prey and SETAR models, see e.g., Stenseth et al. 1997). The procedure is ubiquitous in macroeconomics (for example, auto-regression models of various sorts, see, e.g., Enders 2003), and universal in thermodynamics (see, e.g., Ghez 2001 for discussion of some classical cases using diffusion models) and related bits of physics (e.g., fluid dynamics). And at least in the physical sciences this strategy has a long history of success, fluid dynamics not withstanding.

But there is a second strategy which is widely used in other domains. The strategy is this. If the frequency distribution of a variable v is changing in a population of units, this is because some units are changing their value of v, or units with particular values of v are differentially recruited or lost from the population. Instead of building models in which the equations of state include variables measured on the population—for example, the mean value V of v—one can build models in which the equations govern the behavior of individual components of the population.³ Specifically, if one cares to predict the mean V of v at the next time-step, one does a sequence of three things. First, one builds an equation that predicts the value of v at t + 1 for an individual unit given its values for v and causes (or anyway covariates) $c_1 \dots c_n$ of v at t. Second, one builds a similar model of each cause c_{i} , if that variable is either caused by some c_{i} , or shares a common cause with some c_i , for each c_i and c_i in the first equation. Finally, one identifies this model for the particular population-that is, estimates parameters.

The first two steps result in a dynamic structural equation model (SEM), exemplified below for a system of three causes of v, two of which (c_2 and c_3) are exogenous:

$$v(t+1,i) = I_{v} + \alpha c_{1}(t,i) + \beta c_{2}(t,i) + \gamma c_{3}(t,i) + \varepsilon_{v}$$

$$c_{1}(t,i) = I_{c} + \varphi c_{2}(t-1,i) + \phi c_{3}(t-1,i) + \varepsilon_{c}$$
(3)

Given an identified SEM and the current joint distribution in the population of all the model variables one can predict the frequency distribution over v at t + 1, and hence the value of V at t + 1.

SEMs are a particularly intuitive class of "reductive" models. Reductive models, in the sense at issue here, are distinctive in the nature of their state-variables. The coefficients in such models, as often in population models, are population specific. Unlike population-level models, however, the variables take values that represent properties of units in a population: the variables are measured on individuals rather than populations of individuals. The difference is exemplified by the distinction between measuring the temperature of a gas and measuring the momentum of each of

100,000 gas molecules. From either procedure one can estimate the mean kinetic energy in the gas, though rather more directly in the first case than in the second. The difference in state-variables reflects a difference in the kinds of causes being modeled. Unlike population-level models, reductive models represent the causes of individual behaviors, and attempt to estimate the degree to which each such cause influences the behavior of interest, either for each individual in the population, or on the average for all individuals in the population.⁴

This procedure, which we will call the reductive strategy,⁵ also has a (rather shorter) history of success. It is widely used in sociometric, econometric, and epidemiological studies (see e.g., Ayres and Donohue (2003), D'Agostino et al. (2001), Gallagher et al. (1996), Gornick et al. (1998), and Hoeffler (2002)). And for good reasons. Curiously, they are exactly the reasons that have led philosophers of biology to employ the populationlevel strategy. First, there is no reason to think that the causes of any variable of interest, say death, are the same in any two populations, say distinct countries. While smoking may be an important cause of death in both England and the United States, it undoubtedly has a much less significant effect in, say, the Sudan or Somalia. Second, even when the causes are the same, the degree to which a given cause influences a given effect may change. Obesity is a cause of death in England just as in the States, but a less important cause in the sense that being obese in England is less likely to kill you than being obese in the States. Third, and perhaps worst of all, the effect of a cause can actually change sign from population to population. Increasing the consumption of red meat in the United States will likely increase death rates in various age groups; a similar intervention in the Sudan is likely to decrease death rates in those very same groups. Social scientists build one-off, reductive models for each population of interest because the relevant causes of the behavior of interest vary in these three distinct ways from population to population.

In response to exactly the same problem, philosophers of biology have implicitly endorsed the population-level strategy while social scientists have employed, sometimes to good effect, the quite different reductive strategy. I claim there are good reasons to prefer the latter strategy in evolutionary biology, and therefore to change quite radically the demands we make of an interpretation of ET.

3 Why One-Off Models of Evolving Biological Populations Make Sense

I claim that the reductive strategy makes better sense for population biology than the more standard population-level strategy. We get better predictions, and a better understanding, of what drives evolution in any particular population if we build one-off population-specific models of those populations.⁶ SEM models exemplify this strategy.⁷ In essence, SEM models model the behavior "on the average" of arbitrary individuals in the population, and for this reason are predictively superior to models that employ a single or very few equations of state, re-identified for particular populations but always including the same set of variables measured on the population rather than individuals in the population.⁸ To make the case, I'll need to describe and represent causal relations. To do so, I adopt the interventionist conception of causation, and the graphical causalmodeling framework. I'll assume throughout that dependencies are linear. (I do so for ease of explication—the assumption is almost certainly false for biological populations, but nothing I say hinges on the assumption.) I employ population-level variables in introducing the causal ideas, but the basic assumptions hold for individual-level causes as well.

According to the interventionist conception of causation, C is a direct cause of V relative to a set S of variables if there is some pair of interventions setting the value of each variable in $S/\{V\}$, and differing only in the value to which C is set, across which the probability distribution over V varies. Less formally, there are ways to wiggle C, without directly wiggling anything else in S, which change the probability that V takes a particular value for at least one such value. A direct causal relation between C and V is represented graphically by $C \rightarrow V$. In linear cases, it is common to associate with each edge a path coefficient, representing the strength with which the cause influences the effect. Path coefficients are often estimated by standardized partial regression coefficients, though other statistics are sometimes used. So in the causal graph $C_1 \rightarrow V \leftarrow C_2$, we might assign the name α to the influence of C_1 on V, and β to the influence of C_2 on V, and represent the whole system as in figure 11.1.

If α and β are standardized partial regression coefficients, α and β can be read from (or into) the equation governing *V*: *V*(i) = I + $\alpha C_1(i) + \beta C_2(i) + \epsilon$, where I is an intercept, ϵ is an error term and i indexes units for the model, whether alleles, organisms or populations.

A special kind of causal relation occurs when *V* has two distinct causes *C* and *Z*, where the influence of *C* on *V* depends on *Z*. Light switches and



Figure 11.1



Figure 11.2

circuit breakers are instances of such causal relations: when the breaker is on—but not when the breaker is off—the light switch causally influences whether or not the light is on. In the graphical causal-modeling literature, causes like the switch and the breaker are known as "interactive" causes; elsewhere they are known as context-dependent causes. We can represent this kind of dependence using graphical causal models. To do so, we treat both *C* and α as direct causes of *V*, and treat *Z* as a direct cause of α , as in figure 11.2.

This kind of causal connection is especially relevant for us. Recollect that the problem confronting the sociologists is exactly that for individuals c may be a cause of v in one population but not another, or may remain a cause, but have a different sign or degree of influence. If this is so, then if C and V are population means for c and v respectively, C will be a cause of V in one population, but not another, or remain a cause, but exhibit a different sign or degree of influence on V. It is a reasonable presupposition that if c is a cause of v in this but not that population, there must be some other cause z of v which controls whether or not c influences v, and z must vary between populations. Similarly if c remains a cause in both populations, but differs in its sign or degree of influence. The same is true, mutatis mutandis, for C and V. But there are so many Cs and so many Zs, so many cs and zs, that it is essentially impossible to discover all of them, and even were most such causes discoverable, the available data would be inadequate to reliably identify any model containing them all.

Suppose we have *N* populations, and these populations differ in the following respect: for any two populations P_j and P_k , there is some cause C_j of *V* in P_j , and C_j is *not* a cause of *V* in P_k , where the influence of C_j on *V* is controlled by a distinct interactive cause Z_j . A full model, covering all *N* populations must therefore contain at least *N* direct causes of *V*.



Figure 11.3

Assuming each direct cause interacts with a distinct Z variable, the graphical structure will look something like figure 11.3.

If N is large, it may seem hopeless to attempt any specification of a full model (it certainly seemed so to nearly every philosopher of biology in the 1980s). We then have two choices: abstract away from such variant causes, or build population-specific models for each distinct population.

Abstraction is possible but comes at a price, namely increased error (more on this anon). One-off, population-specific models are therefore initially attractive; but this does not yet justify the reductive strategy: why not build population-specific models in which the state-variables are measured on the population? The reason is entirely epistemic. There is in general insufficient data to do any such thing. To build a model that predicts reliably, one needs many observations of the units whose behavior one is predicting. For population-level models, this unit is the population. If one measures a particular population thrice over at t, t + 1, and t + 2, one has exactly three observations of the population. From such data no equation of state can be responsibly inferred. In fact a tenfold increase in observations (e.g., a thirty-year Long Term Ecological Research (LTER) experiment) would provide sufficient data for a reliable inference only if there were very few state-variables (e.g., if there are six discrete state-variables, the data would necessarily be dangerously over- or under-dispersed for model selection over alternative population-level models). So, for example, the work of Palmer et al. (2008) on the relationship between whistle-thorn acacia and ant species rests in part on data from a decade-long LTER, and yet they of necessity use that data to produce an individual-level model of tree mortality and growth.

How is it possible that data inadequate to produce a population-level model may nonetheless suffice to reliably infer an individual-level model? The same data that constitute three observations of the population might comprise observations at three different times of many hundreds or thousands of constituent individuals, more than enough for reliable selection and identification of models of the behavior of the individual organisms. This is the hidden but overriding epistemic reason for the reductive strategy. Social scientists build population-specific models for particular populations because the causes of the variables they wish to track change from population to population—in other words, the causal structure governing the variables of interest is not uniform across populations. There are causes of these changes in causal structure (the Zvariables), but these variables are so numerous that there is no hope of discovering them all or of identifying models which contain them. Given that social scientists build population-specific models, they must build models of the behavior of individuals rather than populations of individuals, because there is insufficient data to model a population (observed only a handful of times), but often enough data to model individuals (a handful of observations on each of hundreds or thousands of individuals).

For example, if all individuals are governed by the full causal structure given in figure 11.4 but in our particular population z_2 through z_n are invariant and turned "off" while z_1 is invariant and "on," then a model of individual behavior in that particular population requires only that we identify c_1 as a cause of v, and estimate the parameter α_1 current in that population. A few hundreds of observations on units in the population







Figure 11.5

may be sufficient for that chore, and encompassed by only one or two observations of the population.

The epistemic advantage is gained at a price, however. Our model will not recover the whole structure, but only the partial structure represented in figure 11.5—namely, the fact that c_1 causes v. Because the model is population specific, one cannot export it to other populations with any confidence: in those populations z_2 may be on while z_1 is off. The most one can say is that, having identified c_1 as a cause of v for individuals in one population, it is quite possible that it will be a relevant cause in others.

The question, then, is whether biological populations are varied in the way I have suggested social science routinely presumes our particular species is varied. The response that can be elicited from any practicing field biologist is "Yes, of course!" Support for that response may be found in the plethora of time-lagged demographic models with phase-switching (the threshold variable in such models is an interactive cause, see e.g., Stenseth et al. 2004); in the increasingly common results from ecological genomics showing differential expression of genes that code for phenotypic responses that have context-dependent fitness effects (see e.g., Shimizu and Purugganan 2005); and in demographic studies showing interactions between environmental and demographic variables (e.g., Coulson et al. 2002). Still, the question is empirical and deserves a full consideration that, for reasons of space, I cannot give it here. I will make do with a single illustrative example, developed in the following section.

4 An Illustrative Case

Chaine and Lyon (2008) report work on sexual selection in the lark bunting. Male lark buntings are weakly territorial, but compete for matings by display. Chaine and Lyon measured five characteristics of male plumage (body color, proportion of black to brown feathers on the rump, proportion of black to brown feathers on the rest of the body, wing patch size, and wing patch color), and three body-size characteristics (body size, beak size, and residual mass). They also measured the mating success (number of pairings) and offspring produced by birds over a five-year period (five breeding seasons), from 1999 to 2003. Their data suggest that all the phenotypic traits are under some degree of selection. But body size and the percentage of black feathers on the body and rump were under fairly strong selection, at least in some years. What is more, Chaine and Lyon were able to show that the fitness effects of these traits are mediated by female mate choice, and further that female preferences with respect to these characteristics vary from year to year, so that in some years the characters are positively associated with fitness while in others they are negatively associated a preference for small body size, while in others (e.g., 2003) they exhibited a preference for large body size (see fig. 11.2 in Chaine and Lyon 2008).

Sexual selection in the lark bunting thus appears to be a paradigmatic case of context-dependent or interactive causation: certain trait variables cause fitness but the magnitude of this influence, and indeed its sign, depend on some interactive cause (which in this case remains unknown). The value of the unknown interactive cause is clearly not constant across time, and there is no reason to think it is constant over spatially distinct subpopulations of lark bunting at any given time. Hence, a model of sexual selection in the lark bunting must be population specific if it is to be predictively reliable over the short or medium term.

But exactly because this is so, it is essential that any such model employ variables measured on individuals. In the Chaine and Lyon study, for example, measurements of the characteristics, mating success, and fitness of several hundred birds are made over the course of the five breeding seasons. Chaine and Lyon are able from these data to infer a causal structure over individual birds. But the data do not permit one to directly specify, never mind identify, a population-level model. Were a population-level model desired, one could attempt to build one by inferring from the data measured on individual birds to estimated values for population-level variables (mean male body size, mean fitness, mean fitness of large-bodied birds, etc.). But any such attempt would face insuperable difficulties. Suppose one so estimates a suite of n population-level variables, $V_1 \ldots V_n$. Since the measurements are taken from only five breeding seasons, each such variable will be measured at five different times—for instance, $V_1(1999)$, $V_1(2000)$. . . $V_1(2003)$ —and similarly for the other n – 1 population-level variables. Hence, when specifying a population-level model, one's data set
will have a sample size of *five*. Model selection with such limited data is hopelessly unreliable. Conversely, the sample size for inferences to yearand population-specific individual-level models is an order of magnitude higher. The difference in sample size is sufficiently large that while one cannot reliably specify a population-level model, one can infer to an individual-level model—and this is exactly what Chaine and Lyon do.

Chaine and Lyon are not unique in adopting such modeling methods in biology. Although the reductive strategy is not typically announced as such, it is routine in population biology. A second recent example is Palmer et al. (2008) on the relation between ants and acacia, but one can find individual-level SEM models in use in population and evolutionary ecology as early as the 1960s (see Tilley 1968 for a lovely example employing population-specific regression models for species).⁹ In fairness, it is important to note the contrary cases. Some populations (bacteria, viruses, etc.) can be replicated in the laboratory many, many times, and it is sometimes the case that for such organisms it is often much easier to measure variables on local populations than on individual organisms. In such laboratory work, it is not necessary to adopt individual-level models, and quite often models are correctly built at the population level (see e.g., Lenski and Travisano 1994). But in general, wild populations are not subject to such replicated measurements at the population level.

5 Mistakes Compounded

If the foregoing is correct, then it is a serious mistake to think about ET as an abstract theory which can be formalized by some small set of equations or laws which can then in turn be usefully applied to any given population. ET is instead a recipe for building models that predict the trajectories of populations through hyperspaces defined by phenotypic and genotypic frequencies. This recipe says: build models that pay attention to survival, reproductive success, and heritability. ET does not itself require that survival and reproductive success be measured on populations (as a fitness parameter) or on individual organisms: the theory itself permits measurements, and hence models, of either sort. But as it turns out, there is a further, entirely epistemic restriction on competent models: since the causes of survival and reproductive success are so enormously varied, it is often the case that good models of them are population-specific, individuallevel models that track the causes of individual survival and reproductive success rather than the causes of population rates of survival and reproductive success. Failure to recognize this when interpreting ET, and hence failure to recognize the epistemic restriction, have further consequences. There are at least two particularly disagreeable consequences worth noting here: the first concerns prediction, the second explanation.

It is a well-known fact that population genetics models are next to useless in predicting the behavior of actual populations in the short and medium term. The reasons for this are straightforward. Population genetics models generally include only a very small range of state-variables, omitting representation of most of the causes of survival and reproductive success in any particular population. Those causes are often not fixed in value but instead vary over time, and often influence survival and reproductive success interactively. Consequently, the error term in the predictive equations is necessarily large. When putative causes are used to precisify standard models, the data used to do so are nearly always seriously over- or under-dispersed. Hence the model uncertainty and confidence intervals on parameters are large. Predictions will typically be imprecise, and when precise they can not be made with any confidence. It is not accidental that most of the really interesting correct *predictions* from classical population genetics concern long-term equilibria for ideal populations rather than predictions about the future behavior of cod or blue whales or HIV.

A somewhat different worry besets the use of very general abstract population-level models in explanatory contexts. Population-level models attend to the determinants of population behavior, and because the available data will include only a small number of observations of any given population, a population-level model of any particular population can reliably identify only those determinants of its behavior that are influential in all populations. There are two categories of such determinants: those that are nomically connected to genic or genotypic frequencies by a strictcovering law, and those that are connected by mathematical or conceptual necessity. To all appearances there are very few of the former; what remains are a collection of mathematical dependencies. It is not accidental that the important *explanatory generalizations* from population genetics are theorems (e.g., Fisher's fundamental theorem, the Price Equation), or are conceptual truths (e.g., the Breeder's Equation), or are beset by systematic exceptions (e.g., Fisher's explanation of sex-ratios).

There is of course nothing wrong with such mathematical truths or their application to real populations in the search for understanding. There are also domains in which a sufficiency of data is available for developing population-level models. In particular, there are various species in which population size has been measured or can be estimated for many years: when such data are available they can often be used to determine, for example, whether or not there are phase-transitions in the dynamics of population growth.¹⁰ Similarly, laboratory experiments with many replicate populations of bacteria or viruses, for instance, often allow or even require modeling at the population-level. In these latter cases, it is sometimes possible to identify empirical dependencies among population-level variables. But for very many wild populations, the requisite data are simply unavailable. In such contexts, the discoverable dependencies among population-level variables will nearly always be mathematical or conceptual in nature.

To the extent that such dependencies exhaust those we employ in explaining evolutionary events, we deprive ourselves of certain kinds of relevant and interesting explanations. In particular we will be able to explain different evolutionary trajectories for a pair of populations only by appeal to different variable values or, in the best case, different values for some parameter. We will *not* be able to explain different evolutionary outcomes for distinct populations by appeal to different causal structures governing survival and reproductive success in those populations. This for the simple reason that reliable *population-level* models of the two populations will not differ in the variables they identify as relevant to the outcomes in question, or in the formal expression of the dependencies between them.

This constraint on explanation is deeply at odds with longstanding practice in evolutionary ecology. For example, life-history strategies vary among species (and within a species), and different tradeoffs are optimal for different life-history strategies. What is more, specific tradeoffs among life-history traits entrain further selective pressures: that is, given that a species is characterized by a particular tradeoff (e.g., between altricial and precocial reproductive strategies), some but not other variables come to influence reproductive success (e.g., offspring growth rates), and these causes constitute a selection pressure. Characterizing these tradeoffs, their causes, and their consequences is of central explanatory importance in evolutionary biology, because the causes and consequences of such tradeoffs explain differences in evolutionary outcomes between species and indeed whole lineages.

Specific examples are not hard to find. Reznick et al. (1996) show that predation influences a tradeoff between adult mortality and reproductive effort in guppies—in other words, predation is an interactive cause of the degree to which reproductive effort influences adult mortality. When predation is present, the probability of adult mortality is high whether or not an individual over-invests in reproduction; when predation is low,

over-investment in reproductive effort severely compromises the probability of survival, and hence the opportunity for future reproductive success. Similarly, Kolm et al. (2006) show that body size influences the tradeoff between clutch size and egg size in Cichlids, and Kinnison et al. (2001) show that migration, and the distance migrated, influence the same tradeoff in salmon.

At the other end of the causal structure: Bosque and Bosque (1995) show that a particular tradeoff entrains a specific selection process, namely, that nest predation influences growth rate among altricial birds. That is, given that a bird species is altricial, the probability of nest predation causally influences the parameter describing the degree to which the growth rate of an immature individual influences its fitness. Another such case is described by Avilés et al. (2008): among altricial birds with clutch sizes greater than one, nestlings compete for food and this imposes signal selection on those traits that effectively solicit feeding behavior from adults. Avilés et al. (2008) show that such signal selection occurs, and in particular show that traits characterizing the achromatic contrast between nestling and nest are causally influenced by clutch size and nest location (a correlate of nest predation).

While many specific tradeoffs, their causes, and their selective consequences are predicted by theory, confirmation of these predictions nearly always involves modeling-typically modeling of individuals within or across distinct populations. And often the observational data lead to unexpected consequences, producing changes in theory. For example, Martin (1995) showed that, contrary to theoretical predictions, food limitation is much less important in generating life-history tradeoffs in birds than is nest predation. This study is especially interesting because it shows how a sequence of one-off population-specific models can be used to generate general models or theory. Martin's study employs data from studies on 123 specific species. The data for each species allows a diagnosis of the causal structure governing the fitness of individuals in that species. By looking at the results for many species, general patterns can be diagnosed. These patterns are not, of course, exceptionless, but this is just what one expects when distinct populations are characterized by distinct values for interactive causes. Rather than strict-covering laws, one hopes at most for *ceteris-paribus* regularities and, perhaps, a diagnosis of the interactive causes, variation in which produces the exceptions.

The explanatory problem for any understanding of ET that rests exclusively, or even predominantly, on population-level models can thus be described in the following way. The reason, for example, that elephants and humans care for their young while cassowaries do not is that *different causes affect survival and reproductive success* for these species. What is more, *because* elephants and humans care for their young while cassowaries do not, there are yet further differences in the causes of survival and reproductive success among these species. Those causes and the life-history strategies that determine their relevance are lost to any general population-level formulation or formalization of ET. On standard accounts, whether "dynamic" or "statistical," work investigating such causes must be shoehorned in as something extra, something else, something different from ET—applied ecology rather than evolutionary theory. But of course, such work is not something else or something extra—it is the heart and soul and nearly all of the *biologically* interesting content of evolutionary theory. If one wants to be able to explain contrasting behavior by citing the relevance of different causes, population-level models are bad news.

6 Final Considerations

The difference between the population-level and reductive strategies in evolutionary biology is, in its simplest form, the difference between measuring selection by selection differentials and measuring selection by selection gradients (Lande and Arnold 1983). The former coefficient is simply the difference in relative fitness between the most fit and less fit classes. If selection is so measured, it becomes crucially important to understand what fitness is and what role it plays in evolutionary explanations. If ET is regarded as a population-level theory, such explanations are applications of the equations of population genetics, and those equations are, more or less, invariant across populations. Different behaviors can be explained only by appeal to different variable values (initial frequencies for the classes) or by appeal to different parameter values, fitness being the most important among them.

The latter coefficient, a selection gradient, is an estimate of the causal influence of a trait (phenotypic or genotypic) on *individual* fitness, where fitness just is whatever function of survival and reproductive success happens to be of interest in the particular population. If selection is so measured then univocal, generally applicable interpretations of fitness are irrelevant and nearly completely without interest. The relevant explanatory equations are population-specific, individual-level causal models of survival and reproductive success. Differences in population behavior can be explained by appeal to different variable values and different parameter values, but also by appeal to differences in the causal structures governing survival and reproductive success in the populations of interest. These explanations do not depend on the availability of a generally applicable concept of fitness.

This simple difference in how selection is measured has enormous import for how we understand the content of ET and the core formal representations of the role of natural selection in generating evolutionary events. If selection is measured by selection differentials then the core formal representations must be found in population-genetic models. If selection is measured by selection gradients, however, then the core formal representations of selection are no longer to be found in population genetics but in population-specific causal models of survival and reproductive success produced for distinct populations, and in the procedures for reliably producing such models. The latter understanding, I have claimed, is superior in two respects: it permits superior predictive models, and it permits explanations of divergent evolutionary outcomes by appealing to differences in the causal structures which govern survival and reproductive success. It further possesses a versatility: to endorse individual-level models as the core representational scheme for models of selection-driven evolution is not thereby to deny the legitimacy or usefulness of population-level models when data sufficient for their reliable specification is available. The converse has not, at least in practice, been true.

The superiority of the reductive procedure is grounded epistemically. Survival and reproductive success have many interactive causes; those causes are often fixed in value within a population but vary in value between populations; hence, survival and reproductive success have different causes in different populations. Building models that include *all* such causes is epistemically impossible since the available data would not suffice to reliably specify the relevant model, or to identify the model even if it could be otherwise reliably specified. But building population-specific models of the behavior of individuals *is* possible.

These epistemic considerations are themselves not invariant: they do not generally hold of atoms or molecules, and this is why much of chemistry and physics is not tied to the reductive strategy. Statistical mechanics is possible because there are relatively fewer interactive causes of particle behavior. This suggests a test, or anyway a heuristic, for determining where reductive strategies are to be preferred and where nonreductive strategies are to be preferred. Just how many interactive causes do we think there are? And to what extent do these causes vary in value across the "higher"level units, while remaining invariant for all or most "lower"-level units which constitute a given higher-level unit? Where the answer is "many," reductive strategies are promising. Where the answer is "few," populationlevel procedures are recommended. In biology, the answer is "many."

Notes

1. In biology, the state-variables are generally *estimated* rather than measured: one measures a variable on each member of a sample of the population, and from this calculates a sample value for the related state-variable, e.g., the mean of the measured variable. One then infers from this sample value to an estimated population value for the state-variable, via one or another statistical method. Such indirect measurements of state-variables are sometimes not required: the measurement of a temperature is a direct measurement of a population-level state-variable, namely the mean kinetic energy.

2. There are of course differences in the models used for distinct populations: haploid versus diploid, associative mating versus panmictic, etc. But generally these differences are encoded either in different parameter values rather than different variables, or, at most, in extra variables induced by a finer partition of the population into classes.

3. I use lowercase italicized English letters to represent variables measured on individual organisms, uppercase italicized English letters to represent variables measured on populations of organisms, and lowercase Greek letters to represent parameter values.

4. In so-called random-effects models it is assumed that each unit in the population may be characterized by distinct coefficient values, e.g., the effect of c_1 on v(t + 1) for one unit may differ from the same effect for another unit; α is therefore estimated as the *mean* of this effect in the population. In so-called fixed-effects models, it is assumed that the magnitude of these effects is invariant over individuals in the population. In either case, α is not assumed to be invariant over different populations.

5. The definite article is used for convenience. There are of course other reductive strategies; they will not concern us here.

6. Pigliucci and Kaplan (2006) independently defend reductive models in evolutionary biology. Their arguments turn on the explanatory competencies of individuallevel models, though the competencies of concern to them differ from those to which I draw attention in this paper.

7. SEM models are not unique in this respect; they are however both intuitive and especially applicable in biology, the science from which the examples in this paper are largely drawn, and so I employ them here preferentially.

8. So-called agent-based models are the limiting case of reduction to the individual level. In these models, each individual is characterized by a unique equation with

unique parameter values, and both the individual specific equations and parameters may evolve over time. But because *all* parameters in such models may vary across individuals, or even for particular individuals over time, such models cannot be specified or identified from data measured on real populations. Fully reduced agentbased models therefore fail in respect of our purposes here, though they serve other aims well.

9. The results reported in Palmer et al. 2008 are especially interesting: they show that an ant-acacia mutualism is in fact context dependent, and further identify the relevant interactive cause, namely the presence of large herbivores.

10. Even in these cases, however, any attempt to discover the relevant interactive causes will almost certainly require individual-level models, because the potential environmental causes have typically not been measured, and cannot be estimated, at the population level.

References

Avilés, J., T. Perez-Contreras, C. Navarro, and J. Soler. 2008. Dark nests and conspicuousness in color patterns of nestlings of altricial birds. *American Naturalist* 171:327–338.

Ayres, I., and J. Donohue, III. 2003. Shooting down the "more guns less crime" hypothesis. *Stanford Law Review* 55:1193–1312.

Brandon, R. 1978. Adaptation and evolutionary theory. *Studies in History and Philosophy of Science* 9:181–206.

Bosque, C., and M. Bosque. 1995. Nest predation as a selective factor in the evolution of developmental rates in altricial birds. *American Naturalist* 145:234–260.

Bouchard, F., and A. Rosenberg. 2004. Fitness, probability, and the principles of natural selection. *British Journal for the Philosophy of Science* 55:693–712.

Chaine, A., and B. Lyon. 2008. Adaptive plasticity in female mate choice dampens sexual selection on male ornaments in the lark bunting. *Science* 319:459–462.

Coulson, T., E. Catchpole, S. Albon, B. Morgan, J. Peberton, T. Clutton-Brock, M. Crawley, and B. Grenfell. 2002. Age, sex, density, winter weather, and population crashes in soay sheep. *Science* 292:1528–1531.

D'Agostino, R., S. Grundy, L. Sullivan, and P. Wilson. 2001. Validation of the Framingham coronary heart disease prediction scores. *Journal of the American Medical Association* 286:183–187.

Enders, W. 2003. Applied Econometric Time Series, 2nd ed. Hoboken: John Wiley.

Ewens, W. 2004. Mathematical Population Genetics. New York: Springer.

Gallagher, D., M. Visser, D. Sepulveda, R. Pierson, T. Harris, and S. Heymsfield. 1996. How useful is body mass index for comparison of body fatness across age, sex and ethnic groups? *American Journal of Epidemiology* 143:228–239.

Ghez, R. 2001. Diffusion Phenomena: Cases and Studies. New York: Kluwer.

Gornick, J., M. Meyers, and K. Ross. 1998. Public policies and the employment of mothers: A cross-national study. *Social Science Quarterly* 79:35–54.

Hoeffler, A. 2002. The augmented Solow model and the African growth debate. *Oxford Bulletin of Economics and Statistics* 64:135–158.

Kingsland, S. E. 1985. *Modeling Nature: Episodes in the History of Population Ecology*. Chicago: University of Chicago Press.

Kinnison, M. T., M. J. Unwin, A. P. Hendry, and T. P. Quinn. 2001. Migratory costs and the evolution of egg size and number in introduced and indigenous salmon populations. *Evolution: International Journal of Organic Evolution* 55:1656–1667.

Kolm, N., N. B. Goodwin, S. Balshine, and J. D. Reynolds. 2006. Life history evolution in cichlids 2: Directional evolution of the trade-off between egg number and egg size. *Journal of Evolutionary Biology* 19:76–84.

Lande, R., and S. Arnold. 1983. The measurement of selection on correlated characters. *Evolution; International Journal of Organic Evolution* 37:1210–1226.

Lenski, R., and M. Travisano. 1994. Dynamics of adaptation and diversification: A 10,000-generation experiment with bacterial populations. *Proceedings of the National Academy of Sciences of the United States of America* 91:6806–6814.

Martin, T. 1995. Avian life history evolution in relation to nest sites, nest predation, and food. *Ecological Monographs* 65:101–127.

Mills, S., and J. Beatty. 1979. The propensity interpretation of fitness. *Philosophy of Science* 46:263–286.

Palmer, T., M. Stanton, T. Young, J. Goheen, R. Pringle, and R. Karban. 2008. Breakdown of an ant-plant mutualism follows the loss of large herbivores from an African savanna. *Science* 319:192–195.

Pigliucci, M., and J. Kaplan. 2006. *Making Sense of Evolution*. Chicago: University of Chicago Press.

Reznick, D., M. J. Butler, F. H. Rodd, and P. Ross. 1996. Life-history evolution in guppies (Poecilia reticulata) 6. Differential mortality as a mechanism for natural selection. *Evolution: International Journal of Organic Evolution* 50:1651–1660.

Rosenberg, A. 1983. Fitness. Journal of Philosophy 80:457-473.

Sober, E. 1984. The Nature of Selection. Cambridge, Mass.: MIT Press.

Sterelny, K., and P. Kitcher. 1988. The return of the gene. *Journal of Philosophy* 85:339–361.

Shimizu, K., and M. Purugganan. 2005. Evolutionary and ecological genomics of arabidopsis. *Plant Physiology* 138:578–584.

Stenseth, N. C., W. Falck, O. Bjornstad, and C. Krebs. 1997. Population regulation in snowshoe hare and Canadian lynx: Asymmetric food web configurations between hare and lynx. *Proceedings of the National Academy of Sciences of the United States of America* 94:5147–5152.

Stenseth, N. C., D. Ehrich, E. Rueness, O. Lidgjaerde, K. Chan, S. Boutin, M. O'Donoghue, D. Robinson, H. Viljugrein, and K. Jakobsen. 2004. The effect of climate forcing on population synchrony and genetic structuring of the canadian lynx. *Proceedings of the National Academy of Sciences of the United States of America* 101:6056–6061.

Stephens, C. 2004. Selection, drift, and the "forces" of evolution. *Philosophy of Science* 71:550–570.

Tilley, S. 1968. Size-fecundity relationships and their evolutionary implications is five desmognathine salamanders. *Evolution: International Journal of Organic Evolution* 22:806–816.

Walsh, D., T. Lewens, and A. Ariew. 2002. The trials of life. *Philosophy of Science* 69:452–473.

12 Similarity and Species Concepts

Jason G. Rheins

1 Introduction

Conceptions of species that are explicitly based on similarity or resemblance between group members have fallen out of widespread favor despite the apparent importance of morphological criteria of identification for field biologists. To put the matter succinctly: similarity, be it morphological or genotypic, is widely regarded as insufficient to demarcate *natural kinds* or monophyletic branches on the tree of life. Rival views such as the ecological, biological, and evolutionary conceptions of species are regarded as more likely to be successful grounds for delineating real kinds, and many defenders of the now dominant evolutionary-cladistic conception have gone so far as to argue that species are not kinds at all, but individuals.

I will speak of what are normally called species "concepts"—for instance, the "biological species concept"—as "conceptions" since they are conceptions of what species are. The term 'species concept' will be reserved for individual concepts of organisms as species (e.g., "Canis familiaris" and "Homo sapiens"). Hence by 'kind-conceptions' of species I mean a theory of species that understands them to be kinds or concepts. The alternatives I will call "individual-conceptions" of species. By this usage I am not suggesting that an individual can ever be a concept (in my view concepts sort and unite many particulars), but rather that this conception or view of species regards them as individuals. If any kind-conception is to be adopted it must be one based on similarity. The merits of an individual-conception and the extent to which cladism is really individualist will be considered after arguing for the prior point.

I will argue that the rejection of the view that species are groups of organisms united conceptually by similarity relationships is mistaken. It is wrong not merely because the similarity conception of species is superior to the biological, ecological or evolutionary conception of species, but because any kind-based conception of species, including those which make reference to reproductive isolation, niche adaptation or causal-historical phylogenetic relationships, must ultimately be based on similarity relationships of some sort. Insofar as individual-conceptions of species must still make use of classification, they too will be dependent on similarity-based kinds in the identification or demarcation of their particular populations or branches on the tree of life.

This fact has been obscured for at least two reasons. One is that for the most part similarity has been too narrowly construed in terms of only morphological or genotypic properties. This is not without some reasonable basis, as comparative morphology or anatomy has traditionally played a vital role in taxonomy, and though the biological and now evolutionary conceptions have superseded it as the conventional bases of species demarcation, similarity of structure still plays an important role in the higher levels of systematics.

Another and to my mind more fundamental reason is the realist tendency to understand species as natural kinds whose members stand in some sameness relationship through mutual possession of an identical property or causal integration into a unit. Yet, I will claim that a notion of similarity (though not exclusively morphological or genotypic) is implicit even in concepts formed according to the biological and ecological species conceptions. Therefore, if being based on similarity disqualifies species from being objective, then *ecological*, *biological*, and by extension, *evolutionary* species concepts (and hence their respective conceptions) will also be ruled out as subjective.

Kinds can be objective though their members are related to one other by similarity, not sameness. The biological, ecological, and evolutionary conceptions are forms of realism that seek for a qualitative sameness to hold between species members. However, sameness cannot be found among the members of species, or, if it can, then with respect to classification it would be a superfluous matter that would have little bearing on the actual and even ideal formation of biological concepts. Some of the stigma attached to similarity can be wiped away once we see that sameness is a misdirected criterion for identifying species (and kinds more generally).

Moreover, the range of properties in which the members of a kind may be similar to one another in a manner sufficient to justify categorizing them together is more diverse and rich than solely the commonality of appearance or even microstructure. Similarity with regard to descent, reproductive community, adaptation to a niche or any other influential and informative property can also count as criteria. Even if one of these criteria is ultimately selected over the others as most fundamental,¹ the other criteria will remain necessary and influential, if not essential in the strictest sense.² Once we do understand what it means to base one's species concepts on similarities it will be argued that it is plausible to think that one can unite some of the best features of the major species concepts, and that one will be able to avoid at least some of the negative consequences that pluralism about species might imply.

I will not contest at great length the arguments for the Individuality thesis, namely, that species are individuals.³ I find fault with the "entityenduring-across-evolution" arguments for that thesis, and in my last section I sketch out a criticism that this argument is *non sequitur* by deploying a distinction between kinds and systems. My paper will be more concerned with arguing for the presupposition of similarity in all theories that treat species as kinds, and the advantages of a theory based explicitly on similarity. However, I will say why individual-conceptions of species depend on kind-conceptions and are thus subject to the force of my claims about kinds.

2 Similarity or Sameness as the Basis of Concepts and Kinds

Similarity Is Relatively Little Difference in Some Respect

Regrettably, I can only present my views about the nature of similarity and its role in concepts/concept-formation very briefly here, and with little argument besides a few illustrative examples. By 'similarity' I mean the resemblance relationship that holds between two or more things (the "similars") when their differences in some specific respect(s) are dwarfed by their differences to one or more dissimilar things (the "foil").⁴

Unlike sameness, which can be assessed when directly comparing two things, judgments of similarity require a foil. If asked without further prompting, "Are Bill Clinton and Ronald Reagan similar in their political ideologies?" then it is hard to give an unambiguous answer, even if most would say not.⁵ However, if the question ran "Are Bill Clinton and Ronald Reagan similar in political ideology when compared to Vladimir Lenin?" then the answer is certainly yes. With the qualifier "when compared to Margaret Thatcher?" then the answer is a definite no.

Unlike identity and otherness, similarity and dissimilarity must specify a "respect-in-which." Reagan and Thatcher, when compared to Clinton, are similar with respect to their political ideologies, but of the same triad Reagan and Thatcher are *dissimilar* with respect to nationality or sex, while Reagan and Clinton are similar.

Similarity consists in *relatively* little difference, where the difference between the compared units is small when compared to their far greater differences from some other thing(s), the foil, in the same dimension(s) of variation. A very simple example would be grasping a similarity in reaching the concept "red" from three balls of the same size, two in shades of red and one in a shade of blue. We can say that the red ones are similar (similar enough to be the basis of the concept "red") relative to the blue because the differences between them are swamped out by the qualitative difference between each of them and the blue. A concept will be a grasp of a range of possible "measurements" along the relevant axes of variation, as contrasted against all other values that lie outside that range. Something will be red if it has some but any hue falling in the range from which foils like blue, white, and yellow have been excluded.⁶ Forming a concept is identifying and integrating a range of possible measurements that its units can fall into, by seeing those measurements as similar against a foil.7

Realism about Kinds and Sameness

An antirealist about kinds will argue that while the individual members of such a kind are real, the kind itself is not. Some nominalists go further and claim that the only reason that all members of the kind are members is because they have been designated such by the term which is the label for that class; for example, Fido and Lassie are both dogs simply because both are called "dog," or because they resemble a particular things such as Spot that one already calls "dog." I will reserve the term 'nominalism' for this narrower thesis, with the wider range of views simply being called "antirealism."⁸

The realist response, of course, is to say that classes such as dogs and cats are real. The most radical realist position, Platonism, claims that it is the kind-concept as a Form, and not members of the kind, which has ontological priority. The Form F exists independently of any f-particulars, which in turn exist derivatively as reflections of their archetype, F. More moderate views claim that these kinds are real or natural because there is some real sameness possessed by all the members of the kind. The position need only claim that such essences exist, but it is natural to further claim that this essence can to some significant degree explain the other common properties of the kind members, thus further linking the sameness relationship to the objectivity of the kind.

The essence of a kind may not be immediately obvious to those who use the term. I may identify a kind by means of a nominal essence or stereotype that is a conjunctive set of properties that sort members from non-members.⁹ Yet the realist claims that what makes all instances of gold really gold is not that they satisfy the formula "soft, shiny, yellow metal," but rather that the atoms of such samples share the common essential property of having atomic number 79. If, say, pyrite has all other outward properties in common with real gold, but its atomic or molecular microstructure is different, then realists (at least the modern Putnamian sort) deny that it is real gold (Putnam 1973, 700–704).

"Sameness" is typically understood in one of two ways by realists. The stronger way claims that for each member of the kind to share the same property is for each of them to possess some share of one and the same property, a thing that is numerically one. This is the sameness of "my pinky and thumb belong to the same hand," and it is a step back toward Platonism. Although it conforms to the Aristotelian demand that all forms be immanent, it makes particular members belong to a kind by participation in this single, spatiotemporally extended property. In fact, Plato imagined and rejected just such a position as one interpretation of the participation relationship between Forms and the many in his *Parmenides*.¹⁰ The fault that Plato's Parmenides finds with the view is that it fails to account for the many being one by reference to one single form but rather refers to the different portions of what should be indivisible. Since Plato's Forms are changeless and outside of space and time, it is unsuitable for this to be their relationship to the many in the sensible world. It would risk giving the total state of all f-individuals too much of a logically determinative relationship over our characterization of the F-itself. If the many changed, then the Form would change too. Thus, the view is not a very moderate realism at all, but an "immanent Platonism"-that is, extreme realism with the Forms brought into and scattered across space and time. I will refer to this view, variously, as "moderate Platonism" or "immoderate realism."

The traditional form of moderate realism holds that individuals A_1 , $A_2 \ldots A_n$ can stand in a sameness relationship S to one another if they each possess some property P_1 , $P_2 \ldots P_n$ where these respective properties are qualitatively identical, but numerically distinct. This is the sense in which two isomorphic structures are formally the *same*, though perhaps physically or numerically distinct from one another. Thus essential properties are "repeatables." This view, while still realist, takes a significant step away from Platonism. I will call this view "moderate realism proper."

Thus to summarize my distinctions among the types of kind-realisms:

Platonism/extreme realism If many particulars can be jointly recognized as φ , then there is a Being (or thought in God's mind), *F*, that exists independently of any particular φ . A particular is φ or φ -like if and only if it stands in the participation relationship with *F*.¹¹

Moderate realism There is a common property or universal, *U*, that is immanent in all (and only) the particulars that, by possessing it, are φ . *Moderate Platonism/immoderate realism* The Universal or Essence, *U*, that is in all φ -particulars, is numerically one and the same.

Moderate realism proper Each φ -particular has its own numerically distinct Universal, *U1*, *U2*, *U3*... *Un*, but each of these is qualitatively identical with the relevant universal property in every other φ -particular.

3 Realist Species Conceptions

The Biological Conception as Realist

In what ways do alternatives to the similarity conception of species presuppose realism? Let us begin with the biological species conception. Proponents of this view claim that a species is a reproductively isolated group or a reproductive community (Sterelny and Griffiths 1999, 187-189). As Ernst Mayr put it, species are "groups of interbreeding natural populations that are reproductively isolated from other such groups" (Mayr 1969, 26). What exactly counts as a reproductive community is of course somewhat disputed, but the main idea is that if two organisms could in principle be capable of producing viable, nonsterile offspring, then they belong to the same species. That is to say that if I am stranded on a deserted island, out of contact with human civilization and without hope of rescue, and am therefore incapable of finding a mate and passing on my genes, then I have not ceased to be a member of Homo sapiens. Thus, we might refine our definition in the following way: A and B are both members of the same species S if A and B are so constituted that they could reproduce viable, fertile offspring given the opportunity. Yet what if two members of the species, A and C, are both males? That is not a difficult problem. If A and B are members of the same species (where B is a female) and B and C are members of the same species, then by transitivity A and C are members of the same species.¹²

Another problem that arises is what is meant by 'so constituted' in the above definition? Is this a morphological/phenotypic or a genotypic specification? If the former, then are we forced to claim that Chihuahuas and Great Danes are distinct species? A not unrelated problem is raised by the phenomenon of ring species. In this case subgroup A can mate with B, B with A and C, C with B and D, but not A with C nor with D, nor B with D, and so forth. Both problems can presumably be addressed by understanding reproductive community as a gene pool in which genetic information is or can be more or less gradually shared. That is, a reproductive community is that group of organisms with "gene flow." The Chihuahua may not be able to reproduce with the Great Dane, but their genes will eventually be able to commingle through other canines, and subgroup A's genes can reach C's through B.¹³

The significance of this is that without gene flow, it is alleged, evolutionary divergence would occur, but with it subpopulations of the species are kept similar to one another. Thus: "the species that mates together, stays together." This is taken to be a strength of the biological conception, for it is claimed that it recapitulates and reflects much of the findings of modern evolutionary theory. Across an evolutionary timeframe gene flow and reproductive isolation are taken to cause and explain stability and divergence in a population. Various scientists and philosophers have called just this claim into doubt, adducing gaps between phenotypic similarity or niche adaptation on the one hand and gene flow or reproductive isolation on the other (Mishler and Donoghue 1982, 494–496). But for the moment I want to set aside this and other objections to the *biological* species concept and simply show in what way it is (as I have described it) an immoderate realist thesis.

Another way of presenting the immoderate or moderate realist position on natural kinds, and one that will be helpful for determining what kind of realism the biological conception falls into, is to claim that kinds are groups with clusters of like properties unified by *homeostatic mechanisms* (Boyd 1991, 141-142). These mechanisms are taken to account for why property clusters are maintained. In the case of gold, for instance, the atomic number is what accounts for the commonality of properties among all samples of gold. It makes gold samples what they are, thus homeostatic mechanisms are a modern, causal approach to essences, and it is precisely the role of a homeostatic mechanism that "gene flow" is fulfilling in the biological conception. However, unlike the atomic number in each sample of gold, which is a repeated but qualitatively same property, in the case of species there is one property-that of membership in a group with gene flow—that demarcates a species on the biological conception. Thus it is an immoderate realist, not a true moderate realist, view; it is by partaking of or belonging to the one reproductive group that members constitute a kind.

The Ecological Conception as Realist

Now let us turn to the ecological conception of species. Briefly, species members are here united by occupying the same ecological niche, hence they are in competition with each other for the same resources and opportunities for reproduction (Sterelny and Griffiths 1999, 192–193). Where reproductive isolation and gene flow served as the homeostatic mechanisms for species under the biological species conception, being adapted for a niche, and thus being ecologically isolated, "holds together" species under the ecological conception, making them subject to the same environmental and evolutionary forces. Thus, being adapted for a niche is the same property all members of a species share. This conception can be developed into either immoderate or moderate realism. It will be moderate realism if the property in question is adaptive sameness; it will be immoderate realism if each individual is a member of the same species if it is competing for the same place in the same ecosystem (and consequently is so adapted).

The Evolutionary Conception as Immoderate Realism

Finally, the evolutionary species conception organizes organisms into species in terms of lineages. A species is a monophyletic branch of the tree of life. That is, it is all the organisms descended from a population formed in a speciation event until the next speciation event. If one group split into subgroups A and B, then species A is its subgroup's members at the time of speciation and all their descendents until it splits again. If only a small portion breaks off, C, then the main branch might continue to be called A. But if A splits into two or more large groups (D, E . . .), then species A no longer exists.

It has been argued, soundly I believe, that on its own the evolutionary conception of species is insufficient (Kitcher 1984, 323–324). Although the evolutionary conception has a way of determining who the members of a species are given the beginning and end of a branch, on its own it is not sufficient to determine what counts as a speciation event and which individual organisms should be counted as members of which new species at the time of speciation. If what determines that a speciation event has occurred is that two lines of descent have emerged from one lineage which are reproductively isolated from each other or which are adapted for a different niche, then the evolutionary conception is parasitic on the biological or ecological conceptions respectively. Indeed, all that the evolutionary conception would add is that the members of the species, in *addition* to being determined by gene flow or adaptation for a niche, must *also* be descendants.

The addition of evolving lineages to biologically or ecologically carvedout groups might merely add problems to the basic conceptions discussed above. We already saw how the biological conception faced certain problems regarding the transitivity of reproductive relationships. The gaps that might exist were, in that case, all synchronic. That is, at a given time B might be able to reproduce viably with A and C, but A could not with C. The evolutionary version of the conception threatens to add diachronic intransitivity as well.

Assume, for instance, a monophyletic branch of the tree of life that constitutes a species for some period of time sufficient to be significant for evolution. Let us assume that development and adaptation of a species S has occurred during this time though the lineage has been continuous and has not sprouted founding populations of new species. The members of the species at some time t_0 not long after the initial formation of S by speciation might be capable of producing viable offspring with members of S at some later time t_1 and those at t_1 with those at an even later time t_2 , but let us assume that no member of S at t_0 could produce viable offspring with a partner from t_2 . Thus what fixed the membership of the species initially may not coincide with the other principle of demarcation for the species, namely lineage. There might be similar gaps between niche adaptation and lineage if the evolutionary conception uses an *ecological* foundation.

The defender of such an evolutionary conception might try to defend his position by saying that the definition of a species is not the *conjunction* of reproductive isolation (or niche adaptation) and lineage in a monophyletic branch, but rather a *disjunction* of these things. At the point at which members of the same species will not count as different on the first criterion (*biological* or *ecological*) one will resort to the *evolutionary* criterion. However, that raises the question of whether a speciation event has, in fact, occurred. The defender of the evolutionary conception ought to say that in the example I am considering the gradual development between t_0 and t_2 does not constitute speciation: speciation is more or less synchronic, while the differences I have used as objection are diachronic. So he will refine his notion of speciation to include only cases where a species at a given time (or a sufficiently short range of time) breaks into two or more groups defined *biologically, ecologically* or however else.¹⁴

For the purposes of my argument about the role of similarity in species concepts, I can grant the defender of the evolutionary conception this much—for such species conceptions, even if they use different criteria at different times to pick out the members of species, will yield immoderate realist kinds. They will be at least immoderate realism (if not an outright individual-conception) even if their foundation is a moderate realist type of ecological conception because what additionally being a monophyletic group is meant to do is to hold the species members together in virtue of their phylogenetic relations. This makes the lineage a single, unbroken group and the speciation events that generate (or corrupt) a species are meant to be those occurrences that split up these relations. But linkage in a shared web of causal relationships is not mutual possession of a repeatable trait.¹⁵ So if this is kind-realism at all, then it is more plausibly understood as immoderate realism where there is one, common property which is shared by all group members and is not a repeatable.

4 Against the Standard of Sameness in Realist Species Conceptions

General Points about Sameness

There are real examples of sameness. For instance, there is quantitative equality: this week has the same number of days as last week. However, attempts to base concepts on sameness, either as a real essence or a nominal one, are notoriously difficult. Take the famous scholastic example of "man." Both Locke, a conceptualist defender of nominal essences, and Aristotelians, defenders of real essences, define 'man' as the rational animal. For the Aristotelian all men possess the same inherent form or essence in virtue of which they form a natural kind, and that essential property is a rational animal soul. For Locke all men are sorted by the properties of rationality and animality when we abstract from all other properties of men and find that these (and perhaps a few others) are what remains when all other differences are stripped away.

The problem with these views is that it is not clear that there is one property "rationality" that is qualitatively identical in men. Obviously the immoderate realist will have the difficulty of showing in what sense rationality is literally the same property in each man, but even the true moderate realist (as well as the immoderate realist) will have to defend the view that Jan's rationality is qualitatively the same as Marsha's. Yet no two people "reason" exactly the same: they possess different knowledge and beliefs, have different intellectual capacities, speak different languages, and so on. The anti-realist objection is that if we strip away everything that is not *identical* between two different things, we will find nothing left—that is, no *qualitative* property whatever—and if we seek to intuit an essential sameness, we will search in vain.¹⁶ Let us see if such an objection has teeth against the various conceptions of species considered above, which I identified as having realist notions of the nature of species.

The Ecological Conception

Is there any one real property that is the same in or among organisms which share gene flow with each other? Or which are adapted to the same niche? Or which are all part of one lineage? The biological and evolutionary conceptions *prima facie* have much stronger chances, so I will first consider the ecological conception and then discuss those harder cases. On the ecological conception for there to be one property that is the same between all members of a species, there must both be such a thing as a single niche and it must further be the case that the members of a species are adapted for it in the same way. Significant doubt is placed on the first point by dialectical critiques of the niche concept that understand the nature of every niche to depend on the organisms within it, which change the environment around them in their struggles to survive and reproduce (Levins and Lewontin 1985, 53ff.).

Yet even if we can hold a niche constant while varying the members of a species within it, we must recognize that there is variation among these members with respect to their adaptedness or fitness.¹⁷ Indeed, there could be no such thing as natural selection unless there were differences between the fitness of different members of the same population (and therefore, the same species) to compete for resources in their environment. Thus if species concepts are to be based on fitness for a niche, this fitness cannot be an identical property shared or repeated in the species members. Two different members of a species may struggle to survive in *nearly* identical ways. But it is unlikely that any two organisms (at least those produced by sexual reproduction) will be able to achieve exactly the same goals given the same environment.

The Biological Conception

The biological species conception must deal with certain problems of transitivity mentioned above, and it attempts to deal with them through the phenomenon of gene flow. But is there any qualitatively same property between all the members of a species that constitutes gene flow?¹⁸ I am not here questioning whether there is such a phenomenon as gene flow, only that gene flow is one and the same property shared by the members of a species or repeated in them. Organism A stands in certain reproductive relationships to other organisms because of A's genotypic and phenotypic constitution. Organism B might have the same set of relationships as A, but surely not every member of a species will be able to reproduce viably with all the same organisms as every other member of the species. If occasionally this is the case then it is a contingent fact. By contingent I mean that it is not something our concept of gene flow does or could depend on our knowing. We would (and do) attribute gene flow to what is, according to the biological conception, a species even if we knew that the members of that species had different sets of reproductive relationships. Among other things, the idea of gene flow is meant to *solve* the problem of ring species which represent the most difficult and extreme form of the variation in potential reproductive partners I am considering. Indeed, given what we know about the development of species and the variations within them, it would seem to be an obvious mistake to assume that this property will be privy to no ranges. For the very reason that species eventually diverge and members cease to have gene flow with others, the biological conception must allow gene flow to be based not on organisms having the *same* set of reproductive relationships, but *similar* ones which eventually become less similar.

An opponent wishing to maintain and to defend a realist interpretation of the biological species conception will argue that there is a relevant property that is the same here to base gene flow on. That property is membership in a connected net of reproductive relationships.¹⁹ I would challenge the defender's claim that membership in such a network can legitimately be conceived of as a sameness in attribute(s). Here, the sameness between members is understood as their mutual possession of the same property, where 'same' means either numerically one and the same (for immoderate realism) or numerically distinct but qualitatively indistinct (for moderate realism).

If we ask *why it is* that any particular organism belongs to such a network or what membership in such a network really amounts to, then we will have to describe a series of distinct and particular reproductive relationships between organisms. It is these particular relationships that are real and primary, while their aggregate "network," while not illusory, is derivative. What it is for organism A to be "part of the network" is its actual connections—for example, to be immediately connected to B, which in turn is immediately connected not only to A, but also to C, D, and E, with E being connected to F as well as to B. And thus: for F to be a part of that network is for F to be immediately connected to E, which is in turn connected to B, and so on.

A and F might both be members of the same network, but membership in that network amounts to very different things for each them. For A it means having a direct connection to B; a two-step connection to C, D, and E through B; and a three-step connection to F, through B and then E. Membership for F consists in a direct connection to E; a two-step



Figure 12.1

connection to B through E; and a three-step connection to A, C, and D, through E and then B. Their "memberships" are not the same things. In comparison to C's membership, A's and F's memberships may not even count as similar.

A final attempt to defend sameness would be to say that what is the same about membership in the network simply is to be related, in any way whatsoever, to the same set of organisms. A is variously related to B, C, D, E, F; while F is variously related to E, D, C, B, A: so there is something the same about them. If this is the case, then to know what counts as a species requires knowing a great deal about the structure of such a network. This interpretation of sameness is not a very robust formulation of the gene flow condition, but precisely in virtue of its thinning generality it seems a strong candidate for being a real sameness.

Yet in becoming so general, this formulation of sameness in membership reveals an overall weakness in realism as such. The realist is so concerned with finding a sameness that he is willing to sacrifice a more informative system of classification. A species conception that gives some weight to the relative differences in sets of reproductive relationships including distance to network neighbors will likely tell us more than one which abstracts from this information. So unspecific a conception as the one my opponent has just suggested is more likely to have gaps between itself and other properties we deem relevant to species membership including niche adaptation and genotypic and phenotypic similarities. In the quest to find a single sameness the realist will lose precisely what his essences were meant to furnish him with in the first place—namely an explanation for why members of a kind can be grouped together.

However, the more telling objection against the realist's move is that this conception of a species makes a species not a concept at all but merely a set of particulars. A concept or universal has some general property or properties (the same or similar) that unite(s) its members. It is often said that a legitimate real or nominal essence would make no specific reference to particulars in its definition, or if it did, this reference would not be primary.²⁰ It is clear that such a definition of a biological species as discussed above is less an abstraction than a system of particulars united by certain causal relationships or potential relationships (in this case, gene flow). At this point the biological conception no longer names a universal but an individual in a limited sense. At the end of the essay I will return to a discussion of species as individuals or systems, as I prefer to call them, and how this relates to the similarity conception of species for which I am arguing. For the moment what we have found is that the biological conception cannot find a single, same property that is universal in species. All it can do is point to organisms that are related to each other reproductively to some extent, but that causal relation is not a qualitative sameness relation.

The turn to species as causal systems is a step beyond even immoderate realism, for immoderate realism took as its standard for kind-membership the participation in a single shared property by the particular members. If the shared property is "belonging to the causal group," which is not a true qualitative sameness, then the standard for belonging to the kind would be, as a trait, nothing over and above belonging. This does no better than the radical nominalism that says a thing is φ if it is one of the things called " φ ." The only difference is the added requirement that these kinds have certain kinds of causal relationships between each other. Hence the rejection of kind-conceptions altogether, in favor of an individual-conception of species.

The Evolutionary Conception

As I noted earlier, the boundaries of branches of the tree of life will be determined by speciation events, and these in turn will have to be defined by the biological or ecological species conceptions (or some other variant). These defining criteria will either take the form of real universals, which we found to be lacking, or in the case of the biological conception, of a set of individuals connected in some nonarbitrary way. The former, I argued, was a mistaken path. As a consequence, the evolutionary conception will also fail to find real essences uniting species members if it must begin from the biological or ecological conceptions. Nor will it be able to find a real essence in the lineage from an initial group however it is defined, for the individuals in that lineage will not have identical genealogical relationships: organism A is the progeny of B and C, and D is the progeny

of E and F which are not the same as B and C, regardless of whether B, C, E, and F have common ancestors.

Thus the realist species conceptions that were on the table will not work. *Mutatis mutandis*, I suspect that other putative samenesses that any other realist conceptions put forward will likewise fail. We have two choices. The first is that we cease to look for sameness and switch to similarity. When we see that similarity does the job well, while as realists we have at best a distant hope that we have found a "real kind" whose sameness will never be exploded, then we can give up the realist framework as a metaphysics without epistemological credentials. The second option is that we hold fast to sameness and carry out its implications to the point where we reject the notion that species are groups united by logical relationships holding between universal properties, and treat them as individual systems whose parts are united into a whole by causal interactions. If this zeal for sameness is the only reason for adopting the evolutionary conception, then its motivation is suspect.²²

An Epistemological Argument against Moderate Realist Metaphysics

I have argued that kinds of qualitative sameness sought by the current realist, kind-conceptions of species are not really to be found. We can go further, though, and argue against realist approaches to species or kinds as such. If one argues that there is no identical genetic, phylogenetic, morphological or other property shared by the members of any recognized species, then the moderate realist can always push the argument back by pointing to some deeper essence or homeostatic mechanism which, at the moment, might appear to be identical in the various members. However, as soon as a difference is discovered at that deeper level, the sameness they have appealed to will be exploded. This means that what counts as a real kind is always subject to future refutation as new knowledge becomes available. Since we are not omniscient and do not know what future distinctions we will be able to make as new data become available, seeking out kinds that are "real" or "natural" on realist standards is epistemologically foolhardy. Making such metaphysical distinctions will only result in a wrongheaded repudiation of perfectly useful and informative categories (of similars) that turn out not to be real (i.e., not to be the same). To help elaborate and illustrate the argument, consider the following thought experiment:

In atomic chemistry and physics the elements are thought to be natural kinds because all their instances have the same atomic number. Even if their isotopes and ions differ in their number of neutrons and electrons, the proton count remains the same. Let us say, just for argument's sake, that our present forms of measurement fail to distinguish any differences in mass or charge between protons or the up and down quarks that constitute them. Sameness and thus the realness or naturalness of these kinds is assured. Now, suppose that someone finds an even smaller component of the proton, call it a "protonite," and further discovers that while protons have on average 1,000,000 protonites each, their number varies in a small range, going as low as 999,950 and as high as 1,000,050.

On the standards of moderate realism, 'proton' would no longer denote a natural kind. Indeed, it never really did (assuming some degree of externalism about meaning). Rather we would need different "isotopes" of the proton. Suppose they are identified by the number of protonites by which they exceed or fall short of 1,000,000—so for instance, a proton with 1,000,008 protonites would be called a "proton₊₈." Given the flaw in the old 'proton'-kind, the elements, which are defined by their number of protons, would also turn out not to be real kinds. But if, say, all the protons in a given nucleus had to have the same protonite number as all the rest, then element subkinds could be named. For instance: in this atom of "gold" all 79 "protons" are really all proton₋₁₇, and in that atom of "gold" they are all proton₊₂₃. Hence the first is really an atom of $gold_{17}$, and the second is an atom of $gold_{123}$. If the protonite numbers of the "protons" in any given atom tended to vary, then it might well be impossible to rig up new elemental subkinds on the basis of protonite numbers.

There was no way we could have known about the protonites when we started using the proton concept. However, it was appropriate to form the proton concept at that time, given everything we knew, and it is still appropriate to use it and the related element concept in the many contexts where the minute differences between gold₄₈ and gold₄₄₅ are indiscernible and negligible for our practical and cognitive purposes. We would have to know everything that was to be discovered to know which kinds, if any, had real sameness. We do not have such knowledge, and these postulated real essences play no role in our formation of concepts.^{21,22}

Finally, if I am correct that concepts are objective and suitable for science even though they be based on similarity rather than sameness, then the rejection of "proton," "element," and all such debunked kinds simply in virtue of their members not being exactly the same is a grave epistemological error.²³ Valid and much-needed concepts are being stigmatized as "unnatural" or "unreal" in the name of an unrealistic and unnecessary metaphysical standard of sameness.

5 Can Similarity Be Objective?

The Multiplicity of Variations

To argue for a similarity conception of species, two challenges have to be addressed. The first, addressed in the next section, is that a conception of species based on similarity should be able to account for more than morphological or genotypic information, but can contain the important information emphasized in the conceptions I have already mentioned. The second and ultimately more significant challenge is to show how similarity can be an objective basis of species concepts or any concepts at all, as it is widely thought that similarity is an overly subjective and/or theory-laden way of organizing natural phenomena that realists think ought to have "natural divisions." I will deal with the second challenge first since it is the more general and pressing philosophical concern.

Dogs are similar to cats insofar as both are domesticated, and in the same respect dogs are dissimilar to wolves; but in numerous other respects dogs are far more similar to wolves than they are to cats (cf. Aristotle 1831, *PA* 643b3–8). There is no question that the second set of similarities is more relevant to biological classification than the first, so how are we to decide which properties to regard as salient in determining overall similarity, biological or otherwise? Moreover, even if we can determine which dimensions of similarity and dissimilarity are relevant to classification, is not similarity merely a subjective experience of *familiarity* between different things? And if similarity and dissimilarity are variable, at what point does one thing become *decisively* similar to another, and when does it become dissimilar (or are the borders inherently vague)? A defense of the objectivity of basing concepts on similarity must be able to answer these questions.

Purposes of Concepts and Scientifically Significant Similarities

Species concepts can weigh many different dimensions of similarity and dissimilarity in determining membership in a kind. For the purposes of scientific concepts of species, those properties which should be weighted more heavily are those which our best science shows to be most responsible for, explanatory of, or well correlated with the largest number of other relevant characteristics.²⁴

Biology is the science of life, thus the properties that are most biologically significant are those most fundamental to the basic life-processes (growth and development, survival, reproduction and/or replication, death, etc.). Thus the most important characteristics for *our* species-sorting criteria to be able to preserve are those that have direct bearing on organisms' mode(s) of ontogeny, survival, reproduction, phylogeny, and genetic and morphological structure.²⁵

Dimensions of variance that prove to be trivial can be taken out of the judgment process altogether or reserved for making finer distinctions between subspecies. For instance, 'domestication' will be swamped out by much more informative and predictive dimensions of variation such as lineage, reproductive community, comparative morphology, and genetic resemblance. On such bases individual dogs will be classed together as a species, and the same criteria can be used to determine higher-level taxa where dogs would be much more closely related to wolves than to cats. If a distinction is to be raised between domesticated and wild dogs it may be entirely valid for many purposes, and scientists can decide if the differences that account for domestication are of the sort that should count as being biologically significant enough to justify dividing wild and domesticated dogs into subspecies.²⁶

It is necessary to give some precision to this notion of "scientifically significant." To motivate the need for this refinement and to help make it, it will be useful to consider the case of the so-called "lily." Lilies are a favorite example of those philosophers of science who wish to point out that many of our common concepts of (biological) kinds are not "natural" (Dupré 1981, 74–75). The various flowers that the average person would call a lily include plants from a wide variety of chains of descent possessing great variance in their genotypes. What they share is only their overall, outward appearance (at least to the untrained eye).

I have already shown why an immoderate or moderate realist notion like "natural kind" will not work as a basis of concepts for organisms, and why it is an epistemologically fallow theory of classification as such, for it is at best a matter contingent to the way we form and hold concepts that we will ever be able to find exact sameness among properties in nature. What we overwhelmingly tend to find are strong similarities which are the actual basis of our concepts. If one bases concepts on similarity, then one simply cannot make such a distinction and say that "lily" is not a species *only* because it is not a "natural" kind: nature is not pre-carved for us at the joints.

How we carve natural phenomena into concepts will depend on our purposes. By saying this I am not rejecting a basic kind of realism about the objects of concepts in favor of a thoroughgoing pragmatism. Human beings organize the units of concepts, but in so doing they do not directly change them.²⁷ The units' own intrinsic properties, seen but not obscured

in the light of cognitive and practical human needs, form the basis of division and collection. However, as I pointed out above in the case of dogs, there are many different dimensions of variance between units, especially objects as inherently complex as organisms. For everyday purposes it is useful and legitimate to group together domestic cats, dogs, birds, and so on as "pets," and thus to separate poodles and parakeets from jackals and dingoes—yet it would be a disaster if *biologists* organized Siamese cats, borzois, and boa constrictors as "pets" in a higher taxon, while placing lynxes, wolves, and cobras as "wild" in a contrasting taxon. Similarly, "lily" is a fine and good rubric for the flower shop—but a terrible species for the laboratory.

It is at this point one should invoke the notion of "scientific significance." Not every concept we use on a day-to-day basis needs to be based on the properties that our most thoroughgoing investigations of nature show us to be fundamental, where 'fundamental' again means causing and being explanatory of the largest number of other facts. The role a tomato (the part of it we eat, I mean) plays in the lifecycle of the tomato plant explains much about it physically, while how it best serves in various culinary enterprises explains little about the tomato plant's reproduction and survival. It is perfectly valid to classify it as a "vegetable" in the kitchen since we have a cognitive and culinary need to organize those plants which serve well in salads together, and to separate them from the culinary class of "fruit" which is generally much sweeter.

However, the scientist is interested in organizing and explaining the facts within his ken; for the purposes of conceptualization he should choose to focus on criteria which are fundamental in the above sense. The culinary system of classification that treats tomatoes as vegetables rather than fruit is no less "natural" or "valid" or even "objective" than a biological one that makes no systematic use of the vegetable/fruit distinction at all. Both approaches and their subsequent treatments of the tomato plant are perfectly justified by real similarities and differences between different plants, and both need to be organized into retainable groups by humans in the form of concepts. However, the former organization is not scientifically significant and the latter is, because the latter classifies tomatoes on the basis of much more biologically significant characteristics.

Variable Weighting Criteria

Species concepts and concepts of higher-order taxa²⁸ should, then, have their boundaries determined by similarity in the properties that best explain and predict the most facts relevant to the central phenomena of

life: survival, acquisition and use of energy sources, reproduction, and so on. Information about organisms' morphology, genotype, reproductive community, and fitness are probably all highly relevant for grouping them in scientifically significant ways.

The weight of each individual criterion will vary among different kinds of organisms, though. For instance, facts about similarity in reproductive community should come to play a less significant role in categorizing many kinds of plants and certain kinds of animals due to the high frequency of their hybridization, while such facts necessarily cannot play a significant role in classifying asexual organisms.²⁹ This variation in the weight of criteria is not a weakness of the similarity-conception approach, but a virtue. How we categorize organisms, our systematics, ought to be responsive to what we have already learned about them, and occasionally reclassifications will be mandated. We know that many plants and even some animals differ from others with respect to the commonality of hybridization and nondestructive polyploidy. Consequently we know that we cannot always reliably group the most significantly similar plants and animals on the basis of genotypic similarity or gene flow. For organisms that are found to hybridize seldom if ever or cannot survive polyploidy, these factors can play a greater classificatory role. If it is objected that this "pluralism" is only a thin veil for a lack of any unified standard, then we should respond by challenging the monistic notion that all types of organisms should be organized in exactly the same way.

Although it would be highly desirable if one simple algorithm for classification could be used to identify fairly distinct species boundaries among the immense variety of organisms on Earth, the sheer diversity of life and the fact that the organization of scientific species concepts must reflect detailed knowledge about these different kinds of life make it unlikely that such a unitary standard exists. If insights into the nature of life more fundamental than contemporary genetic or evolutionary theory should reveal the grounds for such a method in future, as they might, then it should be adopted.

More importantly though, if no such method can ever be found this will not invalidate the particular ways in which a botanist as against a zoologist, or an ichthyologist as against an entomologist will sort similarities and differences among the organisms within his or her domain. It is not "anything goes." It is for biologists to determine the key dimensions of variation on which to base species divisions, but these are based on what they discover about these organisms and are far from unlimited. Presumably very high-order taxa will have general rules about classifying species within them (e.g., do not weight gene flow too heavily when organizing plants), and more specific taxa below them will have more narrow modifications (e.g., different families of beetles should be distinguished from one another on the basis of such and such).

Just as realism about universals makes the mistake of believing that only the sharing of *one same* property makes a particular species concept objective, forms of systematics based on realism err in assuming that taxonomy as a whole will be objective only if there is *one sort* of same property that every species will have in common among its members.

Ultimately, we need as many species concepts and ways of organizing them as are mandated by our practical and cognitive needs to deal with what we observe. In saying this I do not think I am advocating pluralism in the standard sense of "promiscuous realism" (Dupré 1981, 82ff., Kitcher 1984, 320–327). I am claiming that there is one proper kind-conception for species: the similarity conception. However, the traits that will count the most in determining which organisms are similar enough to be grouped together may vary.³⁰ There *might* even be multiple ways of classifying the same organisms into different species and higher taxa,³¹ but that is a matter to be determined by actual scientific practice and theoretical advances.

How Similar Is Similar Enough?

It is beyond the scope of this essay to give a thoroughgoing explanation of how precise concepts can be formed on the basis of similarity. However, there have been numerous arguments to the effect that similarity cannot be made rigorous and that any mature science will dispense with similarity in favor of precise nomological relationships.³² The most significant worries to address, some of which were raised aporetically above, are the concerns that similarity is inherently subjective or theory-dependent and that it is inherently vague.

Much of the basis of the worry that similarity is theory-dependent is tied up in the issue I have already addressed: there are many dimensions of similarity, and each requires at least tacit specification of a foil. I have suggested how this issue can be dealt with in terms of our cognitive purposes and, in the case of species, scientific significance. However, there is a further worry: whether A will or will not be similar to B, even in some specified respect, will be a subjective matter.

Whether two things are similar or not depends on the "foil" against which they are compared and seen to be similar. A poodle and a Chihuahua are similar to one another relative to a wolf;³³ a wolf and a dog are similar to each other relative to a domestic cat; and a cat and a dog are similar to one another relative to a flat worm. These varying levels of similarity wherein organisms are sorted as similar against their nearest neighbors (i.e., those organisms most like them yet still falling outside the class) form the basis of a hierarchical system of taxa. As one descends from the kingdoms down toward the infima species one may define³⁴ the subclassifications as narrower ranges within the ranges of the immediately higher taxon. So whether two organisms are counted as similar depends on what they are compared to and in which respects they are compared, but the attributes they possess and by which they are sorted are facts in the world, as are the causal consequences for their lives that result from whatever those specific attributes are.

Advantages of the Similarity Conception

There are numerous positive consequences for this position that follow from this approach to forming species concepts. It is an important component of my view that the species level is not more real or somehow different in kind from higher taxonomic levels such as genus, family, and so forth. 'Species' might be reserved for groups that witness particularly robust similarities especially with regard to reproductive community and genotype, but a species is not, therefore, any more or less real than a genus. Some realists might regard this as a disastrous result, but if we are willing to accept the realist position as untenable and observe that similarity-based concepts can be both objective and reflective of the large amounts of scientific information, then we will be pleased to see that certain phenomena that already occur in taxonomy can now be better accounted for and no longer appear to be problematic.

For instance, it is frequently objected that higher-order taxa vary too much in terms of the standards they are reached by and the points at which their divisions are drawn. Yet on the similarity view this variety is neither surprising nor illicit. I have already discussed how different broad classes of species can vary in the weight they give different attributes in determining similarity. Another key point is that the number of levels of higherorder taxa between infima species and kingdom need not be rigid on this view. Some species may be divided at the traditional seven taxonomic levels, others at more than seven, and perhaps some at fewer. The number of divisions necessary will depend on the number of distinctions biologists find it essential to make in elaborating the system that best balances informative classificatory distinctions with retainable simplicity.

Additionally, the range of possible "measurements" that may constitute a concept may have fuzzy edges. When one rejects the realist view of kinds then borderline cases cease to be a major problem, as it is recognized that similarity comes in degree while sameness is all or nothing. It is sometimes suggested that realist concepts suffer the problem of being either-or, all-or-nothing sorters whereas the species they are meant to capture have vaguer boundaries. It would seem to follow that similarity-based concepts are more adequate since their permissiveness of borderline cases can better accommodate the imperfectly established boundaries of actual species.³⁵

Though such an argument would weigh in on the similarity side, the committed defender of the similarity view will not use it since it presupposes such a thing as a "real species" which is carved at nature's joints and independent of the species concept. On the similarity view, individual organisms and their similarities and differences are in nature, but their classification into species (genera, families, etc.) awaits the concept-forming minds of human observers. Thus the similarity view has the advantage of requiring a more modest and plausible ontology with regard to kinds, and it admits that not every identification of an organism under a species concept will be easy or clear-cut precisely because life is so complex and variable. For why should it pretend to be able to do what no theory can? It is a better conception insofar as it jettisons such impossible standards for immaculately clean-cut boundaries for species.

6 How Informative Can Similarity Species Concepts Be?

Similarity and the Biological Conception

Above I have argued for a conception of species that considers them as kinds but would abandon a realist approach to these kinds. I argued that the identical properties meant to unify natural kinds could not be found (or at least are almost never found), regardless of whether one looked for them as gene flow according to the *biological* species concept, niche adaptation following the *ecological* concept, or monophyletic descent as per the *evolutionary* concept. I then claimed that our species kinds were based on similarity,³⁶ and defended the objectivity of concepts based on similarity. I also pointed to a number of the ways such an approach can be applied advantageously to species concepts. I now wish to briefly argue that the information realist conceptions seek to capture can be understood in terms of similarity relationships, and so be incorporated into species concepts, as I have claimed.

Let us first consider the biological conception. Reproductive relationships can be described in terms of similarity in a number of ways. One way is to determine the bases of reproductive viability—that is, the morphological or genotypic traits that sexual organisms must have in order to procreate successfully with one another. These attributes can be conceptualized—again, on the basis of similarity—and can be given weight in determining similarity *qua* species concomitant to their roles as the more or most important (only?) bases of reproductive isolation and gene flow.

Another way to factor reproductive community into determining similarity would be to take stock of which organisms each organism can reproduce with viably. Two organisms of a species may have a few differences in regard to whom they can reproduce with, but so long as there is overlap they can be regarded as similar in this respect. Significant overlap here would be decisive as compared to a total lack of overlap among nonmembers of the species, and marked when compared to a partial overlap that may occur in ring species. For instance, A and B can reproduce with each other and with organisms D, E, F . . . Z, but while B can reproduce with C, A cannot. A and B yet have *very similar* sets of potential partners, so reproductive community can be accounted for by a similarity concept in this way.

Similarity and the Ecological Conception

In section 3 of this essay I argued that the ecological conception most obviously failed to achieve the sameness that a realist interpretation of kinds would demand. It seems abundantly clear from the very fact of natural selection that variation in fitness must exist within a species. The claim that all the members of a species are more or less adapted to the *same* niche does nothing to help the realist case since the niche to which they are adapted must itself be understood as an environmental range to which members are more or less well adapted. This is because of the fact that just as species evolve through time within their environment (and across environments), so too is their environment changed by their actions within it. "Niche construction," if taken seriously, means that the environment the organism struggles in is always changing. So the members of a species must be thought of as being adapted to what is *essentially* the same niche (i.e., extremely similar), but not literally the same.

A serious worry among evolutionary theorists is that accepting the fact of niche construction makes study of the adaptation and evolution of organisms much more complicated, perhaps insurmountably difficult. With the similarity concept of species and a concept of niche based on similarity, one could hopefully accept the insight that the environment is transformed by organisms' attempts to survive and flourish within it and at the same preserve intelligibility within it by keeping stable reference. If one uses the impossible standard of sameness to judge what one observes, then from moment to moment it must be said that the environment has changed—how then could one speak of an organism being adapted to *this* environment over time?

Similarity escapes this Heraclitean morass by showing that *some* difference is inconsequential relative to greater differences. It is true that when a tree's roots absorb water from a lush forest's soil, this soil has itself lost some moisture; but compared to a desert, or even a dryer forest, the soil and the forest are still of their type. Two trees are *ecologically* similar if they will both fare similarly well doing very similar things in environments that are similar.

Similarity and the Evolutionary Conception

Lastly, let us consider one way in which the phylogenetic information focused on by the evolutionary conception might factor into similarity determinations of species. Though there may be other ways to preserve it, I think the best way to let it factor into determining whether organisms A and B are similar (relative to some foil) is to weigh the similarities between A's and B's ancestors. If A's parents were cats and B's were iguanas, then A and B will not be members of the same species. If both organisms' parents were similar enough to be members of the same species, then, unless a speciation event has occurred (signaled by other dissimilarities between A and B), A and B will very likely be members of the same species.

I grant that this way of representing phylogenetic information in terms of similarity does not emphasize species as being monophyletic branches of the tree of life. However, given the possibility of diachronic transitivity mentioned above, the advocate of a similarity approach to species ought not to be committed to the idea that every monophyletic branch will be one and only one species.

7 Species as Systems

Systems versus Kinds

Throughout this essay I have argued against a conception of species as kinds based on sameness and argued for a conception of species concepts based on similarity. Thus I have not dealt directly with an approach that treats species as individuals rather than kinds. We sometimes study a group of objects as a class—that is, we study them as *units of a concept united by a similarity relationship*. At other times, however, we treat them not as a
class but as a system—as *parts of a whole related causally* through various interactions.

For example, sometimes we wish to know the properties of all the instances of a certain kind of salt, and to achieve common, predictable knowledge we must conceptualize all the possible, particular samples of that salt as a class. But at other times we wish to study how samples of that salt behave with other compounds in some solution. There are concepts of certain chemical reactions, but a particular chemical reaction which has particular chemicals interacting is not itself a concept: it is a system. We study systems and can conceptualize them (e.g., reactions A, B, and C are all instances of hydrolysis), yet there is a difference between a system, which is a number of things united causally, and a kind, which is a number of instances conceptually integrated by similarity.

Such a system is what I think philosophers and biologists have in mind when, following a strong form of the evolutionary species conception, they refer to species as "individuals" (see Hull 1978; Ghiselin 1997). I should prefer to reserve that term for singular, unified, physical objects. A series of organisms that reproduce with one another and live in the same niche (as well as their progeny, their progeny's progeny, etc.) is in this sense a number of individuals "held together" by their causal relations, not their similarities per se. Obviously their similarities matter to their causal relationships—for example, there is nothing in the causal history of the evolution of my dog about his ancestors mating with toadstools—but grasping a system of many particulars at once is based on seeing them in their causal interactions, not on conceptualizing them from their similarities and differences.

Nevertheless, understanding any kind of sophisticated system presupposes conceptualization of its relevant parts. In the salt example, we would learn nothing of general interest unless the various sample reagents in the reaction were already classified. In the case of biological systems such as populations we are even more dependent on typological species conceptions. How would we know, short of identifying individually all the members of a population, just which organisms we are studying in the system unless we can have recourse to thinking of a *sort* of organism? For instance: could we fruitfully discuss the ecological dynamics of a given forest if we could not speak collectively of the "birch" trees, but only say tree₁, tree₂... and so on? Even if we had resources allowing us to catalog every tree-specimen in a given area, how would we know which individuals to count as a tree and which to exclude unless we already had some pre-established criteria for selecting what could count as a member (tree₁, etc.)?

And then: how could we distinguish where the deciduous forest ends and the coniferous one begins, unless we had more specific taxa concepts?

If biologists can learn important things about living things, especially about their evolution, by studying phylogenetic systems, then it is their prerogative to do so. Since such systems are very different theoretical objects than species concepts as I described them above, I think it is best to have two separate terms for these two sorts of things. Call the latter a 'species', and the former a 'line of descent', 'monophyletic branch', 'cladogram' or what have you. If it is insisted that both be called "species," then let this at least be done with the understanding that these are two significantly different senses for and *approaches to* "species."

What I object to in the evolutionary individual approach is the claim that the theory of evolution mandates that we understand species only as individuals. Space does not permit me here to discuss at length my arguments against Ghiselin's and Hull's claims that modern evolutionary theory demands that we think of species as individuals; in essence, my view is that their argument is non sequitur. Species are not things that turn out be kinds or individuals. Only particular organisms exist: how we treat them cognitively depends on what we need to know about them. If we wish to understand what they have in common synchronically (and how that similarity is maintained or not diachronically) we shall have to conceptualize them as the units of kinds, that is, we understand them as units or members of a similarity class. If we wish to study them causally interacting with one another, phylogenetically or otherwise, then we will regard them as the *parts* of a system—but even this, as I argued, does not preclude and indeed still requires that we conceptualize them into kinds. Thus, there is no metaphysical question of what 'species' will turn out to be; there is only the scientific, epistemic question of which ways of studying individual living things will be enlightening—and which will not.

So now I will briefly explain why we are justified in forming species concepts—that is, delineating organisms into kinds based on similarity as against merely systems—and why even evolutionary conceptions of species as systems require conceptual/kind-notions of species as well.

Species Systems Need Species Kinds

If the *biological, ecological* or some other species concept is needed to identify speciation events, then it is clear that the *evolutionary* representation requires that species be regarded as in some sense kinds, even if that is for the purpose of studying them as systems. One could try to make phyletic studies independent of kind-conceptions by saying one will simply consider those organisms in a direct reproductive chain. However, such lines of descent will quickly prove to be unworthy of the mantle "species." Without the ability to use speciation events as cut-off points, the evolutionary conception fails to be able to distinguish one species (as kind or system) from another. If all life is traceable back to one or a small number of common sources, and if common ancestry is both necessary and sufficient for common membership in a species, then there will only be one or very few numbers of species to which all organisms belong.

It is clear that common ancestry cannot be sufficient to establish common membership in a species, but it is also unnecessary. For instance, imagine a species S with members A and B at the time of S's formation in a speciation event. A and B both find mates and form reproductive chains of descent a and b respectively. Assume that a and b have no overlap of members in the time prior to the speciation event that ends S. At some point before S ends, there are members of a and members of b which are all members of S yet lack a common ancestor within S. I myself do not know of a way to define these speciation events without reference to kinds or to properties which essentially determine kinds. However, it seems that that is just what will be needed by the defender of the view that species in no way depend on kinds, since ancestry alone is neither necessary nor sufficient.

In Defense of Species as Concepts

It is not only the case that having a notion of species as classes is necessary for viewing them as systems; it is also true that forming concepts of organisms—that is, grouping them on the basis of similarities and differences—is a necessary practice.³⁷ Hull argues that the reason that laws cannot hold good of species is that they are not kinds that can serve as terms in lawful generalizations, but are individuals (Hull 1978, 337, 354– 356). The idea that all sciences must study laws with the mathematical precision of physics is thankfully no longer widely held. Rich generalizations that lack the level of mathematical precision or necessity of physical formulas can be perfectly scientific and extraordinarily valuable, and what is more, they *do* hold good of species.

To see that this is the case we need only turn to the study of the most closely examined species on earth: human beings. Not only do we have extraordinarily detailed *general* knowledge of the human anatomy, we also study the truths that hold good of our behavior and consciousness in psychology; of our various forms of social behavior in anthropology, economics, and political science; and even what the powers and limits of the human cognitive faculty are in epistemology and cognitive psychology. In short, huge measures of the collected knowledge of mankind consist in generalizations about one species—the species collecting this knowledge.

Now, we may never acquire nor need such detailed knowledge about any other sort of organism on earth, but nevertheless there is justification for grouping organisms together on the basis of their similarities to study their common properties. If we wish to understand how to better avoid swarm attacks or better harvest honey or find the cure for the present crisis of hive collapses then we had best study honeybees in a way that will afford us the most power of generalization and prediction: we must study a kind or kinds. We need to group together organisms that are similar enough that they end up acting and reacting in the "same" ways to the "same" things. There is no substitute for this sort of knowledge with its potential for generality and the projection of properties from observed members to future members, and this conceptual approach is clearly indispensable—intellectually, and practically.³⁸

8 Conclusion

I have argued that there is a theoretical and practical necessity of grouping organisms together as kinds so that they can be studied in a way that gives us generalizations with explanatory and predictive power. For certain systems of interacting organisms, including phylogenetic ones, these classes will be needed to help delineate the members or "parts" of the systems in question. Yet these classes cannot be regarded as natural kinds sharing some same property as the realist views claim; they are concepts that sort particulars on the basis of their similarities and differences. Concepts formed on the basis of similarity can be objective, and similaritybased species concepts can take stock of the various kinds of data that most biologists and philosophers of biology deem most relevant to the classification of species.

I hope that I have gone some way in showing that the realist conception of species kinds is misleading and unworkable, while the so-called similarity conception of species has much more to recommend it than is commonly thought. At the least, I hope to have shown that adopting the latter approach has a far wider range of consequences for our understanding of taxonomy than might be otherwise thought, and that the realist conceptions of species must be able to answer several serious philosophical worries about "sameness."

Acknowledgments

This essay was improved immeasurably by the comments, questions, and suggestions of many. In particular I would like to recognize: Michael Weisberg, who guided me through its development; Matthew Bateman, for important recommendations on content and presentation; James Lennox, Jeppe von Platz, and Gregory Salmieri, who generously gave written comments; Stephen Crowley, for his terrific commentary at the 11th INPC; several other participants at the 11th INPC, including Alexander Bird, Corinne Bloch, Judith Crane, Michael Devitt, Peter Godfrey-Smith, Todd Jones, John Matthewson, Bence Nanay, Karen Neander, and Jonathan Schaffer, for their useful questions and suggestions; and not least, the two immensely helpful, anonymous referees of this volume who saved me from a number of mistakes of scientific fact and philosophical obscurity. Any errors or points of unclarity that remain are exclusively the fault of the author.

Notes

1. By 'fundamental' I intend both an epistemological and metaphysical sense. Epistemologically, I mean the criterion most capable of explaining and/or unifying the largest number of the other criteria: e.g., if similarity of descent can explain or help us to retain the knowledge of mutual adaptation to a given environment and gene flow, but not vice versa, then similarity of descent is fundamental to niche adaptation and reproductive community. Metaphysically, the fundamental criterion is the one most causally influential on the most others: it is because two cats belong to the same lineage that they both catch mice or can reproduce with one another, and not the other way around.

2. I will elaborate presently on what I mean by 'essence'.

3. Hull 1976, 1978; Ghiselin 1974, cf. 1997.

4. I use 'resemblance' widely, and do not restrict its usage to sensory-perceptible likenesses.

5. If one can answer, it is because of one's own political views or background knowledge of what views are common in Western politics—these serve as a tacit foil.

6. Building off this conception of similarity, it can be used in an even more comparative way. Within a given range of measurements, all of whose specific measurements would count as similar relative to some foil(s), certain measurements might be *more* similar to one another than others. Imagine five places to eat: four of them are celebrated restaurants and one is a dive. All the elegant restaurants (F_1 – F_4) are similar to one another in "fanciness" when compared to the sawdust-floored crabshack (P), and they can all be called "fancy." (I say "can" so as not to rule out the possibilities of intransitivity. Given a constant background, I doubt they would arise in this case, but that is not crucial to my point.) Now, if F_1 is closer in fanciness to F_2 than it is to F_3 , then it can be appropriate to call F_1 and F_2 "more similar" to each other than F_1 and F_3 are to each other. For example, F_1 and F_2 are both the height of fine dining, and there is only the slightest drop-off in elegance from F_1 to F_2 , while F_3 is a little more casual—so there is a small but noticeable drop-off in fanciness from either F_1 or F_2 to F_3 . Furthermore, suppose that F_4 , while still pretty fancy (at least compared to P), is quite a bit less than F_3 or the others. If F_1 is closer in measurement of fanciness to F_2 than F_3 is to F_4 , then it can also be correct to say that F_1 and F_2 are more similar to each other than F_3 and F_4 are to each other.



7. I have drawn the example of colored balls and this general account of similarity and its role in concept-formation from Rand (1990, 13–15, 139–140), though the other examples, including comparatives in the previous note, and many of the formulations here are my own.

8. I reject calling all anti-realist views "nominalism" because this usage fails to make the crucial distinction between conceptualist and nominalist forms of anti-realism. Simply stated they can be formalized thus:

Conceptualism There is a general representation or concept in the mind, *C*, which sorts and refers to all and only φ -particulars. A particular is φ iff it satisfies the membership criteria or definition of *C*.

Nominalism There are only particular representations, $R_1, R_2, R_3 ... R_n$, in the mind. A word, ' Φ ', can jointly refer to any φ thing whose particular representation (R_x) resembles another particular representation, R_1 , that is called to mind when one says ' Φ '.

9. It can be a *disjunctive* set according to prototype theories.

10. Parmenides unflatteringly compares this view to covering people with a sail (*Parmenides* 131b3–c11): only part of the sail would cover each person, implying that forms are divisible and each of the many only shares in a part of the Form.

11. 'Participation' is a placeholder here for whatever relationship it is by which the many φ things relate to the Form F so as to become φ . Plato considers several possibilities for what this relationship might be, including archetypical resemblance and what I call "participation proper," that is, possessing a share of or being a part of a whole.

12. Transitivity will not hold good of the can-reproduce-with relationship, but might hold good of the shares-a-gene-pool-with relationship.

13. Similarly, sterile members of a species can be included under the notion of 'gene pool' by saying that they have emerged from that set of reproductive relationships.

14. Of course, if gradualism proves to be true this will not be a viable solution, and it seems dubious that our metaphysics and epistemology of species should push us so decisively toward a specialized biological theory—i.e., punctuated equilibrium.

15. See my discussion "Systems versus Kinds," below.

16. Cf. Berkeley's criticisms of Locke (1998, secs. 9, 13).

17. I mean 'fitness' in the classical sense; I do not mean the frequency of an allele.

18. If there is no basis for the moderate realist concept, then it follows that there can be no chance for moderate Platonism. What might appear to be grounds for a moderate-Platonic kind without being sufficient for a proper-moderate-realist kind will in fact be what I call a "system" below.

19. Would the objection that such a conception of species will not work for hybridizing species defeat this response? It may be sufficient to decide against the biological species conception as such, since through hybridization clear nonmembers of a species will have to count as members of the same species as those plants (or animals, etc.) with which they can produce hybrids. But this objection might be just as damaging to a similarity-based version of the biological conception. It might turn out that butterfly A and butterfly B, which we would clearly distinguish into separate species, turn out to have strikingly similar sets of reproductive relationships, and though it may be doubted that such similarity will ever be greater between the two of them than between each of them and what were thought to be members of their respective species, gaps do sometimes exist between resemblance and reproductive isolation. At any rate, I would not like to hang my argument against sameness on something so precarious as whether such similarity relationships are sufficiently robust to maintain them in light of the hybrid objection which defeats sameness. However, see below my contention that similarity concepts can weigh considerations such as gene flow more heavily in the classification of non-hybridizing species than those that do hybridize.

20. Cf. Armstrong's discussion of "pure types" (1989, 9-10).

21. This is basically Locke's argument against real essences as epistemically irrelevant to the formation of abstract ideas (*Essay* III, vi, 9, 19–20). Martha Bolton calls this his "Idea-Theoretic" Argument (1992, 92). For an excellent discussion of the priority for Locke of this argument over his more metaphysical arguments, and of its endurance in the face of putative examples of natural kinds (e.g., the elements of the periodic table), see Guyer 1994, 129–30, 134–140.

22. An externalist could object that even if we do not know about them, the essences might still causally influence our concept-formation. However, these

essences have to *be* there to affect us. Hence, for the many good concepts that turn out *not* to capture sameness, real essences were not there to affect their formation. Therefore, their concept-formation did not involve real essences. Furthermore, if the way we formed our concepts of "debunked" kinds is also the way we formed our concepts of not-yet-debunked kinds, then we have no reason to feign the hypothesis of real essences in the latter case.

23. As we gain knowledge, concepts may be in need of revision and their units of reclassification. The standards by which they were judged similar may change or the respects in which they are similar may turn out to be less significant than other characteristics. I speak here only of cases where the sole "shortcoming" in the classification is the lack of qualitative sameness.

24. By 'relevant' I do not mean other scientifically significant properties (see below), for that would be at best a recursive, but more likely a circular definition. Rather, I mean the physical, psychical, and behavioral properties of the organism broadly construed and not other, highly relational properties depending on human interests and the like such as economic properties (e.g., tigers and blue whales are expensive to purchase, while goldfish are not).

25. I presume here that there are at least some important connections between these properties as well.

26. For example: should wild dogs have somewhat different brains which account for the fact they *cannot* be trained and this is a heritable trait which domesticated dogs do not possess, then this distinction may be scientifically significant and a scientific subspecies distinction might be appropriate.

27. I am leaving aside Hacking's "created social kinds" since Boyd gives an account of how these do not militate against any basic notion of *metaphysical realism*, by which I mean that objects of knowledge are not created or affected directly by the consciousness of the subject. I distinguish this from realism about universals, which I discussed at length above (Hacking 1991, 109, 116–117; Boyd 1991, 143–145).

28. It is a significant feature of my view that in rejecting realism in favor of similarity concepts of species one makes divisions of organisms into species essentially no different than higher-order taxa, whereas other views tend to think of higher-order divisions as significantly less "real" or "objective" than those at the species level. On my view a species is no more or less real than a kingdom, a genus, or an order: a species simply integrates more similarities, including ones based on causal interactions between members of the concept.

29. "Reproductive community" or the lack thereof does not refer here to any lateral gene flow between asexual organisms.

30. It might be objected that we could not decide what special weighting to use unless we already knew what kind of organism we had; in other words, we must

classify an organism in order to decide how best to classify them. Yet clearly we can have a general sense of what sort of organism we are dealing with (e.g., "x" is a plant, not an animal) before knowing exactly what species we are dealing with, and current science builds on earlier science even as it revises.

31. As Kitcher emphasizes (1984, 320ff.).

32. E.g., Goodman 1979, 437–438; Quine 1969, 138. Hacking's placement of some of these views, including Quine's and Russell's, in their historical context is helpful (1991, 112ff.).

33. My previous examples were restricted to individual properties. As I mentioned above, membership within a species, indeed membership in most concepts of complex entities, depends on similarity across many different measurements at once.

34. It is vital to my view that species concepts not be reduced to their definitions as both the advocates of nominal and real essences ultimately do. Membership within a species is not solely fixed by the essence of that kind, by which I mean its defining characteristics. Definitions here are regarded as serving an epistemological/ methodological role of stating the *most* fundamental distinguishing characteristic of a group from its nearest neighbors, not as statements of either inner, metaphysical essences nor of more-or-less freely chosen stereotypic traits. Thus by 'essence' I do not mean "necessary and sufficient condition," for many properties of an organism might be unique but universal to the members of a kind. Similarly, the very notion of any trait of complex organisms must implicitly have built into it the knowledge that mistakes happen in nature, so that in a suitably limited sense a severely handicapped human baby born without the capacity to reason is still human, even if reason is an essential property of Homo sapiens. Lastly, biological diversity and recognition of phenomena such as sister species requires that the manner of specifying essential traits be dynamic sometimes, though still precise. What I mean is that it is rarely the case in defining species that we have a property so clearly distinctive as human reason. Obviously we will need far more fine-grained distinctions to ramify the myriad kinds of beetles. At the same time, these properties (or property ranges, to be more precise) may not be the sorts of things that are easily assigned one term already in the language. For instance, suppose that what distinguished beetle A from beetle B was an average difference of 1mm in wingspan and occupation of different climate zones. There might be no better definitional statements of these species than: "Beetle A is a member of genus G living in environment E_1 with an average wingspan of xmm," and "Beetle B is a member of genus G living in environment E_2 with an average wingspan of x + 1 mm."

35. I do not believe there are concepts, biological or otherwise, which are not based on similarity broadly understood. Thus I am speaking hypothetically to argue why the conception of species as concepts based on similarities is the proper approach.

36. Since my goal was only to show that concepts of species based on similarity could overcome many of the objections against them and display marked advantages, I have not argued that biologists' concepts of species *are* already based on similarity. However, if they treat species as kinds, then the sorts of sameness that the leading conceptions claim to capture are only similarities, which is a very good *prima facie* reason for thinking they are based on similarity.

37. That much knowledge would be lost were we to do otherwise, see Mishler and Donoghue 1982; Kitcher 1984.

38. To take one clear example, taxonomy plays an important role in the identification and conservation of threatened or endangered species (Dubois 2003).

References

Aristotle. 1831. *Aristotelis Opera*, vol. I. Ed. I. Bekker. Berlin: Deutsche Akademie der Wissenschaften zu Berlin.

Armstrong, D. M. 1989. Universals: An Opinionated Introduction. Boulder: Westview Press.

Berkeley, G. 1998. *A Treatise Concerning the Principles of Human Knowledge*. Ed. J. Dancy. Oxford: Oxford University Press.

Bolton, M. 1992. The idea-theoretic basis of Locke's anti-essentialist doctrine of nominal essence. In *Minds, Ideas, and Objects: Essays on the Theory of Representation in Modern Philosophy*, eds. P. Cummins and G. Zoeller. Atascadero: Ridgeview.

Boyd, R. 1991. Realism, anti-foundationalism, and the enthusiasm for natural kinds. *Philosophical Studies* 61:127–148.

Dubois, A. 2003. The relationships between taxonomy and conservation biology in the century of extinctions. *Comptes Rendus Biologies* 326 (Suppl. 1):S9–S21.

Dupré, J. 1981. Natural kinds and biological taxa. Philosophical Review 90:66-90.

Dupré, J. 2001. In defense of classification. *Studies in History and Philosophy of Biological and Biomedical Science* 32:203–219.

Goodman, N. 1979. Seven strictures on similarity. In *Problems and Projects*. Indianapolis: Bobbs-Merrill.

Ghiselin, M. 1974. A radical solution to the species problem. *Systematic Zoology* 23:536–544.

Ghiselin, M. 1997. Metaphysics and the Origin of Species. Albany: SUNY Press.

Guyer, P. 1994. Locke's philosophy of language. In *The Cambridge Companion to Locke*, ed. V. Chappell. Cambridge: Cambridge University Press.

Hacking, I. 1991. A tradition of natural kinds. Philosophical Studies 61:109–126.

Hull, D. 1976. Are species really individuals? Systematic Zoology 25:174–191.

Hull, D. 1978. A matter of individuality. Philosophy of Science 45:335–360.

Hull, D. 1980. Individuality and selection. *Annual Review of Ecology and Systematics* 11:311–332.

Kitcher, P. 1984. Species. Philosophy of Science 51:308-333.

Levins, R., and R. Lewontin. 1985. *The Dialectical Biologist*. Cambridge, Mass.: Harvard University Press.

Locke, J. 1975. *An Essay Concerning Human Understanding*. Ed. P. H. Nidditch. Oxford: Oxford University Press.

Mayr, E. 1969. *Principles of Systematic Zoology*. Cambridge, Mass.: Harvard University Press.

Mishler, B., and M. Donoghue. 1982. Species concepts: A case for pluralism. *Systematic Zoology* 31:491–503.

Plato. 1901. Opera, vol. II. Ed. J. Burnet. Oxford: Clarendon Press.

Putnam, H. 1973. Meaning and Reference. Journal of Philosophy 70:699–711.

Quine, W. V. O. 1969. Natural kinds. In *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Rand, A. 1990. *Introduction to Objectivist Epistemology*. Expanded, 2nd ed. Ed. H. Binswanger and L. Peikoff. New York: Meridian.

Sterelny, K., and P. Griffiths. 1999. Sex and Death. Chicago: University of Chicago Press.

13 Species Concepts and Natural Goodness

Judith K. Crane and Ronald Sandler

1 Introduction

Philippa Foot (2001) has defended a form of natural goodness evaluation in which living things are evaluated by how well fitted they are for flourishing as members of their species, in ways characteristic of their species. She has argued, further, that assessments of moral goodness (virtue and vice) in humans are, mutatis mutandis, of the same evaluative form. (For similar naturalistic approaches see Hursthouse 1999, Macintyre 1999, Geach 1977, and Sandler 2007). If this natural goodness approach is to provide an adequate explanation of moral evaluation, issues need to be addressed at several levels. First, is this form of natural goodness evaluation of living things biologically and philosophically plausible? Second, can the account be carried over to natural goodness evaluations of human beings? Third, can natural goodness evaluations ground or otherwise explicate moral evaluations? This paper primarily concerns the first of these issues. In particular, since organisms are to be evaluated as members of their species, how does a proper understanding of species affect the feasibility of natural goodness evaluations? We defend a pluralist understanding of species on which a normative species concept, such as that employed by Foot, is viable and can support natural goodness evaluations. However, given the account of species defended, natural goodness evaluations and, by extension, the natural goodness approach, do not garner justification in virtue of employing a scientifically privileged conception of species. The natural goodness approach does not depend upon naturalism alone. It is only justified given particular metaethical and normative commitments that are independent of naturalism.

2 "Species" in the Natural Goodness Approach

In *Natural Goodness*, Foot's "constructive task is . . . to describe a particular type of evaluation and to argue that moral evaluation of human action is of this logical type" (2001, 3). The particular type of evaluation she identifies, "which is attributable only to living things themselves and to their parts, characteristics, and operations, is intrinsic or 'autonomous' goodness in that it depends directly on the relation of an individual to the 'life form' of its species" (2001, 26–27). 'Life form' is a term introduced by Michael Thompson (1995) and adopted by Foot. It refers to the characteristic features of the lifecycle of members of a species. The idea of a life form and its central role in the natural goodness approach are further explicated in section 4.

Crucial to natural goodness evaluations are "Aristotelian categoricals" (originally discussed in Thompson 1995) of the form 'Ss are F' (or equivalent), where 'S' is a variable for a species and 'F' is a variable for a predicate that provides substantive specification of the life form of individuals of a species—for example, rabbits are herbivores or warblers begin moving south in the autumn. Aristotelian categoricals are not meant to describe statistical generalities; nor do they describe incidental features of organisms. They are distinguished from statistical and incidental descriptions by their teleological character. For Foot, "Aristotelian categoricals give the 'how' of what happens in the life cycle of that species. And all the truths about what this or that characteristic does, what its purpose or point is, and in suitable cases its function, must be related to this life cycle. The way an individual should be is determined by what is needed for development, self-maintenance, and reproduction: in most species involving defense, and in some the rearing of the young" (2001, 32-33). An Aristotelian categorical "speaks, directly or indirectly, about the way life functions such as eating and growing and defending itself come about in a species of a certain conformation, belonging in a certain kind of habitat" (2001, 33).

Aristotelian categoricals provide a standard for individuals of a species, so that "evaluation of an individual living thing in its own right, with no reference to our interests or desires, is possible where there is intersection of two types of propositions: on the one hand, Aristotelian categoricals (life-form descriptions relating to the species), and on the other, propositions about particular individuals that are the subject of evaluation" (Foot 2001, 33). The natural goodness form of evaluation therefore depends upon the viability of Aristotelian categoricals. Implicit in Aristotelian categoricals is a certain conception of species—one in which conspecifics share the life form described by the Aristotelian categoricals. One possible

worry about the natural goodness approach is that what Foot and others mean by 'species' may be importantly different from what biologists typically mean. We maintain that the natural goodness approach does indeed use a nonstandard species concept, but we argue that this worry is nonetheless misguided. In the following section, we take a closer look at species concepts and argue that there can be multiple legitimate species concepts, which allows that the species concept of the natural goodness approach may still be viable.

3 The Species Problem and Species Pluralism

There is no uniform definition of 'species' in biology, but rather a host of competing species concepts. This has contributed to a family of issues known as "the species problem" (see e.g., Stamos 2003). At one level, this is simply the problem of determining which (if any) of those currently on offer is the correct species concept. The fact that biologists and philosophers of biology have been unable to resolve this question has given rise to the further question of whether there is one correct account of species we should be attempting to articulate (species monism), or whether we can accept a plurality of species concepts, each of which is useful in different contexts (*species pluralism*). A third and related question is whether species taxa are real natural categories into which biological organisms are divided based on their fundamental features. Species realism accepts this claim, and thus that species are natural kinds. The conventionalist position denies that species are natural kinds and suggests instead that species taxa represent convenient and useful ways to organize the living world into groups, but do not reflect the fundamental features of living things. The realist intuition that species are natural kinds is a large part of the motivation for attempting to develop a single species concept that articulates the fundamental features of biological organisms which divide them into natural groups. Whether or not a species concept succeeds in this, it still provides an account of the species category that spells out what sorts of features unify a group of organisms into a species and make organisms conspecific. It will indicate where the boundaries are between distinct species taxa and generate a classification of organisms. Different species concepts generate different classifications. Below are some of the most important species concepts used by biologists.

Biological Species Concept A species is a group of interbreeding natural populations that is reproductively isolated from other such groups (Mayr and Ashlock 1991, 26).

Evolutionary Species Concept A species is a single lineage of ancestral descendant populations of organisms which maintains its identity from other such lineages and which has its own evolutionary tendencies and historical fate (Wiley 1978, 18).

Ecological Species Concept A species is a lineage (or a closely related set of lineages) which occupies an adaptive zone (ecological niche) minimally different from that of any other lineage in its range and which evolves separately from all lineages outside its range (van Valen 1976, 233).

Phylogenetic Species Concept A species is a group of organisms, including a common ancestor and all of its descendants (a monophyletic group), that is the smallest diagnosably distinct such group (see Cracraft 1983; Mishler and Brandon 1987).

Phenetic Species Concept A species is a group of organisms with a great deal of overall similarity in their intrinsic characteristics, including both morphological and genetic characteristics (see Sokal and Crovello 1970).

Morphological Species Concept A species is a group of organisms that differs morphologically from others, that is, in terms of measurable anatomical features (see Cronquist 1978; Kitcher 1984; Stamos 2003).

For species monists (Ghiselin 1987; Hull 1987; Sober 1984), there can be at most one correct species concept, and thus it is an important task of biological systematics to resolve the dispute among rival species concepts, and to provide a single account of the species category. Monism is not committed to the essentialist claim that there are intrinsic (nonrelational) features that all (or even most) members of a species share, or that are essential to the organisms that make up a species. Ghiselin, Hull, and Sober all reject these forms of essentialism. Monism is committed to the view that the species category ought to be characterized by a single set of features shared by all species taxa. It is also consistent with species monism that the single best way to classify organisms into groups needn't classify them by their fundamental features. But it is hard to see what would motivate monism in that case. If no species concept divided organisms into species based on a set of fundamental features of organisms, it seems unimportant that we use only one species concept. Moreover, given the plurality of species concepts used by biologists, it is not obvious how one of them could be the best if not for the reason that it captures something fundamental about the living world that the others do not.

Species pluralists (Dupré 1993; Ereshefsky 2001; Kitcher 1984, 1987) maintain that we can accept a number of different species concepts, which need not be rivals. Species pluralists are impressed by the fact that different species concepts are used—and are useful—in different contexts. Biologists

with different concerns and different research projects are categorizing organisms in different ways, and referring to different kinds of groups as "species." For example, the Biological Species Concept is most useful when trying to distinguish groups of organisms whose geographic ranges overlap. Where populations do not overlap geographically, we typically need to use other criteria (perhaps linked to reproductive isolation, e.g., bird song) to distinguish or lump together populations of organisms as species. The Biological Species Concept also provides no way to distinguish populations of asexually reproducing organisms into species. Biologists who study such organisms are not concerned about identifying breeding populations, and so would adopt a different species concept. Nor is the Biological Species Concept very useful in paleontology, where we have little information relating to reproductive isolation. Paleontologists are primarily interested in the evolutionary succession of populations with changing patterns of genotypic and phenotypic traits. Thus paleontologists are likely to be interested in discerning similarity groups rather than breeding populations, and may be talking about morphological rather than biological species.

In other contexts, biologists may use the Ecological or the Evolutionary Species Concept, both of which would lump together as a single species populations that to do not exchange genetic material due to geographic isolation, so long as those populations occupy the same ecological niche or maintain the same evolutionary tendencies, respectively. If, however, an isolated population becomes subject to different selection pressures, such that it acquires new evolutionary patterns, it would be considered a distinct evolutionary species as well as a distinct biological species. Once it occupies a new ecological niche, it becomes a distinct ecological species. The Phylogenetic Species Concept splits all such populations into distinct species so long as they are monophyletic groups that are diagnosably discernible on a variety of different grounds. Groups that are reproductively isolated but occupy the same ecological niche or have the same evolutionary tendencies may be considered distinct phylogenetic species. In addition, once a population branches off and acquires its own evolutionary tendencies or occupies a new ecological niche, the Phylogenetic Species Concept would not recognize the original population from which it branched as a distinct species since it does not include all of its descendants, though other species concepts would recognize such populations as species.

Species pluralism helps to alleviate the difficulty of competing species concepts since it allows us to accept that these species concepts are

not necessarily rivals. Pluralists believe there may be alternative ways of carving the organic world into populations, each of which generates groups that play a role in biological theorizing and that deserve to be called "species." In addition to helping make sense of the species problem, pluralism is also naturalistically and metaphysically plausible. The monist idea that there is a single best way to divide organisms into species seems inconsistent with biological practice. More importantly, a large part of the motivation for monism is an adherence to the realist idea that organisms are divided into species according to a single set of fundamental features. But this assumption is highly suspect. The fact that there is a variety of species concepts with different practical applications shows that different features can generate useful and explanatory classifications. There is no reason to suppose there is a single set of explanatory features that is "fundamental."

Suppose, for example, that we consider phylogeny fundamental, so that evolutionary history is what demarcates populations into groups: every organism that shares the evolutionary history of a certain population will belong to the same taxon. (This taxon may be at a higher level than species, but this ensures that all the descendants of a population will be included in its taxon, and thus that all taxa are monophyletic.) But if we take phylogeny to be fundamental in this way, what do we make of the variety of other features that are used to divide organisms into groups? Can they be subsumed under the explanatory umbrella of phylogeny? This is implausible for several reasons. Features other than phylogeny are used to make finer-grained distinctions between populations than can be done using just phylogeny. Populations with the same evolutionary histories may be distinguished into species by ecological niche or reproductive isolation, for example. Phylogeny alone doesn't explain these differences. Moreover, biological organisms are as inextricably ecologically situated as they are phylogenetically situated, and ecological situatedness is crucial for understanding why organisms and populations have the characteristics they have and behave as they do. Indeed, the ecological situatedness of populations turns out to be important for understanding phylogeny, since environmental changes are crucial in explaining evolutionary history. So it is not the case that phylogeny is more explanatorily fundamental than ecological factors. Yet phylogeny does capture something important about life, which is why the Phylogenetic Species Concept is a powerful and influential species concept.

Perhaps those features that give us finer-grained species divisions are more explanatorily fundamental, since they can explain biological divisions that courser-grained species concepts cannot. But things are not this simple. Different species concepts do not straightforwardly differ in terms of making finer- and courser-grained distinctions among species. They cross-classify organisms into different kinds of groups. The general point is that species have common phylogenies, common ecological niches, common genetic features, are members of common reproductive communities, and so forth. These are all important explanatory features, yet they appear to resist reduction into a single set of fundamental features. The fact that they generate classifications of different kinds of groups, classifications which are incompatible with one another because they crossclassify organisms, suggests that this is not merely a result of our failure to understand the causal structure of the world. No single species concept identifies *the* fundamental causal structure of the biological world because there is no *single* set of fundamental features.

For these reasons, species pluralism is the more plausible view. By accepting species pluralism, it is possible to make room for species concepts that serve a variety of explanatory projects—perhaps even those of ethics. However, no pluralist should accept that all species concepts are equally legitimate. To do so would be to reject a minimal scientific realism in favor of full-blown relativism. Biological reality must place some constraints on what counts as a legitimate species concept. Otherwise species divisions would not need to correspond to anything real and any species concept would be as good as another: we would have to accept "the suggestions of the inexpert, the inane, and the insane" (Kitcher 1987, 190). But what makes one species concept legitimate and another not?

Pluralists have offered a variety of approaches to this question, which are surveyed by Ereshefsky (2001, 158–162). For Kitcher, the organisms of a species must be related to each other by "biologically interesting relations" (1984, 309). There are a variety of such relations, and various biological theories and research areas focus on different ones: to know which species concepts are legitimate, we look to the experts in biology. Those species concepts in current use in accepted biological fields are considered the legitimate ones. Ereshefsky looks to the aims of biological taxonomy, and identifies legitimate species concepts as those that promote those aims. He rejects certain species concepts (e.g., the Phenetic Species Concept) on the grounds that they do not adequately promote the aims of biological taxonomy. Both Kitcher and Ereshefsky see the aims and projects of biologists as privileging certain species concepts over others; Ereshefsky privileges an even narrower set of biological projects than does Kitcher. But an adequate species pluralism should do more than appeal to expert biologists as the arbiters of legitimate species concepts. It should explain why some species concepts are legitimate while others should be rejected. Below, we attempt to provide a set of necessary conditions for carving the living world into groups that rule out "inexpert, inane, and insane" ways of doing so; but ruling out such species concepts as illegitimate does not require that we privilege a narrow set of biological aims and purposes. These conditions rule out certain species concepts that intuitively ought to be rejected. In the absence of arguments for further necessary conditions, we will regard a species concept that satisfies these conditions as legitimate.

1. A legitimate species concept needs to classify organisms into groups of a certain kind, since the point of a species concept just is to divide and organize organisms. This is something that all species concepts are intended to accomplish. If there are no groups of the kind that a species concept recognizes, the concept fails. For a species concept to be legitimate there must exist groups of organisms that correspond to the names of the taxa generated by the species concept. Such groups must have discernible boundaries, though these need not be precise. Alternative species concepts can differ with respect to whether to count certain groups as species, but as long as those groups exist, the different species concepts might all be legitimate. For example, syngameons (or multispecies) consist of populations with distinct ecological roles, but which frequently interbreed and produce fertile offspring (Ereshefsky 2001, 4.1; van Valen 1976.) On the Biological Species Concept, the whole population would be considered a species, while on the Ecological Species Concept, the smaller populations, rather than the whole group, would be considered species. There is no question about whether the groups exist, only about whether they constitute species. On the other hand, assuming one version of creationism is false, a species concept that defined a species as a group of organisms, including an initial pair created by God and all of its descendants, would not meet this criterion. Nothing corresponds to a species name that alleges to pick out a group of that kind.

2. A legitimate species concept must distinguish organisms into taxa by features that are biological properties of organisms or of groups of organisms. A biological property is a property that only a biological entity (either an organism or a group of organisms) can have. Only biological organisms can have morphological and genetic features with respect to which we can sort them into groups. Interbreeding relations and reproductive isolation are biological properties of populations of organisms, as are evolutionary histories. Weight is not a biological property. If a species concept split

organisms into species A, weighing 50 pounds or more, and species B, weighing less than 50 pounds, it would not satisfy this condition (though it satisfies the first). Even a property like citizenship is not biological as it is not necessary that a citizen of a certain country be a biological entity: in principle, a synthetic robot could attain citizenship. Thus if we were to sort certain organisms into groups by their country of citizenship, those groups would not be biological taxa, and any species concept that generated such a classification would not be legitimate.

These initial conditions are meant to guarantee that the names of particular species taxa that fall under a species concept refer to groups of organisms that are characterized biologically. Any species concept that generates taxa the names of which either do not refer at all, or refer to groups that are not biological groups, is not a legitimate species concept. It is important that the groups are biological in the sense that they are distinguished by biological properties. But it is not required that the species concept itself be one used in some list of accepted biological fields, or more narrowly, used by biological taxonomists. What should be required is that the species concept has some explanatory function. Suppose, for example, that a species concept divides organisms into those with eyes and those without eyes: the two groups exist, and having eyes is a biological property. The difficulty is that eyes have evolved independently more than forty times (Mayr 1982, 611), so the possession of eyes appears to have very limited explanatory relevance. It provides no indication of evolutionary relationships, reproductive relationships, or ecological situatedness, and very little grounds for making predictions or inferences to other biological properties. In order to rule out such explanatorily weak species concepts, we suggest a third condition:

3. A legitimate species concept must be explanatorily useful. It must help make sense of the world in terms of organizing it, understanding it, making predictions, and so on. The features that divide organisms into species are related to phenomena we wish to explain in such a way as to contribute to an explanation of those phenomena. We wish to explain how species maintain genetic and morphological stability across many generations: interbreeding relations and environmental pressures both contribute to an explanation of such stability. We wish to understand and organize evolutionary history: dividing organisms into monophyletic groups helps to accomplish that. The explanatory usefulness of a species concept need not be restricted to the aims of biologists, however. In this respect, our version of species pluralism is more permissive than either Kitcher's or Ereshefsky's. The groups of organisms picked out by a species concept, and the features used to differentiate species, must serve to explain something in need of explanation and must organize the world in theoretically useful ways. The organic world plays a large role, and may serve as an explanatory basis, in a variety of phenomena. But there is no reason why classifications of organisms should be restricted to those that contribute to generating biological knowledge. In particular, the natural goodness approach looks to the living world for explanations of *evaluation*, including normative evaluations of human behavior. How such evaluations are possible is in need of explanation, so if a species concept can aid in this—that is, if it helps to explain and enable a form of evaluation that, when properly employed, justifies particular normative evaluations—it would satisfy the third condition.

These three conditions leave room for multiple species concepts that divide the world into different biological groups according to different biological properties. We do not believe it is necessary that a species concept "carve nature at its joints." The intuitive idea behind this metaphor is that there is a single set of fundamental features that tracks natural biological categories, so that species are natural kinds. A principle motivation for our version of species pluralism is that it makes little sense to pick out one of the sorts of features that systematists use to classify organisms as distinctly fundamental. Different species concepts identify different biologically significant features to classify organisms, and their significance is not reducible to or derivative from one type of feature. As a result, no single species concept isolates the "fundamental" properties of the living world. If no species concept does this, there is little motivation to adopt a monist approach to species, and little reason to think of species as natural kinds. If we wish to speak of nature's "joints" we should say there is no single set of biological joints in nature, but rather many, which cross-cut the beast. (The "joint" metaphor begins to breaks down here. The organic world can be carved along a variety of dimensions, so long as we carve it into biological groups-we might say, so long as we wind up with cuts of meat of some sort or other.)

A final consideration regarding the legitimacy of species concepts concerns the relevance of the species *rank*. Biologists who frame species concepts are often concerned with the question of which groups deserve to be ranked at the level of 'species' proper, as opposed to higher or lower taxonomic categories. The Biological Species Concept is particularly strong at identifying groups at the species rank (at least for sexually reproducing organisms), since it identifies reproductively isolated populations as being of special importance. For the Phylogenetic Species Concept, however, rank is less important: "species" are the groups at the tips of the branches of the phylogenetic tree. One can make divisions at the tips of the branches as fine as one likes, many of which are of little theoretical importance. (In fact, which groups are at the tips of the branches is transitory; over time, many of them will be displaced by their descendant populations.) The conditions for the legitimacy of species concepts outlined above do not include that a species concept picks out groups that qualify as occupying a special taxonomic rank, which alone deserves the name 'species'. The groups need only be real, biological groups of a certain kind, such that identifying groups of this kind is explanatorily useful.

4 Natural Goodness in Nonhuman Organisms

Foot adopts Thompson's (1995) notion of "species" or "life form" in grounding the natural goodness approach. Although it clearly has a basis in biology, Thompson does not see his species concept as one that is necessarily used by biologists, which is why he prefers to speak of "life forms." What exactly is the species concept that Foot and Thompson are using, and is it a legitimate species concept? Neither Foot nor Thompson adequately answers these questions. They find the legitimacy of their species concept in the fact that it is used in ordinary language, "natural history" descriptions of living things-the sort expressed in television nature programs and by naturalists more generally. We spell out below what we believe to be the species concept implicit in the natural goodness approach, demonstrate its role in generating evaluations of organisms, and show that it is a legitimate species concept, even if not one used by most biologists. Since this species concept helps to explain (and enable) a form of evaluation that generates normative evaluations of organisms, we call it the Axiological Species Concept.

4.1 The Axiological Species Concept

The idea of a *life form* is central to the Axiological Species Concept. Thompson appears to use 'life form' interchangeably with 'species'; as Foot uses 'life form', it refers to a feature shared by members of a species. Our usage of 'life form' more closely follows Foot's. We take a life form to be constituted by the set of Aristotelian categoricals that specify what characteristically happens in the lifecycle of members of a given species. Foot's examples of Aristotelian categoricals include *rabbits are herbivores, cats have four legs, the deer is an animal whose form of defense is flight,* and *the peacock has a brightly colored tail.*

Foot and Thompson emphasize that Aristotelian categoricals are not meant to be universally quantified, since taken in that way most of them are clearly false. For Foot, the Aristotelian categoricals that constitute a life form are teleological, for only such categoricals will yield evaluations. The blue tit has a round blue patch on its head will not yield evaluations about individual blue tits, because, as far as we know, "the colour of the head plays no part in the life of the blue tit" (2001, 30); hence there is nothing defective about a blue tit without said blue patch. In an Aristotelian categorical, there is an "expectation of an answer to the question 'What part does it play in the life cycle of things of the species S?" (2001, 32). Aristotelian categoricals purport to describe the characteristics that members of a species have in order to, for example, maintain themselves and reproduce in their distinctive way (i.e., the way described by the Aristotelian categoricals). The teleological character of the Aristotelian categoricals also ensures that they are not merely descriptions of what is statistically normal; they do not "come from the counting of heads" (2001, 31). Even if a minority of species members reaches maturity and acquires many of the features or engages in many of the behaviors described (consider e.g., sea turtles), the Aristotelian categorical is still thought to be true.

Given the nature and role of Aristotelian categoricals in natural goodness evaluations, we suggest the following species concept is operating in the natural goodness approach:

Axiological Species Concept An axiospecies (or axiogrouping) is a biologically related group of organisms that shares a life form, as described by a set of Aristotelian categoricals.

"Biologically related group of organisms" indicates that the *grouping criteria* for this species concept are biological. Grouping criteria separate organisms into groups, draw boundaries between groups, and determine whether an organism belongs to a particular group (Mishler and Brandon 1987). In this case, members of a group must be related by interbreeding relations, parent-offspring relations, sharing the same ecological niche, monophyly, or any set of biological features that groups organisms together as a biological unit. The Axiological Species Concept is not specific with respect to which biological features may be used to identify the boundaries of an axiospecies. It requires only that the features be recognizable by biologists as delineating biologically significant groups. This captures the sense in which the notion of 'species' used by the natural goodness approach is grounded in biology. The natural goodness approach attempts to ground an account of evaluation in biologically significant groups.

Grouping criteria are necessary but insufficient conditions for a group of organisms to constitute a species. Grouping criteria are distinguished from ranking criteria (Mishler and Brandon 1987), which determine which groups may be ranked as groups of the right sort. The ranking criteria for the Axiological Species Concept determine which groups of organisms identified by the biological grouping criteria qualify as axiospecies, and which are merely biologically significant groups. The sharing of a life form is the ranking criterion in the sense that only those biologically related groups of organisms that share a life form, as expressed by a set of Aristotelian categoricals, count as axiospecies. In order for a biological group to share a life form, the members of the group must share certain goods or ends, such as self-maintenance, reproduction, and sociability, which are realized in characteristic forms and achieved by characteristic means. (It is easy to slip here into talking about goods for a species, but neither we nor Foot believe it is plausible that a species as a whole has a good. The goods we are referring to should be understood as goods for the members of a species [Sandler and Crane 2006].) If it is characteristic of the members of a biologically related group to strive toward a state/activity G, such as selfmaintenance or reproduction, then G counts as a good for the members of the group. "Striving" means expending energy toward, varying behavior in manners required to achieve, and so forth. It consists of displaying forms, processes, and behaviors, under certain conditions, which are conducive to achieving G. In saying that certain strivings are characteristic of members of a group, we do mean to be saying that the large majority of members of the group strive toward G, and in characteristic ways—though these may (and often will) be indexed to sex, life-stage, or environment, for example. Foot is quite right that the Aristotelian categoricals do not depend on the "counting of heads" in that we do not count how many species members achieve their ends. But we do count the strivings-the forms, processes, and behaviors-to determine the ends of the members of the group, and how they are characteristically pursued and (when accomplished) attained.

As we understand the Axiological Species Concept, the life form of a species is shared by *all* organisms belonging to the group, as determined by the biological grouping criteria, even though the particular Aristotelian categoricals are not true of all members of the group. Some members may fail in achieving the ends, or fail to exhibit the characteristic forms, processes, and behaviors by which group members typically strive toward the ends—but they all share the same ends, and hence have a common life form. It is because the life form applies to all members of the group that

we can identify an organism as a member of a group (by the appropriate biological criteria) and then evaluate it in terms of the life form of that group.

In our discussion of species concepts in section 3, we noted that it is not a necessary condition for legitimacy that a species concept identify groups that occupy a special taxonomic rank, one that alone deserves to be called "species." Some species concepts do aim to do this, but the species rank has no special status for the Phylogenetic Species Concept, for example. An interesting feature of the Axiological Species Concept is that the species rank, as typically understood in biological systematics, is not privileged as the only level at which organisms may be evaluated. The Axiological Species Concept does aim to identify biological groups of a special sort, and uses life form as a ranking criterion to identify them. A biological group that satisfies that ranking criterion will generate evaluation standards, and it is groups that generate evaluation standards that the Axiological Species Concept aims to identify. Biological groups with discernible life forms will not all be at the species level, as typically understood in biological systematics. There are Aristotelian categoricals about placental mammals, to the effect that placental mammals have a characteristic way of reproducing, and about mammals generally, that they have a fourchambered heart. To make this point clear, it will be useful to refer to the biological groups that generate evaluation standards as "axiogroupings" rather than "axiospecies." In fact, none of Foot's examples are at the species level-deer, rabbits, and cats are all at higher taxonomic levels. Axiogroupings will also be at lower taxonomic levels. Foot is explicit that "the Aristotelian categoricals must take account of subspecies adapted to local conditions" (2001, 29). Foot's peacock example is particularly interesting. Peacocks do not constitute a species in standard taxonomies, partly because there are several species of peafowl. (Peafowl do not constitute a genus, either, but a group consisting of two genera, Pavo and Afropavo.) But the Aristotelian categorical is about just the males of this group, the peacocks, and their brightly colored tails. What makes peacocks an axiogrouping is that that they comprise a biologically related group of organisms (not related by interbreeding relations or monophyly, but by morphology or occupying a similar ecological niche, perhaps) which have a discernible way of achieving their ends—particularly that of mating with peahens. Individuals may be evaluated in virtue of belonging to such an axiogrouping, and evaluations can be generated from any of the axiogroupings to which an organism belongs. (The natural goodness approach therefore accommodates Copp and Sobel's (2004) point that there is no reason to

focus on the particular "species" to which an organism belongs, rather than the other kinds to which it belongs, when determining evaluation standards.)

4.2 Evaluations of Organisms

The Axiological Species Concept generates evaluations of organisms because the Aristotelian categoricals that constitute life forms yield *norms* for individual members of an axiogrouping. While the Aristotelian categoricals are not themselves universally quantified, they apply to all members of the group in the sense that all members of the group have a common life form, including common goods. It is the fact that all members of a group F have a set of common goods that allows the derivation from the Aristotelian Categoricals to universally quantified norms stating that all members of group F are *supposed to* have characteristic C. The norms are true of all members of the group may be evaluated. Given the norm that *all warblers are supposed to begin moving south in the autumn* and given that a particular bird is a warbler (based on the biological criteria), we are in a position to evaluate that particular bird with respect to its migratory behavior.

A crucial part of the derivation of the norms from the Aristotelian categoricals is the attribution of the life form, including a set of goods/ends G, to all members of the group. How is the extension of the life form to those that do not display its characteristics justified? The Axiological Species Concept is bounded by biological criteria. The shared life forms apply to all members, as determined by the biological criteria. This sort of extension is not unusual. Under the Biological Species Concept, infertile organisms are considered members of a breeding community, and thus members of their parents' species, though they do not interbreed with other members. Under the ecological species concept, individuals (even populations) born and raised in captivity, which live under quite different ecological conditions and fill quite different ecological niches from their wild counterparts, are nevertheless considered conspecific with them. Similarly, under the Axiological Species Concept, organisms are conspecific if they are members of a biological group with a characteristic life form, even if one of them does not fully exhibit that life form.

The Aristotelian categoricals and the norms derived from them are not intended to be statistical generalities, and it should be clear that they are not. Even in cases in which a minority of species members displays the attribute, such norms may be generated. Only a small minority of sea turtle hatchlings successfully make it alive to the open ocean once hatched. Nonetheless, this is what they are *supposed* to do, and arriving at the ocean counts as a good for all of them, as can be seen by the fact that this is what the large majority of them attempt to do. The norms are not descriptions of what would be *beneficial* to individual organisms, either. The behavior of worker bees is often not in the interest of individual worker bees, but still falls under the norm of what such bees should do. Consider too, tigers living in big cat sanctuaries. Many of these animals have never lived in the wild and could not survive in their natural habitat. While certain natural behaviors may be of no benefit to these individual tigers in their current environment, which contains neither prey nor predator, such tigers are still subject to the evaluative claims that they are not functioning as they should. They are defective tigers, which is why they have been placed in such sanctuaries.

An important feature of the Axiological Species Concept is that it relies on biological criteria for an organism's belonging to an axiogrouping, and generates distinct criteria for evaluation of an organism based on the life form of its axiogrouping. The criteria for being *an F* are distinct from the criteria for being a *good F*. If the membership criteria and the evaluative criteria were not distinct, there could be no evaluative discrimination among members of an axiogrouping. Only an organism previously determined to be an F can be evaluated as a good or a defective F.

4.3 The Legitimacy of the Axiological Species Concept

The Axiological Species Concept is not on the standard list of species concepts used by biologists. While it does defer to standard biological ways of grouping organisms in its grouping criteria, it also contains a distinct ranking criterion, one that is explicitly teleological and is meant to identify those biological groups that generate evaluation standards for their members. Given the teleological component and the focus on evaluation, the Axiological Species Concept is not likely to be useful for most purposes for which biologists need a species concept. Since the natural goodness approach is thinking of "species" rather differently than biologists do, it seems that it could not get off the ground were species monism to be true: the Axiological Species Concept is not a plausible candidate for the single best way to classify biological organisms. Fortunately, as we have argued, species pluralism is the more plausible position. Provided the Axiological Species Concept satisfies the necessary conditions for legitimacy, it can reasonably be accepted along with other species concepts.

In order for a group to constitute an axiogrouping it must be a biological group. That is, it must be a group of organisms related by biological features, recognizable by biologists as a biologically significant unit. This is sufficient to satisfy the second condition that the taxa recognized by a species concept be biological groups. In order to satisfy the first condition, it must be the case that there exist biological groups of the kind described by the Axiological Species Concept. There must be biological groups with shared life forms expressible by Aristotelian categoricals that describe the ways in which organisms within the group characteristically achieve certain goods/ends. In picking out biological groups with shared goods/ ends, the Axiological Species Concept aims to carve nature at its *teleological* joints. It looks for characteristic forms, processes, and behaviors that are conducive to characteristic realizations of ends like self-maintenance and reproduction, and identifies as axiogroupings those biological groups with such common ends and characteristic ways of attaining them.

If there are teleological features in nature—for instance, the goods and ends implicit in Aristotelian categoricals—then the Axiological Species Concept aims to use them to pick out groups of organisms that are subject to normative evaluation. If there is no teleology in nature, then the Axiological Species Concept fails. One reason for thinking there are such goods and ends is that we can observe organisms striving to attain them. Living organisms display certain forms, processes, and behaviors that involve expending energy and altering behavior in characteristic ways toward the attainment of ends. We do not observe such strivings in non-living natural phenomena. As these forms, processes, and behaviors are characteristic of certain biological groups, there are groups of organisms that correspond to the names of axiogroupings, including the peacocks, the cats, and the placental mammals.

The explanatory power of the Axiological Species Concept is shown by the fact that it is used to explicate how organisms are (and can be) evaluated as being good or defective members of their kind—that is, it explains the form of evaluation. (It does not itself justify any *particular* normative evaluations of individuals. Such evaluations are generated or justified when the form is employed to evaluate particular individuals as members of the axiogroupings to which they belong.) This is sufficient to meet the third condition. A skeptic may doubt that there is anything here to be explained. Perhaps the claim that *warblers are supposed to begin moving south in the autumn* is reducible to a non-normative claim, or perhaps it is not really a coherent statement. We have argued that these norms are reducible neither to statistical generalities nor to descriptions of what would be beneficial to individual organisms. If the skeptic is left maintaining that these norms are simply not coherent, the defender of Axiological Species Concept has a clear advantage. Such claims certainly appear to be coherent: we make them all the time, and have no trouble discerning their content. The Axiological Species Concept, and the teleological features it employs, explain what makes such norms possible, and in a naturalistic way. In addition, there is much more at stake than normative judgments about nonhuman biological organisms. Moral philosophers have struggled to make sense of moral facts. If one category of moral evaluation—evaluations of human character—is a variety of natural goodness evaluation, then the Axiological Species Concept has an explanatory role in the area of ethics.

In spelling out the Axiological Species Concept implicit in the natural goodness approach, we have shown that the approach is naturalistically and philosophically coherent and plausible. We have not shown that the natural goodness approach is the only possible approach to generating normative claims regarding individual organisms. (For an alternative approach, see Post 2006.) Nor have we shown that it is superior to the alternatives, such as an interest-based approach. But that the natural goodness approach is biologically and philosophically tenable implies that it needs to be considered alongside other possible approaches.

5 Transitions: From Nonhumans to Humans and From Health to Morality

The natural goodness approach is not undermined in virtue of employing the Axiological Species Concept, even if this is not a species concept typically used by biologists. This is not to claim that the Axiological Species Concept can do all the work the natural goodness approach aspires to accomplish. In addition, it must be established that the biological group *Homo sapiens* (or something very close to it) is an axiogrouping—that is, has a life form describable by Aristotelian categoricals which enables natural goodness evaluation of individuals of the group. Further, it must be established that this form of evaluation can be a significant part of the basis for moral evaluations. A full defense of the natural goodness approach is beyond the scope of this paper, but we offer the following sketch of how these challenges might be met.

It is clear that there is a biological group *Homo sapiens*, one that is delineated (at least) by genetic and phylogenetic criteria. Aristotelian categoricals will be as appropriate to this biological group with respect to physiology or biological functioning (including the physiological components of cognitive and psychological functioning) as they are to groups of nonhumans—for instance, *human beings have clotting factors in their blood*. This enables natural goodness evaluations of human beings with respect to health: a human being whose blood lacks the proteins necessary for blood coagulation is a poor specimen (in that respect) of its species, in the same sense that a rhododendron that never flowers is a poor specimen (in that respect) of its species.

But there are complications in the transition of natural goodness evaluations from nonhumans to humans. The aspects of human individuals which are relevant to promotion of goods for individuals of the species (e.g., self-maintenance and reproduction) are not limited to bare biological functioning, but include as well desires, emotions, and actions (from reason and from inclination) (Hursthouse 1999). Moreover, the goods constitutive of human flourishing are more diverse than those of rhododendrons, or even porpoises. They include survival, self-maintenance, reproduction, and sociability-but also autonomy and knowledge, for example (Sandler 2007). In addition, the ways in which the goods are realized (and pursued) by human beings are not nearly so circumscribed by our biology as they are with other species. Our biology constrains us to some extent-for instance, human infants cannot survive on their own and we all require some interpersonal relationships. But human social systems and approaches to raising children have varied widely over time, among cultures, and between individuals, and they continue to change and develop. Human beings have the capacity to imagine a way of going about or realizing something, judge it as good, devise ways to attempt to accomplish it, and, if it works, pass it on to others (actively or passively) (Foot 2001; Sandler 2007). Therefore, although human beings do not have a characteristic way of realizing goods in the same sense as do other species, we do characteristically go about the world in a rational way. As Hursthouse explains, "A 'rational way' is any way that we can rightly see as good, as something we have reason to do" (1999, 222). These (and other) complications require the "mutatis mutandis" qualification when transitioning natural goodness evaluations from nonhumans to humans; however, they are not alterations to the *form* of evaluation. Individual humans are still to be evaluated on how conducive their parts, processes, and behaviors are to realizing the ends appropriate to their life form. This is true even as, for the reasons above, what constitutes both these ends and the characteristic pursuit of them is not solely determined by biological facts about Homo sapiens, but rather as well by the rationality, culture, and technology which our biology enables, yet which shape and provide novel possibilities for our life form.

The transition from evaluations of health (i.e., biological functioning) to moral evaluations is also complicated. A person who has hemophilia is not a *morally* poor specimen in virtue of that condition. Moral evaluations (i.e., evaluations of virtue and vice) involve evaluating those aspects of human beings that remain after the bare biological parts and processes are separated out-namely, their emotions, desires, and actions (from both reason and inclination). But even this does not fully accomplish the transition. What is additionally required is provided by the "rational way" human beings characteristically pursue and realize the appropriate goods. (Again, characteristic features need not be exhibited by all members of the biological group. That *Homo sapiens* characteristically goes about the world in a rational way does not imply that nearly all people act rationally nearly all of the time.) That there is a variety of character traits that a human being might have and that our characteristic way of going about the world is rational provides the basis for evaluating character traits in light of their conduciveness to promoting endorsable (rightly-seen-as-good) realizations of the goods of the human life form (Hursthouse 1999; Sandler 2007). This is moral evaluation-or very close to it.

These several transitions place considerable weight on how our rationality modifies the approach from nonmoral evaluation of nonhuman organisms to moral evaluation of humans. Given the relationship between rationality and morality, it could not be otherwise. However, it does raise the question of how much work the natural goodness approach is ultimately doing once the transition is complete, and whether in the process of transition all the heavy normative work is ceded to an independent account of rationality (Thompson 2008). To be sure, the natural goodness approach applied to human character evaluations (as briefly sketched above) is not as tightly tied to biology as is normative evaluation of nonhuman organisms. Evaluations of human character traits are not merely biological appraisals. The ends constitutive of human good are not fixed by the biological facts about us. Human virtue is not reducible to good biological functioning. Nevertheless, significant aspects of the natural goodness approach remain. These include the form of evaluation (i.e., that aspects of the organism are evaluated according to their conduciveness to promoting certain ends), as well as that the biological facts about us significantly (albeit not exclusively) inform the content of the ends (i.e., human flourishing) and what constitutes endorsable forms of pursuit and realization of those ends. Moreover, that this account does not reduce

evaluation of character traits to evaluations of biological function helps it avoid some common objections to naturalistic accounts of virtue: that there are frequently situations in which being an imperfect biological specimen of *Homo sapiens* is not detrimental, or may even be conducive, to a person living well; that people often have worthwhile goals other than, or even inimical to, their biological flourishing; and that naturalistic accounts must countenance (or even require) us to act aggressively, vengefully or xenophobically, for example, if humans characteristically have impulses to act in these ways (Sandler 2007, 2008).

It is hoped that the foregoing sketch of a path forward motivates the sense that the natural goodness approach can be successfully transitioned from nonmoral evaluations of nonhuman organisms to moral evaluations of humans. If it can, then the Axiological Species Concept has an explanatory role in how a form of moral evaluation of human beings—that is, character evaluations—is possible.

6 Conclusion

In order for the natural goodness approach to explain how living things are evaluated as having intrinsic goodness, it must employ a normative conception of species, which we have articulated as the Axiological Species Concept. Given the plausibility of species pluralism, the use of such a species concept is not objectionable. That the Axiological Species Concept is in fact a legitimate species concept is shown in part by the fact that it picks out real groups of organisms that are characterized biologically and that share a common life form. Further, the Axiological Species Concept provides an explanatorily powerful way of dividing the living world into groups, by focusing on its teleological features. We have shown that the Axiological Species Concept makes possible a form of normative evaluation of living things as members of their species. This form of evaluation, mutatis mutandis, may be applicable to evaluation of human beings, including evaluations of human character traits. The Axiological Species Concept therefore enables the natural goodness approach. It does not, however, entirely justify that approach. Given the pluralist account of species, there is no basis in biology or metaphysics for privileging this species concept as more "true" or "real" than other viable species concepts. Nor is there a basis in biology or metaphysics for privileging the natural goodness evaluations it enables as the model for moral evaluations (Copp and Sobel 2004). The justification for modeling ethical evaluations on natural goodness evaluations must come from commitments in ethics, not biology. For example, the justification can come from an ethical naturalism which maintains that human flourishing should be understood from an ethological perspective (not that of one's genes or ecosystem), that virtue and vice concern emotions, desires, and actions (not bare physiological functioning), and that ethical norms have a certain level of generality (even if indexed to environment/culture, which natural goodness evaluations accommodate). Therefore, while it has been established that the natural goodness approach is not problematic in virtue of its use of the Axiological Species Concept or Aristotelian categoricals, and that these provide a naturalistic ground for evaluations of natural goodness, we have not given a full defense of the natural goodness approach as an explanation of moral evaluation.

Acknowledgments

For helpful comments and discussions, we wish to thank John Danley, Robert Epperson, Joseph LaPorte, Skip Larkin, Greg Littmann, Chris Pearson, Matthew Slater, Jim Stone, Jay Odenbaugh, and three anonymous referees.

References

Copp, D., and D. Sobel. 2004. Morality and virtue: An assessment of some recent work in virtue ethics. *Ethics* 114:514–554.

Cracraft, J. 1983. Species concepts and speciation analysis. *Current Ornithology* 1:159–187.

Cronquist, A. 1978. Once again, What is a species? *Biosystematics in Agriculture: Beltsville Symposia in Agricultural Research* 2:3–20.

Dupré, J. 1993. The Disorder of Things: Metaphysical Foundations of the Disunity of Science. Cambridge, Mass.: Harvard University Press.

Ereshefsky, M. 2001. *The Poverty of the Linnaean Hierarchy*. Cambridge: Cambridge University Press.

Foot, P. 2001. Natural Goodness. Oxford: Oxford University Press.

Geach, P. 1977. The Virtues. Cambridge: Cambridge University Press.

Ghiselin, M. 1987. Species concepts, individuality, and objectivity. *Biology and Philosophy* 2:127–143.

Hull, D. 1987. Genealogical actors in ecological roles. *Biology and Philosophy* 2:168–183.

Hursthouse, R. 1999. On Virtue Ethics. Oxford: Oxford University Press.

Kitcher, P. 1984. Species. Philosophy of Science 51:308–333.

Kitcher, P. 1987. Ghostly whispers: Mayr, Ghiselin, and the "philosophers" on the ontological status of species. *Biology and Philosophy* 2:184–192.

Macintyre, A. 1999. *Dependent Rational Animals: Why Human Beings Need the Virtues*. Chicago: Open Court Press.

Mayr, E. 1982. The Growth of Biological Thought. Cambridge, Mass.: Belknap Press.

Mayr, E., and P. D. Ashlock. 1991. *Principles of Systematic Zoology*, 2nd ed. Columbus: McGraw-Hill.

Mishler, B. D., and R. N. Brandon. 1987. Individuality, pluralism, and the phylogenetic species concept. *Biology and Philosophy* 2:397–414.

Post, J. F. 2006. Naturalism, reduction, and normativity: Pressing from below. *Philosophy and Phenomenological Research* 73:1–27.

Sandler, R. 2007. Character and Environment. New York: Columbia University Press.

Sandler, R. 2008. Natural goodness, natural value, and environmental virtue: Responses to Katie McShane and Allen Thompson. *Ethics, Place, and Environment* 11:226–235.

Sandler, R., and J. K. Crane. 2006. On the moral considerability of Homo sapiens and other species. *Environmental Values* 15:69–84.

Sober, E. 1984. Sets, species, and evolution: Comments on Philip Kitcher's "Species." *Philosophy of Science* 51:334–341.

Sokal, R., and T. Crovello. 1970. The biological species concept: A critical evaluation. *American Naturalist* 104:127–153.

Stamos, D. N. 2003. The Species Problem. Lanham, Md.: Lexington Books.

Thompson, A. 2008. Natural goodness and abandoning the economy of value: Ron Sandler's Character and Environment. *Ethics, Place, and Environment* 11:218–226.

Thompson, M. 1995. The representation of life. In *Virtues and Reasons*, ed. R. Hursthouse, G. Lawrence, and W. Quinn. Oxford: Oxford University Press.

van Valen, L. 1976. Ecological species, multispecies, and oaks. Taxon 25:233-239.

Wiley, E. O. 1978. The evolutionary species concept reconsidered. *Systematic Zoology* 27:17–26.

14 How to Think about the Free Will/Determinism Problem

Kadri Vihvelin

Common sense says that we have free will. We make choices. And while we are sometimes in a position where we make a choice while being mistaken about what our options are, this is not always, or even usually, the case. Ordinarily, when we make a choice we really do have the choice we think we have. You can stop reading this paper right now, or you can read on. You really can do either of these things. It's up to you. You have a choice.

But *for all we know*, determinism is true. Determinism is the thesis that the state of the entire universe at any time, together with the laws, is logically sufficient for the state of the universe at any later time.

Common sense either doesn't know, or doesn't take seriously, the thought of determinism. But as soon as a philosopher explains the thesis of determinism, common sense sees the problem: the truth of determinism means the absence of free will. As William James (1956, 150–151) put it:

Determinism professes that those parts of the universe already laid down absolutely appoint and decree what the other parts will be. . . . [N]ecessity on the one hand and impossibility on the other are the sole categories of the real. Possibilities that fail to get realized, are, for determinism, pure illusions; they never were possibilities at all.

We can sum up commonsense thinking with the following simple argument.

The No-Choice Argument for Incompatibilism

1. I have free will only if I sometimes have a choice about what to do.

2. I have a choice about what to do only if there is more than one thing that I am able to do.

3. If determinism is true, then what I do is (always) the only thing that I am able to do. (I am never able to do otherwise.)
4. Therefore, if determinism is true, I never have a choice about what to do.

5. Therefore, if determinism is true, I have no free will.

I think that this way of thinking about the free will/determinism problem is the correct way.¹ There is something as plain as the nose on our face—the fact that we often have a choice about what to do—and this fact *seems* to be in direct conflict with determinism. I see the free will/ determinism problem as the problem of saying whether this apparent conflict is real.

This may seem obvious—I hope it does seem obvious. But this is not how the literature sees the problem.

1 Through the Lens of Moral Responsibility (How Not to Think about Free Will)

In the contemporary literature, the free will/determinism problem is almost invariably viewed through the lens of moral responsibility. It is, of course, widely agreed that having free will is a *necessary* condition of being a morally responsible agent. But most contemporary discussions of the free will/determinism problem forge a much stronger link between questions about moral responsibility and questions about free will. Let's begin with some particularly striking examples:

the pessimists reply, all in a rush, that *just* punishment and *moral* condemnation implies moral guilt and guilt implies moral responsibility and moral responsibility implies freedom and freedom implies the falsity of determinism . . . [The optimist replies that] people often decide to do things, really intend to do what they do, know just what they're doing in doing it . . . But it is here that the lacunae in the optimistic story can be made to show. For the pessimist may be supposed to ask: But *why* does freedom in this sense justify blame, etc.? (P. F. Strawson 1962, 2–4)

if people did not regard themselves and one another as responsible beings, life would be unrecognizably different from what it actually is. But the concept of responsibility is a mysterious one which tends, on examination, to become increasingly opaque and to threaten variously to be incoherent or impossible or universally inapplicable. ... We can express the problem of responsibility in the form of the question "How, if at all, is responsibility possible?" And we can express the problem of free will in the form of the question: "What must our relation to our wills be," or better, perhaps, "What kinds of beings must *we* be if we are ever to be responsible for the results of our wills?" (Wolf 1990, 4) I believe that libertarians hold the views they do because they believe that any satisfactory conception of freedom of the will must allow us to justify our ascriptions of moral responsibility. They might be persuaded to accept a compatibilist account of freedom of the will if it could be shown that the fact that we are free, in that sense, could be used as the basis for a justification of the claim that we are morally responsible for our actions. (Bok 1998, 25)

Are we free agents? That depends on what you mean by 'free'. In this book the word 'free' will be used in what I call the ordinary, strong sense of the word. According to which to be a free agent is to be capable of being truly responsible for one's actions. (G. Strawson 1986, 1)

As I understand it, true moral responsibility is responsibility of such a kind that, if we have it, then it *makes sense*, at least, to suppose that it could be just to punish some of us with (eternal) torment in hell and reward others with (eternal) bliss in heaven. (G. Strawson 1994, 9–10)

These passages, from philosophers whose views about free will and the free will/determinism problem are otherwise very different,² have in common the assumption that if we are not morally responsible, then we are also never free agents; to be a free agent is, according to these philosophers, to have what it takes to be morally responsible for our actions (and thus for blame to be deserved, or just, or justified).

I think that this way of approaching the free will/determinism problem is a mistake. If we think about free will in this way, we are imposing a heavy burden on a class of natural facts—those facts, whatever they are, in virtue of which we have free will. We are saying that a class of natural facts constitute *free will* facts only if they play the role of being *the justifier* of praise, blame, and other practices associated with moral responsibility. Free will is assigned the burden of bridging the "is–ought" gap, of explaining why moral responsibility is *moral* responsibility.

If we think about free will in this way then we rule out, by stipulation, the possibility that so far as freedom is concerned, we have what it takes to be morally responsible, but we lack *something else* required for moral responsibility. For instance, insofar as we hold people responsible for what we believe were *their* past actions, it seems plausible that moral responsibility requires that we are numerically the same person through time. If that's right, then a philosophical argument for the conclusion that no one is ever numerically the same person over time is an argument, independent of free will, for the claim that no one is ever morally responsible (Parfit 1984). Another example: arguably our practice of holding other persons responsible includes the belief that we have the *authority* to demand certain kinds

of conduct from other people and to respond adversely—with sanctions, such as public acts of blame—when these demands are not met.³ If that's right, then a philosophical argument for the conclusion that we *never* have this kind of authority to make demands on other people is also an argument, independent of free will, for the claim that no one is ever morally responsible.

And if we think about free will this way, questions about what it is to have free will and to act freely (questions in the domain of the philosophy of action and metaphysics) take a back seat to questions about the nature of *just or justified* or *deserved* praise and blame (questions in the domain of moral philosophy). We start worrying about the nature of moral responsibility and whether anyone is ever *really* morally responsible—and moral responsibility becomes, as Wolf notes, a mysterious thing. It seems urgent, then, to provide an analysis or account of moral responsibile. And in discussing these questions the question of free will either gets left behind or becomes problematic in all the ways that moral responsibility is problematic. And these problems have little or nothing to do with *determinism*.⁴

The problem of moral responsibility—what it is and what it takes for someone to be morally responsible—is an interesting and important problem, but it is a *different problem* than the free will/determinism problem.

Not all philosophers who view the free will/determinism problem—and more generally, questions about free will—through the lens of moral responsibility see the connection in the strong way that P. F. Strawson, Wolf, Bok, and Galen Strawson see it. But the following kind of view is very common:

Granted, it's a mistake to conflate free will and moral responsibility, and granted, it is intelligible that someone—a child, for instance—might have free will without having what it takes to be morally responsible for her actions. But we care about freedom only because we care about moral responsibility, and our intuitions about what it is to have free will are clear only insofar as they are linked to moral responsibility. So we should understand the free will/determinism debate as a debate about whether determinism would deprive us of the kind of freedom that is required for moral responsibility.⁵

In my conception, compatibilism is the view that determinism is compatible with whatever sort of freedom is sufficient for moral responsibility, while incompatibilism is the view that determinism is not compatible with this sort of freedom. (Pereboom 1995, 42, n.2)

What is needed is an argument against the view that the freedoms that soft determinists have advocated are sufficient for moral responsibility. (Ibid., 43, n.7)

As a theory-neutral term of departure, free will can be defined as *the unique ability of persons to exercise control over their conduct in the fullest manner necessary for moral responsibility.* (McKenna 2009, sec. 1.1)

And let's reserve the technical term *morally free* to describe actions that are free in the sense required for moral responsibility. Then we can characterize the Free Will Thesis as the claim that some actions are morally free, and compatibilism as the claim that the Free Will Thesis is compatible with determinism.⁶ (Markosian 1999, 258)

This way of thinking is an improvement insofar as it does not impose upon an account of free will (free agency, 'moral freedom') the burden of being *the* justifier of blame, and insofar as it does not preclude the possibility that something *other* than lack of free will prevents us from being morally responsible.

But I think that this way of thinking about the free will/determinism problem is, nevertheless, mistaken: moral responsibility still gets all the attention. Since free will (free agency, 'moral freedom') is functionally defined as *whatever it is* that plays the role of satisfying the freedom requirement of moral responsibility, we cannot rely on commonsense assumptions about what counts as free will or free agency. If we want to defend the claim that we have the kind of freedom that plays this moral role, or that this kind of freedom is compatible with determinism, we will need to provide an *analysis* or *account* of this "moral freedom." And the only way of testing the adequacy of our account will be by using our intuitions about *when* agents are praiseworthy, blameworthy, or in some other way morally responsible for what they do.

Derk Pereboom exploits this fact with his notorious "Four Case" argument against compatibilism (Pereboom 1995, 2001). He claims that the compatibilist must provide an account, or at least sufficient condition, of what it is to satisfy the freedom component of moral responsibility, and he describes four cases and argues that at least one of these cases is a counterexample to each of the best compatibilist accounts, and a challenge to the possibility of *any* compatibilist account.⁷ Pereboom's claims about his cases have been criticized, but a more fundamental objection has been overlooked. If we reject the lens of moral responsibility when approaching questions of free will, a compatibilist can respond to Pereboom's argument by saying: "I don't know about moral responsibility. That is a different and difficult thing. My claim is only that *free will* is compatible with determinism. I don't have an analysis of free will, nor can I even provide a sufficient condition for having free will. But I know it when I see it, and nothing in your argument gives me any reason to believe that determinism would rob me of free will "

Defenders of this way of thinking about the free will/determinism problem would, I suspect, dismiss this objection on the grounds that they are interested only in *moral* freedom and *moral* responsibility. Some even go so far as to deny that we have *any* intuitions about free will independently of intuitions about praiseworthiness, blameworthiness, and moral responsibility.

But I don't think that this is right (and not just because I disagree with the claim about intuitions).⁸ There are different *kinds* of questions we can ask about free will, about moral responsibility, and about the relevance of determinism to free will and to moral responsibility.

Existential questions

1. Do we have free will?

2. Are we ever morally responsible for anything?

Compatibility questions

3. If determinism were true, would we have free will? (More generally, is determinism compatible with free will?)

4. If determinism were true, would we be morally responsible? (More generally, is determinism compatible with moral responsibility?)

Common sense believes that we have free will and believes that we are morally responsible for at least some of the things that we do. Common sense is not familiar with the philosophical term 'determinism', but once the thesis is explained to someone, that person can give answers to the two Compatibility questions. The most common *first response*, on learning what the thesis of determinism says, is to answer "no" to both questions that is, to say that if determinism were true, we would be neither free nor morally responsible.

But these four questions are different, and we should not define our terms in a way that precludes the possibility that their *answers* might be different. For instance, answers to the Compatibility questions *could be*: "No, if determinism were true we would not have free will—but we would still be morally responsible."

Or:

"Yes, if determinism were true we would still have free will—but we would not be morally responsible."

Alternatively, the answer to the free will question might be "don't know" and the answer to the moral responsibility question might be "no," or "yes." Or vice versa.

The problem with viewing free will strictly through the lens of moral responsibility—as Wolf, Bok, P. F. Strawson and Galen Strawson did

above—is that it conflates questions 1 and 2 and thus also 3 and 4, *forcing* the same answers to the two Existential questions as well as the same answers to the two Compatibility questions.⁹

Pereboom's, McKenna's, and Markosian's way of viewing free will through the lens of moral responsibility is more subtle, since it permits someone to answer "yes" to questions 1 and 3 about free will while answering "no" to questions 2 and 4 about moral responsibility. But since this view *defines* free will as the kind of freedom that is required for moral responsibility, the following combinations of answers are ruled out, *by stipulation*:

"Yes (1), we have free will and yes (2), we are morally responsible. But *no* (3), if determinism were true we would not have free will; and *yes* (4), if determinism were true we would still be morally responsible."

"Yes (1), we have free will and yes (2), we are morally responsible. *I don't know* (3) if we would still have free will if determinism were true; but *yes* (4), if determinism were true we would still be morally responsible."¹⁰

I do not defend either of these combinations of answers. I believe (1) that we have free will and (2) that we are morally responsible for some of our actions, and I believe (3) that determinism is compatible with free will and also (4) with moral responsibility. But other intelligible positions, such as those I have listed, should not be ruled out *in advance* by the terminology we use to discuss the free will/determinism problem.

One could object: Maybe you are right to complain that some philosophers don't take sufficient care to distinguish questions about free will from questions about moral responsibility—but surely you can't deny that free will, *as we ordinarily understand it*, is a necessary (though not sufficient) condition of moral responsibility? After all, if we don't have free will we are helpless as a falling rock and blame is *always* unjust. So we should impose the following constraint on accounts of, or claims about, free will: that they be accounts of, or claims about, the *kind of thing* that is a nontrivially necessary condition of moral responsibility.

Until recently, I would have agreed.¹¹ But I now think that we should not insist, *starting out*, that free will *must* be something that is necessary for moral responsibility. If we insist on this, we are immediately forced into taking a position concerning a philosophical literature that appears to be hopelessly stalemated—the literature concerning Harry Frankfurt's famous argument against the Principle of Alternate Possibilities (Frankfurt 1969).

Here's the problem. Suppose we have the commonsense view that to have free will is to be able to do otherwise, at least sometimes. And suppose

we also say that having free will is a necessary condition for being morally responsible. So far, so good. But Frankfurt and his defenders argue that someone may be morally responsible even if she is *never* able to do otherwise. If they are right, then either free will is not a necessary condition of moral responsibility or we may have free will even if we are *never* able to do otherwise.

Now, I think that Frankfurt (1969) and his followers are mistaken, and I have written several papers arguing this (Vihvelin 2000, 2004, 2008a). But I don't want to have this argument every time I write a paper on free will. So let me be explicit. In this paper, I will not assume that free will has *any* conceptual connection with moral responsibility—not even the apparently innocuous connection of being a necessary but not a sufficient condition of moral responsibility. Perhaps free will is necessary for moral responsibility; perhaps it is not. That is not my concern here. My concern here is *only* with the question of whether free will is compatible with determinism.¹²

I believe that viewing the free will/determinism problem through the lens of moral responsibility has distorted discussion of the problem. Our view of ourselves as free *and* responsible agents is so central to our lives that there is a danger of being lost in one big problem: "How, if at all, is moral responsibility possible?" (Or the only slightly smaller version of this problem: "How, if at all, is the freedom required for moral responsibility possible?") And when we think in these terms, our views tend to polarize in one of two directions: Either free will ("moral freedom") becomes as problematic as moral responsibility (Strawson 1986) or moral responsibility and free will ("moral freedom") become so unproblematic that it's hard to see what the fuss was about.¹³

I believe that if we take away the distorting lens of moral responsibility, we will be in a better position to solve the easier, or at least smaller, or, at least, nonmoral or premoral problem of whether determinism would rob us of the everyday kind of free will we unreflectively suppose we have almost all the time—the kind of free will that licenses talk of alternatives and options, choices, unrealized possibilities, missed opportunities, and wasted abilities. We will need to say more about what free will is—but I will start with determinism.

2 How Not to Think about Determinism

Roughly stated, determinism is the conjunction of two claims: (i) that we are *no exception* to the laws that govern the universe; and (ii) that those

laws state *sufficient*, as opposed to merely necessary, or probabilistic, conditions. More precisely, determinism is the thesis that a complete description of the intrinsic state of the world at any time t and a complete statement of the laws of nature together entail every truth about the world at every time later than t.

Let's call a possible world "deterministic" iff the thesis of determinism is true at that world; indeterministic iff the thesis of determinism is false at that world. There are different ways a world might be indeterministic. Most radically, a world might be indeterministic by being what I will call a "lawless world"—that is, by being a world where there are no laws at all. A less radical way in which a world might be indeterministic is by being what I'll call a "limited law" world—that is, by being a world where there are laws but some laws are limited in scope or application (cf. O'Connor 2000; Clarke 2003). The least radical possibility, and the one that corresponds most closely to what quantum physics appears to tell us about the actual world, is that a world is indeterministic by being what I will call a "probabilistic world." A probabilistic world is a world where there are laws, and the laws are all-encompassing rather than limited in scope or application, but at least some of the fundamental laws are probabilistic rather than deterministic.

Determinism is not an ontological thesis: it neither entails physicalism nor is entailed by it. There are possible worlds where determinism is true and physicalism false—for instance, worlds where minds are nonphysical things that nevertheless obey strict deterministic laws.¹⁴ And there are possible worlds (perhaps our own) where physicalism is true and determinism is false.

It is also important to distinguish determinism from several claims about causation:

UC Every event has a cause.

UEC Every event has an event-cause.

UEEC Every event has an event-cause, and no other kind of cause.

UC is the weakest claim. It says that every event has a cause, but it doesn't say anything about the entities that are the causes of events; it is consistent with the claim that the causes of events are (or include) events, facts, omissions, persisting objects or agents, or some other kind of entity.¹⁵ UEC says that every event has at least one event-cause, but it leaves open the possibility that events also have other kinds of causes. UEEC makes the strongest claim; it says that every event has a cause, and that the causes of events are *always and only* events. We should not assume that

determinism either entails or is entailed by even the weakest of these claims about causation.

UC, the thesis that every event has a cause, might be true at a world where determinism is false. Consider a world (perhaps our world) where the fundamental laws are probabilistic: at such a world, there are no sufficient causes. But it would be a mistake to conclude that there is no *causation*. It is now generally agreed by philosophers who work on the question that causation may be "chancy" or probabilistic.¹⁶ More controversially, UC might be true at a world that is neither deterministic *nor* probabilistic. If there can be causes without laws (if a particular event, object, or person can be a cause without instantiating a law) (Anscombe 1971), then it may be true—even at a lawless or limited-law world—that every event has a cause.

It is less clear still whether every deterministic world is a world where UC is true. Whether this is so depends on what the correct theory of causation is, and in particular, it depends on what the correct theory of causation says about the relation between causation and law. If there can be laws without causation, for instance, there may be deterministic worlds where UC is false. What is clear, however, is that we should *not* make the assumption, almost universally made in the older literature, that determinism is (or is equivalent to) the thesis that every event has a cause. This is an important point, because some of the older arguments in the literature assume that to deny determinism is to claim that there are uncaused events.¹⁷

Determinism should also not be confused with naturalism. There are different ways of formulating naturalism, but for our purposes we can define it as the conjunction of two claims which state that humans are no exception to the causal and nomological workings of the universe.

No Exception (Laws) We are no exception to the laws that govern the universe.

No Exception (Causation) When we cause things to happen, we do so in a way that differs in complexity, but not in kind, from the way that natural objects cause things to happen.

Note what these claims say, and what they do not say.

No Exception (Laws) says that the laws apply to us in the same way that they apply to other things, while *leaving it entirely open what those laws are*, and without saying anything about *how* laws "govern" or "apply" to things. That is, No Exception (Laws) is neutral with respect to the claim that the actual world is a deterministic, probabilistic, or limited-law world and it is also neutral with respect to rival philosophical accounts of the nature of laws. No Exception (Laws) merely says that *we* are not a special case, so far as the laws are concerned.¹⁸

No Exception (Causation) says that our causal interactions with the world are of the same kind as the causal interactions of other things with each other. This leaves it entirely open what these other causal interactions are like, and it leaves it entirely open what the correct theory of causation is. No Exception (Causation) merely states that *we* are not a special case, so far as causation is concerned.¹⁹

Though determinism is distinct from naturalism, there is a relation between them that is worth noting. In the literature, one sometimes hears that "determinism is the worst form of naturalism," and that a compatibilist is someone who not only believes we have free will (because she believes that naturalism is true and that free will is compatible with naturalism), but who also believes that we would have free will even if science told us that the "worst form of naturalism" is true—that is, if the fundamental laws turn out to be strict deterministic laws. And this seems right. A libertarian incompatibilist, by contrast, may be a naturalist because she believes that the fundamental laws are probabilistic or limited in the right kinds of ways. But her beliefs about free will are contingent on what science tells us: she draws the line at the "worst-case scenario." If science told us that determinism is true, the libertarian incompatibilist would give up her belief that we have free will.

The upshot of our investigation of determinism and related claims is this: since we do not know whether determinism is true or false, we cannot think of the free will/determinism problem as the problem of reconciling our belief in naturalism with our belief that we have free will. Even if we think that the true threat to free will is naturalism, and that determinism is merely the "worst case" of naturalism, we need to address the question of whether *determinism*, and not just naturalism, is compatible with free will. If we are incompatibilists, we need to argue that determinism would preclude us from having the free will we might otherwise have. And if we are compatibilists we need to defend, not only the claim that we *in fact* have free will, but also the claim that we would have free will *even if* determinism turned out to be true.

3 How to Think about the Free Will/Determinism Problem

I propose three constraints on what kind of problem the free will/determinism problem should be. 1. It should be a problem about something obvious and apparently undeniable like "there are tables and chairs" or "things continue to exist through time." Let's call this the "Moorean fact about free will."²⁰

2. Determinism should be prima facie incompatible with the Moorean fact about free will. (Everyone's first thought, on grasping what the thesis of determinism says, is "But if that were true, we wouldn't have free will.") 3. Indeterminism should not be prima facie incompatible with the Moorean fact. (It takes philosophical work to get people to see that indeterminism might also be problematic for free will.)

'Free will' defined so that it is conceptually linked to moral responsibility does not fit these constraints. At one time, the fact that we have the kind of freedom that justifies blame and punishment might have counted as a Moorean fact—but this is no longer true. These days it is easy to raise doubts about moral responsibility without saying a word about determinism.

'Free will' defined so that it is conceptually connected to something valuable or "worth wanting" also fails to fit these constraints. So does 'free will' defined in some vague (and vaguely Kantian) way like "that which makes us dignified and worthy of respect" or "that which makes us different from everything else in nature." You may wonder, then: what could free will possibly be, if it is to satisfy these constraints?

I gave my answer at the beginning of the essay: we should understand the free will/determinism problem as a problem about *choice*. Our core conception of ourselves as agents with free will consists in our belief that we are often in situations where we both *make and have a choice*.

You offer me cheesecake or mousse; I ask for mousse. The road forks and I must decide between a scenic (but winding) coastal route and the boring (but speedy) freeway; I take the freeway. The speeches are over and the vote has been called, "All in favor"; I raise my hand. We are in situations like these daily, and in such situations all of the following seem to be true:

1. We believe that we have options—that is, we believe that there are different things we are able to do, that we can select alternative courses of action. Call these options "A" and "B."

2. We deliberate (or ponder, ruminate, consult our feelings, etc.) between A and B.

3. We decide to do A. (We form an intention, make up our minds, make a choice.)

4. We do A.

5. The belief that we had before we made up our mind was a true belief. While we were deliberating, we really were able to do A *and* we really were able to do B.

We are, of course, sometimes mistaken about what our options are. We ponder the menu, trying to decide between mousse and cheesecake, unaware that the mousse is not available today. We ponder going out for a drive, having forgotten that our car is at the garage. In these situations, 1 to 4 may be true, while 5 is false. This shows that we are not entitled to infer, from the fact that we *deliberate* between A and B, that A and B are both *options* for us. We may be mistaken about what we are able to do. But it seems incredible to suppose that we are *always* mistaken.

Call the belief that we are often, or at least *sometimes* in situations where our beliefs about our options are correct, and thus 1 to 5 are all true: "Choice." I take Choice to be an uncontroversial part of our commonsense view of ourselves as agents with free will (and also the core part of that view). Choice seems as obviously and undeniably true as other facts that only a philosopher would question: the fact that there are tables and chairs, the fact that things continue to exist through time, the fact that I have a pair of hands. I call Choice a Moorean fact. Like the fact that there are tables and continuing things, the fact that we are sometimes able to do more than one thing is more obvious and self-evident than any *analysis* we can give of it. We do not need to give arguments or provide an analysis in order to defend our belief in a Moorean fact.

In saying this, I do not mean to suggest or imply any of the following: that we should never be convinced, by argument, to reject a Moorean fact;²¹ that we have a priori knowledge of Moorean facts; that Moore's response to the skeptic is correct;²² that Moorean facts are philosophically unproblematic. In particular, I don't mean to suggest that if something is a Moorean fact there is nothing more to be said about what kind of fact it is. (Compare the debate between perdurantists and endurantists about the two different ways in which it might be true that things continue to exist through time.) But I do claim that we need to be convinced *by very good argument* before we should give up our belief that a Moorean fact is indeed a fact.

It would be lonely if I were the only philosopher who saw things this way. Luckily I am not. I will close this section with a couple of quotes from Peter van Inwagen, who is largely responsible for how I think about the free will/determinism problem.²³

The problem of metaphysical freedom is a problem so abstract that it arises in any imaginable world in which there are beings who make choices. Consider some important choice that confronts you . . . Consider the two courses of action that confront you. I'll call them simply A and B. Do you really not believe that you are able to do A and able to do B? . . . It seems clear to me that when I am trying to decide which of two things to do, I commit myself, by the very act of attempting to decide between the two of them, to the thesis that I am able to do each of them. (van Inwagen 1998, 373)

The free-will thesis is the thesis that we are sometimes in the following position with respect to a contemplated future act: we simultaneously have both the following abilities: the ability to perform that act and the ability to refrain from performing that act. (van Inwagen 2008, 329)

4 How Not to Think about the Consequence Argument

We need to be forced by a very strong argument to give up our belief that a Moorean fact is a fact. Appeal to intuition is not an argument. We already know that taking the determinism thesis seriously causes people to doubt the Moorean fact of Choice. The philosophical problem is to decide whether these intuitions, this common way of responding to the problem, is justified.

There is only one *argument* in the literature for the conclusion that if determinism is true, the Moorean fact of Choice does not obtain.²⁴ It is the Consequence Argument, due, most famously, to van Inwagen (1983). Here is an informal statement of the argument: If determinism is true, then our acts are the *logical consequence* of the laws of nature and events in the remote past; but we don't have a choice about what happened before we were born, and we don't have a choice about what the laws of nature are; therefore, we don't have a choice about the *consequences* of those things, *including our own present acts*.

It can't be denied that this way of stating the argument tends to produce the desired effect. Our attention is directed to the apparently undeniable fact that (i) we have no choice about the remote past or the laws of physics, and (ii) at a deterministic world *everything*, including our own present acts, can be logically deduced from facts about the remote past and the laws.²⁵ Thinking in these terms causes us to think: "If determinism is true, then *everything has already been decided*, long before I was born. There is nothing that remains up to me. I *never have a choice about anything*, not even in those situations in which it seems clear as day that I am *able to do more than one thing.*"

The argument has convinced many. It has, I think, made incompatibilism respectable again. But it has also been widely misunderstood, and it has not received nearly the attention from compatibilists that it should have received. Here is a typical response, from Hilary Bok:

I have described two conceptions of possibility: the broad compatibilist conception of possibility and possibility tout court . . . van Inwagen's argument against compatibilism presupposes rather than establishes, the claim that the sense of possibility relevant to libertarian freedom is possibility tout court. (1998, 97, n.4)

By "two conceptions of possibility," Bok means two conceptions or senses of 'could have done otherwise'. She is accusing van Inwagen of begging the question by using 'could have done otherwise' in a way that entails that a person could have done otherwise only if determinism is false; that is, only if the person's doing otherwise is compossible with the actual past and the laws.

Here is van Inwagen's response to this kind of compatibilist response:

Many philosophers, in attempting to spell out the concept of free will, use the phrase 'could have done otherwise'. I did so myself in *An Essay on Free Will*. Nowadays, however, I very deliberately avoid this phrase. I avoid it because 'could have done otherwise' is ambiguous and (experience has shown) its ambiguity has caused much confusion in discussions of free will . . . A whole chapter of Daniel Dennett's first book on free will was written to no purpose because he didn't realize that 'could have done' sometimes means 'might have done' (and this 'might' is ambiguous; it has both an ontological and an epistemic sense) and sometimes 'was able to do'.

I want to make what seems to me to be an important point, a point that is, in fact, of central importance if one wishes to think clearly about the freedom of the will: compatibilists and incompatibilists mean the same thing by 'able'. And what do both compatibilists and incompatibilists mean by 'able'? Just this, what it means in English, what the word means. And, therefore, 'free will', 'incompatibilist free will', compatibilist free will' and 'libertarian free will' are four names for one and the same thing. If this thing is a property, they are four names for the property *is on some occasions able to do otherwise*. If this thing is a power or ability, they are four names for the power or ability to do otherwise than one in fact does. (2008, 333)

In these passages, van Inwagen makes two important points—that it is a strategic mistake to discuss questions of free will in terms of 'can' or 'could', and that it is a mistake to think that compatibilists and incompatibilists mean different things by 'able'. I agree with both points. It doesn't follow, however, that whenever we say, in ordinary English, that someone is able to do something, we always mean the same thing. Perhaps the locution is ambiguous between two or even three different meanings.²⁶ But since we

are anchoring our discussion to the Moorean fact of Choice, this should ensure that we are discussing the same thing.²⁷

The question before us, then, is whether determinism has the consequence that the ability to do otherwise that we take for granted whenever we make choices is an ersatz ability, an illusion, no ability at all. Could it be that we are *always* mistaken when we believe that we have a choice about what to do?

We need to take a closer look at a more precisely formulated version of the Consequence Argument. The version I will discuss is due to David Lewis in his classic "Are We Free to Break the Laws?" (1981), which van Inwagen commends as "the finest essay that has ever been written in defense of compatibilism—possibly the finest essay that has ever been written about any aspect of the free will problem" (2008, 330).

5 How to Think about the Consequence Argument

Think of the argument as a reductio. A compatibilist is someone who claims that the truth of determinism is compatible with the existence of the kinds of abilities that we assume we have in typical choice-situations. Let's call these "ordinary abilities." The Consequence Argument claims that if we suppose that a deterministic agent has *ordinary* abilities, we are forced to credit her with "*incredible* abilities" as well.

Here's the argument: Suppose, for reductio, that determinism is true and that someone, call her Dana, did not *in fact* raise her hand but had the *ordinary ability* to do so. If Dana had exercised her ordinary ability—if she had raised her hand—then either the remote past or the laws of physics would have been different (would *have to* have been different). But if that's so, then Dana has at least one of two incredible abilities—the ability to change the remote past or the ability to change the laws. But to suppose that Dana has either of these incredible abilities is absurd. So we must reject the claim that Dana had the ordinary ability to raise her hand.

The first thing to note about this argument is that it relies on a claim about counterfactuals. The argument says that one of these counterfactuals is true:

Different Past If Dana had raised her hand, the remote past would have been different.

Different Laws If Dana had raised her hand, the laws would have been different.

Both of these counterfactuals seem incredible to us because we are not used to thinking in terms of determinism, but we are good at evaluating counterfactuals, especially counterfactuals we entertain in contexts of choice. And when we contemplate Dana's options, we assume that:

Same Past and Laws If Dana had raised her hand, the past and the laws would still have been exactly the same.

So: both Different Past and Different Laws strike us as false. But that doesn't mean they *are* false—and if determinism is true, at least one of these counterfactuals *is* true.

Our best theory of counterfactuals—David Lewis's (1979) theory—tells us that our commonsense assumption about the past is correct and Different Past is false.²⁸ According to Lewis, if Dana had exercised her ordinary ability to raise her hand, the past would have been exactly the same as it actually was until a time shortly before she chose to raise her hand.²⁹ So Dana's ordinary ability to raise her hand does not entail an incredible ability to change the past.

Lewis's theory of counterfactuals does, however, have the consequence that Different Laws is true. But Different Laws *sounds* more incredible than it is. If we do not understand Lewis's theory of counterfactuals, we might think that Different Laws means:

New Set of Laws If Dana had raised her hand, our laws would have been replaced by a new and entirely different set of laws.

But this is not how we evaluate counterfactuals. We evaluate counterfactuals by considering the *closest* worlds where the antecedent is true, and the closest worlds are worlds where there are no gratuitous changes from actuality. To suppose a wholly new and different system of laws is to suppose a gratuitous change. Lewis's theory tells us that if Dana had exercised her ordinary ability to raise her hand, the past would have been exactly the same until the occurrence of a small and inconspicuous event-an extra neuron firing in Dana's brain—which marks the divergence from our actual history. Lewis calls this event a "divergence miracle" (or "law-breaking event"), since it is an event that is unlawful by the standards of our laws. But he tells us that it's a mistake to think that at these non-actual closest worlds one of our laws has been replaced by a contrary law; rather, we should think that one of our laws has been replaced by something that might be described either as an "almost-law" or, perhaps, as a version of the broken law, "complicated and weakened by a clause to permit the one exception" (Lewis 1973). So we should understand Different Laws as

Slightly Different Law If Dana had raised her hand, one of our laws would have been replaced by a *slightly* different law.

You may say: "This hardly helps! If Dana had raised her hand, a lawbreaking event—the divergence miracle—would have occurred. So the compatibilist remains committed to the claim that Dana has an incredibly ability—the ability to *break* the laws."

But—and this is Lewis's main point in his reply to van Inwagen—the ability to break a law is the ability to *cause a law-breaking event*. Dana has this ability only if she has the *ability to cause* the divergence miracle; but she doesn't have this ability. The divergence miracle is over and done with by the time that she acts.³⁰

There is more to be said, but I believe that Lewis's reply succeeds. (I also believe, though I cannot defend it here, that the success of his reply does not require the truth of any particular philosophical account of lawhood.) Thus the Consequence Argument fails. Compatibilists do not need to hide behind different senses or conceptions or kinds of ability, nor do they need to fear that the ability to do otherwise that a deterministic agent has is a "weaker" or "less robust" ability than the ability the libertarian claims we have. In the relevant sense, the *ordinary sense* that we use when we deliberate for the purpose of making a choice, deterministic agents are able to do otherwise.

6 How to Think about Free Will

I would like to close by saying a bit about the nature of the Moorean fact of Choice. Consider a typical choice situation: I am considering whether to vote "yes" by raising my hand. I think for a moment, and then keep my hand lowered. I refrained from voting "yes," but I was able to vote "yes" because I was able to raise my hand. I had the ordinary ability to do so. Nothing stopped me from doing so. Or so it seems.

What are the truth-makers of this fact—the fact that I was able to vote "yes" by raising my hand? A simple answer, once widely accepted, but now just as widely rejected, is:

Simple Conditional Analysis 'I was able to raise my hand' = if I had chosen (decided, intended, tried, wanted, etc.) to raise my hand, I would have raised my hand.

More generally, at one time the debate about the free will/determinism problem was regarded as a debate about whether the phrases 'could have done otherwise' or 'is able to do otherwise' can be *analyzed* in terms of some kind of counterfactual or subjunctive conditional. Compatibilists argued that 'can' is "hypothetical" or "conditional" or "constitutionally iffy"; incompatibilists argued that 'can' is "categorical" and not analyzable. There was a rich literature of cases, and while the discussion did not formulate the problem in terms of what I have been calling the Moorean fact, it was, in my view, an improvement over the current free-will literature in this respect: it did not view free will through the lens of moral responsibility. The debate was focused on the meaning of 'can' or 'could' or 'is able to' claims, rather than their moral or normative significance.³¹

It is an unfortunate fact about the history of the free-will literature that this debate came to an end sometime in the mid-1960s, shortly before the development of the Lewis/Stalnaker possible-worlds semantics for counter-factuals (Stalnaker 1968; Lewis 1973). This debate came to an end because *everyone* agreed that the attempt to provide a conditional analysis of 'could have done otherwise' was doomed in principle.³² Compatibilists had to look to other strategies for defending compatibilism, and the years since then have seen a wide variety of different strategies: the Strawsonian program (Wallace 1994); Frankfurt's (1969) argument against the Principle of Alternate Possibilities and the rise of Semi-Compatibilism;³³ accounts of free will in terms of one's "deep" or "real" self;³⁴ and, more generally, the increasing tendency to view the free will/determinism problem through the lens of moral responsibility.

But I digress. Here's an example of the kind of case that convinced everyone that the attempt to provide a conditional analysis of 'is able to do X' is flawed *in principle*.

Mary is under general anesthesia, which temporarily prevents her from thinking and thus prevents her from deciding or choosing to do anything. If she chose (decided, intended, etc.) to raise her arm, she *would have to be* conscious, and if she were conscious she would succeed in raising her arm (if she so chose). But since she is *in fact* unconscious, she is unable to raise her arm. So it is true that if Mary chose to raise her arm, she would, but false that Mary is able to raise her arm.

The moral drawn was that abilities are "categorical" rather than "iffy." Mary's unconscious state has changed certain categorical properties of her brain, the properties that enable her to make choices on the basis of reasoning and deliberation. It is because she lacks these categorical properties that she lacks the ability to *decide or choose whether to raise her arm*. And because she lacks the ability to decide or choose to raise her arm, she also lacks the ability *to raise her arm*.

I think that this way of diagnosing the case is a mistake. I agree that the Simple Conditional Analysis is false; and I agree that this case shows that *part* of the truth-maker of an ability-claim is something that is "categorical," or as we would perhaps say today, has a *causal basis* in the *intrinsic properties of the person* (Lewis 1997). It is because general anesthesia has changed the intrinsic properties of Mary's brain that she (temporarily) lacks the ability to make decisions and choices on the basis of deliberation. But I don't think that the ability to raise one's arm requires the ability to deliberate about the pros and cons of arm-raising: a very young child might have the first ability without having the second. I think that we should say that Mary has temporarily lost the second ability—the ability to *decide or choose* to raise her arm—while retaining the ability the child has. After all, general anesthesia is not a paralytic; it hasn't changed the intrinsic properties of the parts of Mary's brain and body that enable Mary to *move her arm*.

My aim here, however, is not to defend an analysis of ability but to say a bit more about the nature of the Moorean fact of Choice. The lesson to be learned from Mary's case is that *at least part* of the truth-maker of the Moorean fact is constituted by the intrinsically based abilities that a choicemaker has.³⁵ When I refrain from raising my hand, I am different from Mary in the following respect: I am conscious, and I exercise, and therefore possess, the ability to deliberate for the purpose of deciding whether or not to raise my hand. I also have the ability to keep my hand lowered, as well as the ability to raise it; I exercise the first ability; do not exercise the second.³⁶ We may sum this up, somewhat vaguely, by saying that I have "what it takes" both to *choose* to raise my hand and also to act on my choice by *raising my hand*.

But having the intrinsically based abilities to *choose* and to *do* both the acts you are contemplating doing does not suffice for the truth of the Moorean fact. I may have the intrinsically based ability to choose and to do some act that I am contemplating doing without being *able*, in the *sense relevant to Choice*, to do that thing. For suppose that I am contemplating whether to stay home or to go out for a walk. I have the intrinsically based ability to go for a walk: my legs aren't broken, I'm not paralyzed, nor am I disabled in any other way (I don't suffer from agoraphobia, for instance). I decide to stay inside and do so: I made a choice. But unbeknownst to me, I *didn't have the choice I thought I had*. For I am a prisoner—the doors are in fact all locked and armed guards are posted outside, ready to use whatever force it takes to prevent me from going for a walk.

In this situation, I have what we would ordinarily call the "ability" to go for a walk (the intrinsically based ability), but I lack what we would ordinarily call the "opportunity." I have what it takes to go for a walk, but unfortunately, I am (unwittingly) in a situation in which something extrinsic prevents me from exercising my ability. But this kind of situation is not the norm. Ordinarily, when we deliberate and make a choice between what we take to be two options (A and B) we possess the ability and *also* the opportunity to *choose* and also to *do both* of these things. Ordinarily, when I choose to stay home rather than go for a walk, I have the opportunity as well as the ability to go for a walk (and, of course, the opportunity as well as the ability to *decide* whether to stay or to go).

How should we understand this ability-opportunity distinction? I suggest the following: What *abilities* you have is determined by a subset of your *intrinsic properties*; ability facts are, very roughly, facts about what you are like "beneath your skin." What *opportunities* you have is determined by a subset of the facts about your *surroundings*. Abilities are shared by duplicates; you and your atom-for-atom duplicate have exactly the same abilities. Opportunities are a function of your situation; in swapping places with your atom-for-atom duplicate, you would also swap opportunities.³⁷

This is a partial, partly stipulative definition; the intention is to provide a mutually exclusive and jointly exhaustive classification of all the ways in which it might be true—of a particular person, on a particular occasion—that she *doesn't have a choice* about something. On my proposal, there are just two ways in which this might be true. The fault could be *in her* because she lacks, temporarily or permanently, something that would enable her to choose or to do that thing. If so, then I say that she lacks either the ability to choose or the ability to do that thing. Or the fault could be *in her surroundings*, that is, something *extrinsic to her* either prevents or would prevent her from choosing or doing that thing (perhaps by failing to supply her with what she needs to successfully choose or to do that thing. If so, then I say that she lacks the opportunity to choose or to do that thing.

The borderline between ability and opportunity may be unclear or disputed insofar as it is unclear, or in dispute, where the boundaries of the *person* (at the relevant time) are. I suggested a physical boundary ("under the skin"), but this is overly simple since implanted devices controlled by nefarious neurosurgeons are under a person's skin but are not part of *her*. But we can set these borderline and difficult cases aside so far as the free will/determinism problem is concerned. For that problem is to decide whether determinism really has the shocking consequence it appears to have: that no one *ever* has a choice about anything.

Here is an operational test (not a definition) of when someone is able to do something in the sense relevant to Choice. Suppose that I have the *ability to do* something—to go for a walk or get to the airport in time to catch a plane. And suppose, as we do ordinarily suppose, that I have the ability as well as the opportunity *to choose* whether or not to do that thing. (I'm not asleep, or too drunk to think, or pathologically depressed, or hopelessly distracted by the noise from my neighbors, etc.) If so, then I also have the opportunity to *do* that thing if it is true that *if I now tried to do it, I would succeed.* (Or, perhaps, would have a good, or good enough, chance of succeeding.) The cases in which I am locked in my room or wake up far too late to catch my flight are cases in which I have the ability but lack the opportunity; the difference lies in the fact that the counterfactual is false. (The Simple Conditional Analysis, though false, was on the right track.)

I propose that this is all that there is to the Moorean fact of Choice. The Moorean fact is the fact that we are often in situations in which we exercise our ability to choose *and* do A, while having the ability as well as the opportunity to choose *and* do B instead. Any argument for the claim that the truth of determinism is incompatible with the Moorean fact must, then, be an argument for one of two surprising claims: that determinism has the consequence that our abilities go out of existence whenever we don't exercise them or that determinism has the consequence that we lose our opportunities whenever we don't take advantage of them.

Acknowledgments

I am grateful for comments from Joe Campbell, John Carroll, Randolph Clarke, Janet Levin, Michael McKenna, Rebekah Rice, Terrance Tomkow, Gideon Yaffe, the audience at the INPC conference in Idaho in March 2008, and the audience at the University of Victoria Undergraduate Philosophy Conference in Victoria, British Columbia in March 2009.

Notes

1. I don't mean to suggest that I endorse the No Choice argument—I am a compatibilist. But I do think this argument succinctly sets out the *problem*.

2. P. F. Strawson, Wolf, and Bok are all compatibilists, but their ways of defending compatibilism are very different. Galen Strawson argues that "true moral responsibility" requires a kind of self-making that is impossible for finite, non-godlike creatures like us.

3. For a discussion of this last requirement on moral responsibility, as well as an interesting discussion of two senses or "faces" of responsibility, see Watson 2004.

4. Galen Strawson is quite clear about the problem being our nature as finite beings, rather than determinism. The compatibilists who think of free will this way—Wolf, Bok, and P.F. Strawson—do not see determinism (as opposed to naturalism) as the primary problem either.

5. This way of articulating the view is mine, but the view expressed is widespread in contemporary discussions of the free will/determinism problem.

6. Markosian argues that if agent-causation is possible and relevant to moral freedom and moral responsibility, then it is compatible with determinism.

7. Pereboom says that compatibilists must provide sufficient conditions for moral responsibility because "it is essential to their case that we can attribute moral responsibility in certain standard conditions" (2001, 101).

8. I do not care much for the recently overused term "intuition," but I will be arguing that our intuitions—or better, core commonsense beliefs—about free will are based on everyday cases of making and having a choice. The choice *may* be one for which we hold the person responsible, but it need not be.

9. P. F. Strawson, Wolf, and Bok answer "yes" to the two Existential questions and "yes" to the two Compatibility questions. Galen Strawson answers "no" to the two Existential questions, and "no" to the two Compatibility questions; however, on his view it makes no difference whether determinism is true—either way, we would not have free will or be morally responsible.

10. It might look as though Markosian (1999) can allow these combinations of answers because he explicitly introduces 'moral freedom' as a technical term and because he issues two disclaimers: he does not assume that moral responsibility and moral freedom require the ability to do otherwise, and he acknowledges that moral freedom "may be very different from 'freedom' in some other popular senses of the word." But Markosian in effect appropriates 'free will' as a term for moral freedom by defining the Free Will Thesis as the thesis that some actions are *morally free* and compatibilism as the thesis that the Free Will Thesis is compatible with determinism. Someone who wants to respond to the two Compatibility questions by answering "no" and "yes," respectively, is forced to say something like: "If determinism were true, there would be some sense of 'free' in which I would be less free than I believe I am, but I would still have the freedom that makes the Free Will Thesis true, so I would still be morally responsible" (1999, 2).

11. Thus I recently wrote: "the 'can' relevant to free will is the 'can' that we have in mind in contexts in which we raise questions about moral responsibility" (Vihvelin 2004, 428). I was closer to being right in an earlier paper: "I think that Simple Compatibilism is basically right, not just as an account of freedom of action, but also as an account of free will. It's not, however, a theory of moral responsibility and much confusion will be avoided once we realize this" (Vihvelin 1994, 141).

12. I would argue, though not here, that although the correct (compatibilist) solution to the free will/determinism problem leaves the moral responsibility/ determinism problem unsolved, since it is silent on the question of moral responsibility, it provides us with the following *conditional solution* to the problem: *If moral responsibility is possible for non-godlike creatures like us*, then it is compatible with determinism.

13. Arguably, this is what happens with the "reactive attitudes" account of moral responsibility defended, famously, by P. F. Strawson (1962). Free will drops out of the picture altogether. A hard determinist could accept his view.

14. Van Inwagen (1998) describes such a world, a world where there are only angels.

15. For a defense of fact-causation and omissions as causes, see Mellor 2004; for a defense of property-instances as causes, see Paul 2004; for a defense of states as causes, see Thomson 2003; for a defense of primitive agent-causation, see O'Connor 2000.

16. A world with chancy causes might be a world where some events lack causes altogether, but it need not be. The following seems to be a metaphysical possibility: a world where every event has a chancy cause.

17. This is the argument widely known as the "determined or random" dilemma.

18. This means that a libertarian or some other kind of incompatibilist—a hard determinist or someone agnostic about the truth or falsity of determinism—does not need to reject No Exception (Laws). Different libertarian accounts differ concerning the kind of laws that are required to accommodate free will. O'Connor (2000) has some interesting suggestions concerning a kind of limited laws account; Kane (1996) and Clarke (2003) explore varieties of probabilistic world accounts.

19. Note that a libertarian (or other incompatibilist) need not reject No Exception (Causation). This is most clear in the case of libertarians who explicitly endorse event-causation and the claim that free will is possible at probabilistic worlds (cf. Kane 1996), but even libertarians who defend agent-causation and insist that free will is possible only at limited-law worlds (cf. O'Connor 2000) can endorse No Exception (Causation), provided that they argue that agent-causation is a species of primitive object-causation.

20. David Lewis (1996) characterizes a Moorean fact as "one of those things that we know better than we know the premises of any philosophical argument to the contrary." I like this definition, but it may be stronger than I need. If you prefer a more epistemically neutral characterization, you may think of a Moorean fact as a commonsense platitude or belief that seems so obviously true that we typically don't bother saying it. I thank Joe Campbell for pressing me on this point.

21. Here I disagree with some of the stronger claims about the evidential status of Moorean facts made by recent defenders of common sense against the skeptic; for good discussions, see Kelly 2004 and Lycan 2001.

22. Nor am I making *any* claim about knowledge. The free will/determinism problem isn't a problem about whether we *know* that we have free will or whether we *know* that free will is compatible with determinism; it's a problem about whether we *have* free will and whether our having free will is compatible with determinism. (It's a problem within metaphysics, not epistemology.) If the epistemic skeptic is right, then we don't know that we have hands, let alone that we have free will. My intent in calling Choice a Moorean fact is only to point out that our belief that we have free will is as firmly embedded within common sense as our belief that we have hands, that the world didn't begin to exist five minutes ago, that there are tables and chairs and other minds. It has whatever epistemic status these other beliefs have.

23. I offer these quotes as evidence that van Inwagen sees the free will/determinism problem in the way that I do—as a problem about Choice. But in a recent paper, van Inwagen (2008) describes a *larger* problem that he calls "the problem of free will," but which I think is more accurately described as "the free will*–moral responsibility* problem," because it is the problem of "finding out" which of three "seemingly unanswerable" arguments is "fallacious," and the conjunction of the conclusions of these three arguments is that *moral responsibility* does not exist. The problem that van Inwagen sets out is an important problem. But for the reasons I give in my section on the "lens of moral responsibility," I believe that we should not think of it as "the problem of *free will.*"

24. I mean that there is only one *serious* argument. There are, of course, lots of bad arguments; in particular, there are lots of arguments based on various fatalist confusions. For criticism of some of these arguments, see Vihvelin 2008b.

25. I would argue, however, that these apparently undeniable facts are *not* Moorean facts.

26. Some plausible candidates for the different things we might mean by 'S is able to do X': "S has the skills, competence, or know-how required to do X"; "S has the skills required to do X *and also* the physical capacity to use these skills (she has what it takes to do X)"; "S has what it takes to do X *and* nothing stands in her way" (she's in the right place at the right time; she's got the means and the opportunity; there are no impediments or obstacles, etc).

27. This is an oversimplification, but it will do for now. See my final section for more details.

28. For a good discussion of Lewis's theory of counterfactuals, see Bennett 2003.

29. Why "shortly before"? Because if Dana were to suddenly raise her hand, out of the blue, without choosing, intending, or even wanting to do so, she would not be exercising any ability; her hand-raising would be something that "just happened" to her. What Lewis has in mind is something like the following: Dana thinks of something she didn't actually think of, and this thought gives her a reason (or something she takes to be a reason) to raise her hand.

30. For more details on why Lewis's reply does not commit compatibilists to the claim that free deterministic agents have incredible abilities, see Vihvelin 1991, 2008b, and 2011.

31. For more details about why I think the search for a conditional analysis of ability was abandoned prematurely, see Vihvelin 2004.

32. For an overview of the debate, see Berofsky 2002.

33. Frankfurt 1969 spawned an enormous literature: a good introduction is Widerker and McKenna 2003; for semi-compatibilism, see Fischer 1994 and Fischer and Ravizza 1998.

34. This literature is huge. The classic starting points for deep-self views are Frankfurt 1971 and Watson 1975; Wolf (1990) criticizes deep-self views and Watson (2004) responds to Wolf's criticism.

35. What are these intrinsically based abilities? An incompatibilist might argue that they are "agent-causal powers," and some compatibilists would agree (cf. Markosian 1999). In Vihvelin 2004, I argue that they are dispositions or bundles of dispositions, differing in complexity but not in kind from intrinsically based dispositions like fragility.

36. Arguably these are the same ability. I do not assume this here because there are other cases in which the two abilities are clearly different—for instance, cases where we deliberate between going for a walk and going for a bike ride.

37. I am assuming, of course, that the laws are held constant.

References

Anscombe, G. E. M. 1971. *Causality and Determination: An Inaugural Lecture*. Cambridge: Cambridge University Press.

Bennett, J. 2003. A Philosophical Guide to Conditionals. Oxford: Clarendon Press.

Berofsky, B. 2002. Ifs, cans, and free will: The issues. In *The Oxford Handbook on Free Will*, ed. R. Kane. Oxford: Oxford University Press.

Bok, H. 1998. Freedom and Responsibility. Princeton: Princeton University Press.

Clarke, R. 2003. Libertarian Accounts of Free Will. Oxford: Oxford University Press.

Collins, J., N. Hall, and L. A. Paul, eds. 2004. *Causation and Counterfactuals*. Cambridge, Mass.: MIT Press.

Fischer, J. M. 1994. The Metaphysics of Free Will. Oxford: Blackwell.

Fischer, J. M., and M. Ravizza. 1998. *Responsibility and Control*. Cambridge: Cambridge University Press.

Frankfurt, H. 1969. Alternate possibilities and moral responsibility. *Journal of Philosophy* 66:820–839.

Frankfurt, H. 1971. Freedom of the will and the concept of a person. *Journal of Philosophy* 68:5–20.

James, W. 1956. The dilemma of determinism. In *The Will to Believe and Other Essays*. New York: Dover. Originally published in 1884, in James's *An Address to the Harvard Divinity Students*.

Kane, R. 1996. The Significance of Free Will. New York: Oxford University Press.

Kelly, T. 2004. Moorean facts and belief revision, or Can the skeptic win? *Philosophical Perspectives* 19:179–209.

Lewis, D. 1973. Counterfactuals. Cambridge, Mass.: Harvard University Press.

Lewis, D. 1979. Counterfactual dependence and time's arrow. Noûs 13:455-476.

Lewis, D. 1981. Are we free to break the laws? Theoria 47:113-121.

Lewis, D. 1996. Elusive knowledge. Australasian Journal of Philosophy 74:549-567.

Lewis, D. 1997. Finkish dispositions. Philosophical Quarterly 47:143–158.

Lycan, W. 2001. Moore against the new skeptics. Philosophical Studies 103:35-53.

Markosian, N. 1999. A compatibilist version of the theory of agent causation. *Pacific Philosophical Quarterly* 80:257–277.

McKenna, M. 2009. Compatibilism. In *The Stanford Encyclopedia of Philosophy*, ed. E.N. Zalta, http://plato.stanford.edu/archives/win2009/entries/compatibilism/.

Mellor, D. H. 2004. For facts as causes and effects. In *Causation and Counterfactuals*, ed. J. Collins, N. Hall, and L. A. Paul. Cambridge, Mass.: MIT Press.

O'Connor, T. 2000. The Metaphysics of Free Will. Oxford: Oxford University Press.

Parfit, D. 1984. Reasons and Persons. Oxford: Clarendon Press.

Paul, L. A. 2004. Aspect causation. In *Causation and Counterfactuals*, ed. J. Collins, N. Hall, and L. A. Paul. Cambridge, Mass.: MIT Press.

Pereboom, D. 1995. Determinism al dente. Noûs 29:21-45.

Pereboom, D. 2001. Living without Free Will. New York: Cambridge University Press.

Stalnaker, R. 1968. A theory of conditionals. In *Studies in Logical Theory*, ed. N. Rescher. Oxford: Blackwell.

Strawson, G. 1986. Freedom and Belief. Oxford: Oxford University Press.

Strawson, G. 1994. The impossibility of moral responsibility. *Philosophical Studies* 75:5–24.

Strawson, P. F. 1962. Freedom and resentment. *Proceedings of the British Academy* 48:1–25. Reprinted in *Freedom and Resentment and Other Essays*. New York: Methuen & Co. References are to the reprint.

Thomson, J. J. 2003. Causation: Omissions. *Philosophy and Phenomenological Research* 66:81–103.

van Inwagen, P. 1983. An Essay on Free Will. Oxford: Clarendon Press.

van Inwagen, P. 1998. The mystery of metaphysical freedom. In *Metaphysics: The Big Questions*, ed. P. van Inwagen and D. Zimmerman. Oxford: Blackwell.

van Inwagen, P. 2008. How to think about the problem of free will. *Journal of Ethics* 12:337–341.

Vihvelin, K. 1991. Freedom, causation, and counterfactuals. *Philosophical Studies* 64:161–184.

Vihvelin, K. 1994. Stop me before I kill again. Philosophical Studies 75:115–148.

Vihvelin, K. 2000. Freedom, foreknowledge, and the principle of alternate possibilities. *Canadian Journal of Philosophy* 30:1–24.

Vihvelin, K. 2004. Free will demystified: A dispositional account. *Philosophical Topics* 32:427–450.

Vihvelin, K. 2008a. Foreknowledge, Frankfurt, and ability to do otherwise: A Reply to Fischer. *Canadian Journal of Philosophy* 38:343–372.

Vihvelin, K. 2008b. Compatibilism, incompatibilism, and impossibilism. In *Contemporary Debates in Metaphysics*, ed. T. Sider, J. Hawthorne, and D. Zimmerman. Oxford: Blackwell.

Vihvelin, K. 2011. Arguments for incompatibilism. In *The Stanford Encyclopedia* of *Philosophy*, ed. E.N. Zalta, http://plato.stanford.edu/archives/fall2008/entries/ incompatibilism-arguments/>.

Wallace, R. J. 1994. *Responsibility and the Moral Sentiments*. Cambridge, Mass.: Harvard University Press.

Watson, G. 1975. Free agency. Journal of Philosophy 72:205-220.

Watson, G. 2004. Two faces of responsibility. In *Agency and Answerability*. Oxford: Clarendon Press.

Widerker, D., and M. McKenna, eds. 2003. *Moral Responsibility and Alternative Possibilities*. Vermont: Ashgate.

Wolf, S. 1990. Freedom within Reason. Oxford: Oxford University Press.

Contributors

Alexander Bird University of Bristol Andrea Borghini College of the Holy Cross Judith K. Crane Southern Illinois University, Edwardsville Michael Devitt The City University of New York Bruce Glymour Kansas State University Peter Godfrey-Smith The City University of New York Marc Lange University of North Carolina, Chapel Hill Noa Latham University of Calgary Bence Nanay University of Antwerp and University of Cambridge Jason G. Rheins University of North Carolina, Chapel Hill Ronald Sandler Northeastern University Matthew H. Slater Bucknell University Roy Sorensen Washington University in St. Louis Achille C. Varzi Columbia University Kadri Vihvelin University of Southern California Neil E. Williams State University of New York at Buffalo

Index

Abelard, Peter, 132, 149-150 Absence, 19, 21, 113-127, 212. See also Voids Accident, 7, 9-10, 15, 27, 45, 55-62, 65, 67, 69-70, 80, 86-87, 89, 93, 96, 105, 207, 219, 243 Adams, Ernest, 135, 150 Adaptation, 16, 156, 254, 259-263, 275-277, 282, 292, 302 Agency, 25, 314-315, 317, 320-321, 324-325, 328, 330, 335-336, 338-339 Albon, S. D., 249 Anscombe, G. E. M., 322, 338 Antibiotics, 204-205 Antirealism, 256, 262, 283 about species, 22, 155, 159, 163, 168 Ariew, André, xi, 181-182, 192, 197 Aristotle, 7-12, 25, 119, 124, 126, 132-133, 149-150, 257, 262, 269, 287, 290, 299-306, 310 Armstrong, David, 26, 28, 56, 66, 73, 96, 100, 104, 109-111, 170, 184, 192, 284 Arnett, Frank C., 223, 229 Arnold, Stevan J., 246, 250 Arthritis, rheumatoid, 23, 122, 199-230 Ashlock, Peter D., 291, 311 Atomists, 116, 119, 124 Atran, Scott, 42, 51 Averroes, 124

Aviles, Jesus M., 245, 249 Ayres, Ian, 235, 249 Balshine, S., 250 Barker, Meg J., 197 Barrow, John D., 75, 82 Bealer, George, 14, 28 Beatty, John H., 231, 250 Bennett, Jonathan, 79, 82 Berkeley, George, 145, 284, 287 Berofsky, Bernard, 338 Berzi, Vittorio, 80, 82 Bigelow, John, 14, 28, 109-111 Biological conception of species. See Species concept, biological Biological kinds. See Kinds, biological Biological science(s), 18, 25, 140, 142-146, 156-159, 162-163, 166-168, 205, 224, 226 Bird, Alexander, 14, 20, 26, 28, 71, 82, 85-96, 110-111, 282 Birds, 1, 126, 138, 143, 166, 241, 245, 271, 293, 303 Bjornstad, Ottar N., 251 Black, Robert, 111 Blame, 314-319, 324 Bloch, Daniel A., 229 Bok, Hilary, 315-316, 327, 334-335 Bolton, Martha Brandt, 284, 287 Bolzano, Bernard, 134, 150 Boorse, Christopher, 226, 229

Borghini, Andrea, 228, 230 Born, M., 80, 82 Bosque, Carlos, 245, 249 Bosque, Maria Teresa, 245, 249 Bouchard, Frédéric, 231, 249 Boundaries, 11, 21-22, 118, ch. 7 passim, 170, 207-208, 217, 222, 225, 266, 271-272, 275, 291, 296, 300, 333 Boutin, Stan, 137, 251 Boyd, Richard, 17-19, 23, 28, 49-50, 159, 169-171, 175, 191-192, 201, 204, 207, 222, 224-226, 229, 259, 285, 287 Brandon, Robert N., 231, 249, 292, 300-301, 311 Brentano, Franz, 132-136, 150 Brigandt, Ingo, 197 Butler, Mark J. IV, 250 Calmette, Joseph, 131, 150 Campbell, Keith, 177, 192-193 Campbell, Scott, 47, 50 Caplan, Arthur L., 229 Carnap, Rudolph, 50, 145 Carter, Brandon, 75, 82 "Carving nature at its joints," 1-2, 19, 98, 116, 142, 152, 146-147, 294, 298 Plato's Phaedrus, 1-2, 17, 116, 142, 152 Taoist allegory of, 1-2 Casati, Roberto, 124, 127, 149 Catchpole, E., 249 Categoricals, 290, 299-306, 310 Causal relationships, 39-40, 185, 221, 236-237, 262, 266, 278 Causation, 121, 184-185, 236, 241, 321-323, 335-336 agent, 335-336 interventionist conception of, 236 singular, 184–185 Çelik, Can, 230 Cells, 141, 164, 210-213, 218, 221-227

Chaine, Alexis S., 242-244, 251 Chalmers, David, 31, 109, 111 Chan, Kung-Sik, 251 Chargaff, Erwin, 45-46, 48, 50 Choice, 103, 122, 137, 140, 147, 241, ch. 14 passim Circularity, 59, 61, 87, 285 Clarke, Randolph, 321, 334, 336, 338 Clutton-Brock, Tim, 249 Collier, John, 111 Common sense, 126, 130, 134, 136, 142, 148, 158, 313, 318, 336-337 Compatibilism, 316-317, 327-328, 331, 334-335, 338 Concepts, 13, 18, 21, 87, 104, 131, 137, 139-140, 170 phenomenal, 103 species (see Species concept) Conditional analysis, 110, 330-331, 334, 338 Conditionals, 20, 26, 36, 38, 40, 42-44, 57-60, 64-68, 71-72, 80-81, ch. 4 passim, 99, 102–103, 105–110, 192 Consequence argument, the, 326-330 Conservation of energy, 58, 62-63, 73, 75 Contingency, 21, 65, 79, 87, ch. 5 passim, 144, 201, 263 Contingentism. See Laws, contingentist view of Conventions, 18, 21, 126, ch. 7 passim, 148, 228 Copeland, B. Jack, 149-150 Copp, David, 302, 309-310 Cortens, Andrew, 145, 150 Coulomb's law. See Laws, electrostatic Coulson, Timothy N., 240, 249 Counterfactual dependence, 38, 44 Counterfactuals, 20, 26, 36, 38, 42-44, 57-60, 64-68, 71-72, 80-81, ch. 4 passim, 330. See also Subjunctive facts backtracking, 20, 91-93 material, 99, 102, 105, 107-108

subjunctive (see Subjunctive facts) supervenience, 103 Counterlegals, 20, 71-72, 77 Cracraft, Joel, 155, 165, 170-172, 292, 310 Crane, Judith K., 24-25, 301, 311 Crawley, Michael J., 249 Crovello, Theodore J., 156, 173, 292, 311 Cummins, Robert, 188, 193 Curzon, Lord of Kedleston, 130-132, 138, 150 D'Agostino, Ralph B., 235, 249 Daly, Chris, 187, 193 D'Amico, Robert, 122, 127, 222, 229-230 Darwin, Charles, 146, 150, 169, 171, 180-181, 193 Davidson, Donald, 184-185, 193 da Vinci, Leonardo, 133, 150 Dawkins, Richard, 193 De Queiroz, Kevin, 169, 171 Descartes, Rene, 124 Determinism, 25, ch. 14 passim Development, 120, 164, 188-189, 261, 264, 269, 290 Devitt, Michael, 16, 22, 27-28, 157-158, 161, 163, 169-171, 190, 193 De Vries Robbé, Peter, 223, 229 Dewey, John, 42, 50 Diagnosis, 202-203, 224 Dirac, Paul, 119 Disease, 18, 23, 122, ch. 10 passim Disorders, 123, 213-214, 228 Dispositions. See Properties, dispositional DNA, 16, 33, 39-40, 45-46, 50, 144, 187, 210 Donohue, John J. III, 16, 30, 235, 249 Dretske, Fred, 26, 28, 109, 111, 188, 193

Dubois, Alain, 287 Dummett, Michael, 26, 148, 150 Dupré, John, 16-18, 22, 28, 157, 160, 162, 169, 171-172, 179, 186, 193, 224, 229, 270, 273, 287, 292, 310 Dynamic structural equation model. See Models, dynamic structural equation Eco, Umberto, 143, 151 Ecological conception of species. See Species concept, ecological Edworthy, Steven, 229 Ehrenfest, Paul, 62 Ehrenhaft, Felix, 44 Ehrich, Dorothee, 251 Ehring, Douglas, 184, 193 Einstein, Albert, 63, 66, 80, 82 Elder, Crawford L., 179, 193 Eldredge, Niles, 155, 165, 172 Electromagnetism, 63, 66, 71-72, 75, 113 Elgin, Catherine, 141, 151 Ellis, Brian, 14, 26, 28, 71, 81-82, 109-111, 179, 193, 215, 225, 227, 229 Engelhardt, H. Tristam, Jr., 229 Environmental factors/pressures, 209-210, 226, 297 Ereshefsky, Marc, 22, 28, 155-157, 160-170, 172, 176, 183, 185-187, 191, 193, 292, 295-297, 310 Essences, 7-17, 22-23, 26, 53, 121, 170, 175-179, 182, 186, 191, 199-209, 217-219, 223, 228, 256-257, 259, 262, 266-268, 280, 286 historical, 17, 176 of individuals, 7-9, 12-13 of kinds, 9-16, 22-23, 26, 53, 177-179, 182, 186, 191, 199-209, 217-219, 223, 228, 256-257, 259, 267-268, 280

Essences (cont.) nominal, 7, 11, 14, 16, 23, 121, 170, 199, 206, 257, 262, 266, 286 process (see Kinds, process) real, 7-11, 16, 23, 170, 199, 262, 266, 284-286 Essentialism, 7, 12-17, 19-23, 27, 66, 71, 146, 157, ch. 9 passim, 176-180, 183, 192, 199-209, 222-224, 227, 259, 266, 292 anti-, 7, 176-180, 183, 192, 292 biological, 16-17, 22-23, 27, 146, 157, 176, 179-182, 205, 259, 266, 292 scientific, 12-15, 19-21, 66, 71 Euclid, 133, 151 Events, 3, 18, 20-22, 63, 92, 94-95, 131, 136, 140, 184-185, 244, 247, 260-262, 266, 277, 279-280, 321-322, 326, 329-330, 336 Evolutionary conception of species. See Species concept, evolutionary Evolutionary theory, 23, 163, 179, 187, 231-232, 235, 242, 245-247, 259, 272, 279 dynamic vs. statistical interpretations of, 23, 231, 246 Ewens, Warren, 233, 249 Explanation, 2, 7, 9-10, 16, 27, 56, 74, 77, 120, 159, 162-168, 183, 185-187, 265, 190, 192, 204, 244-247, 289, 297-298 biological, 27, 162-168, 183, 186-187, 190, 192, 244-247, 265, 297-298 historical, 163-166 structural, 16, 27, 163-165, 167-168 Facts, 27, 55, 58, 85-86, 93-96, 325-326, 335-336 contingent/accidental, 27, 79, 86-87, 102 Moorean, xi, 181-182, 192, 197 nomic (see Laws) subjunctive (see Subjunctive facts) sub-nomic, 58, 85-86, 93-96

Falck, Wilhelm, 251 Febvre, Lucien, 139, 151 Ferraris, Maurizio, 146, 151 Feynman, Richard, 63, 82-83 Fiat entities, 137-138, 145 Figures, 11, 125-126, 135-136 Fischer, John Martin, 338, 340 Fitness, 23-24, 231-233, 240-242, 245-247, 263, 273, 276, 284 FitzGerald, G. F., 67, 80-81 Fitzpatrick, Ray, 230 Flier, Frank, 223, 229 Fodor, Jerry, 55-56, 74, 83 Foot, Philippa, 24-25, 289-291, 299-302, 307, 310 Forms, 1, 8-9, 142, 256-257, 283 Foster, John, 79, 83 Four case argument against compatibilism, 317 Frankfurt, Harry, 319-320, 331, 338-339 Franklin, Benjamin, 118 Franklin, James, 44, 47, 50 Franklin, Laura, 200, 229 Free will, 25-26, ch. 14 passim Frege, Gottlob, 27, 137, 150-151 Frequencies, 33, 189, 231-234, 242-243, 246, 284 genic, 231-234, 242-243, 284 trait, 189, 246 Gale, Richard, 118, 127 Gallagher, Dympna, 235, 250 Galton, Antony, 133, 149, 151 Geach, Peter, 49-50, 289, 310 Gendler, Tamar S., 175, 194 Genesis, 142 Genotypes, 188, 270, 272, 274 Ghez, Richard, 234, 250

Ghiselin, Michael T., 17, 29, 169, 172, 179, 191, 194, 278–279, 282, 287, 292, 310–311
Gigerenzer, Gerd, 42, 50
Glenn, Sigrid S., 194 Glymour, Bruce, 23-24 Godfrey-Smith, Peter, 19, 194 Goheen, Jacob R., 250 Gold, 10-11, 57-60, 85-86, 176, 183-187, 204, 206, 216, 218, 257, 259, 268 Goldman, Alvin, 46, 51 Goodman, Nelson, 5-6, 19, 22, 29, 33-38, 41-43, 51, 55-57, 83, 120, 140, 145, 147, 149, 151, 224, 229, 286-287 Goodness, 19, 24-25, 289-291, 298-302, 306-310, 320, 334 moral, 19, 289-290, 306, 308, 320, 334 natural, 24-25, 289-291, 298-302, 306-310 Goodwin, N. B., 250 Gorini, Vittorio, 80, 82 Gornick, Janet G., 235, 250 Grenfell, Bryan T., 249 Griffiths, Paul E., 17-18, 29, 155-157, 161, 165-167, 169, 172-173, 179-182, 191-192, 194, 224, 229, 258, 260, 288 Grow, Gerald, 119-120, 126 Grue. See Induction, new riddle of Grundy, Scott, 249 Guyer, Paul, 284, 287 Hacking, Ian, ix-x, 2, 4, 6, 17-18, 26, 29, 179, 194, 202, 230, 285-286, 288 Half-life, 54-55, 67, 71, 89, 91 Harman, Gilbert, 51 Harris, Tamara, 250 Hawthorne, John, 175, 194 Hawthorne effect, 34, 49 Health, 220, 222-223, 227, 306-308. See also Disease Heaviside's equation, 66-67, 80-81 Heil, John, 109-111 Heller, Mark, 147-149, 151 Hempel, Carl, 3, 5, 26, 29, 74, 77, 83, 120

Hendry, Andrew P., 250 Hershel, William, 113 Heterogeneity argument, the, 159 Heymsfield, Steven B., 250 Hoeffler, Anke E., 235, 250 Holes, 119, 122-124, 127 Homeostasis, 210-212, 217, 226-227 Homeostatic mechanisms. See Mechanisms, homeostatic Homeostatic property cluster theory. See Kinds, homeostatic property cluster (HPC) Hull, David L., 17, 29, 169, 172, 176, 178-179, 191, 194, 278-280, 282, 288, 292, 310 Hume, David, 12, 48, 51, 83 Humeanism, 97, 99, 102, 105, 107-109, 111, 225 Hunt, Bruce J., 80, 83 Hursthouse, Rosalind, 289, 307-309, 311 Husserl, Edmund, 26, 144, 151 Huygens, Christian, 124 Idealism, 145 Ideas, 11, 181, 284 Identity, 13, 15, 19, 21, 104, 107, 132, 137, 140, 147, 255 necessity of, 13, 21, 104, 107 psychophysical, 13, 104 Immoderate realism. See Platonism, immoderate/extreme realism Incompatibilism, 313, 316 Individuality thesis, the. See Species, as individuals

Individuals, 9, 11, 13, 21, 23–24, 147, 175, 177, 181–182, 188, 234–236, 239, 245, 248–249, 257, 290, 303, 306–307 Induction, 5–6, 17, 19, ch. 2 *passim*, 56, 120, 201, 204, 217, 225 counter, 37, 49 habits of, 41–42

new riddle of, 5, 34-37, 120

projection, 6, 19, 56

Infection, 200, 204, 217, 221 Inheritance, 115, 117, 188-189, 218, 231 Jackendoff, Ray, 135, 151 Jackson, Frank, 35-38, 42, 51 Jakob, Robert, 230 Jakobsen, Kjetill S., 251 James, William, 313, 339 Jefferson, Thomas, 132, 139, 146 Kane, Robert, 336, 336-339 Kaplan, Jonathan, 18, 30, 248, 250 Karban, Richard, 250 Keating, Paul, 125 Kelly, Kevin, 51 Kelly, Thomas, 336, 339 Kendell, Robert E., 228, 230 Kim, Jaegwon, 184 Kinds, 2-7, 10, 13-19, 21-27, 34-35, 39, 41, 47, 53-57, 69, 71, 78-79, 85, 88, 116-117, 119-126, 132, 141-143, 155, 158-159, 161-169, 175-180, 183-184, 186-187, 189, 191-192, 199-210, 214-220, 222-228, 253-254, 256, 258-259, 267-268, 270, 275, 281, 284, 291, 298 biological, 16-17, 22, 176, 179-180, 183, 186-187, 189 causal-explanatory natural, 17, 22, 158-159, 166-168 essence of (see Essences, of kinds) homeostatic property cluster (HPC), 17-19, 23, 169, 192, 201-202, 205, 207-210, 217, 219, 224-226, 259 natural, 2-7, 10, 13-19, 21-27, 34-35, 39, 41, 47, 53–57, 69, 71, 78–79, 85, 88, 116-117, 119-126, 141-143, 155, 158-159, 161-169, 175-178, 183-184, 186-187, 191-192, 199-210, 214-220, 222-228, 253-254, 256, 258-259, 267-268, 270, 275, 281, 284, 291, 298

nominal (see Essences, Nominal) para-natural, ch. 6 passim process, 18, 141, 201, 210, 214-222, 226-227 realism about, 22, 25, 204-205, 208-209, 226, 256, 262, 274-276, 281 reference to, 13, 27, 121-122, 132 Kingsland, Sharon E., 233, 250 Kinnison, Michael T., 245, 250 Kitcher, Philip, 16, 18, 22, 27, 29, 31, 155-157, 159-164, 166-168, 172, 186-187, 194, 231, 251, 260, 273, 286-288, 292, 295, 297, 311 Koch, Robert, 199-200 Kolm, N., 245, 250 Kornblith, Hilary, 27, 29, 170, 172 Krebs, Charles J., 251 Kripke, Saul, 6, 12-16, 29, 101, 103, 110, 121-122, 175, 179-180, 191, 194, 199, 215, 230 Kuhn, Thomas, 172

Ladyman, James, 110-111 Lande, Russell, 248, 252 Lange, Marc, 4, 19-21, 26-27, 29-30, ch. 4 passim, 214, 228-229, 232 Langman, Rodney E., 196 LaPorte, Joseph, 4, 17, 27, 30 Latham, Noa, 21 Laws, 4, 18, 20-21, 26, chs. 3-5 passim, 244, 323-325, 330 best system account of, 26, 56, 66, 68 biological, 4, 18, 244 contingentist view of, 21, 97, 101-103, 107-108 fundamental, 20-21, 55-57, 62-65, 69-73, 75, 80, 89-92, 96, ch. 5 passim, 323-325, 330 necessitarian view of, 21, 97, 100-103, 107 - 109special science, 98 supervenience, 102-103, 105-108

Lenski, Richard E., 244, 252 Lesson, René-Primevère, 143 Levins, Richard, 265, 290 Lévy-Leblond, Jean-Marc, 64, 83 Lewens, Tim, 188, 193, 196-197, 199, 233, 253 Lewis, David K., 26, 30, 35, 48, 51, 56, 66, 68, 73, 79, 83, 91, 96, 109, 111, 123-124, 127, 134, 144, 146-147, 150-151, 330-334, 338-341 Lewontin, Richard, 16, 30, 187-189, 193-194, 199, 265, 290 Lidgjaerde, Ole Christian, 253 Lierse, Caroline, 14, 28, 109-111, 217, 229, 231 Lineage, 156, 159, 171, 193, 246, 262-266, 268, 272, 284, 294 Linnaean hierarchy, 22, 146, 152, 155, 164, 167-169, 172, 312 Locke, John, 7, 10-13, 142, 152, 170, 232, 264, 286, 289, 290 Lorentz, Hendrik, 63-64, 82-83 Lotka-Voltera model. See Models, Lotka-Voltera Lycan, William G., 338, 341 Lyon, Bruce E., 242-244, 251 Mach, Ernst, 119, 127 Macintyre, Alasdair, 291, 313 Mackie, John L., 217, 224-225, 230, 232 Mallet, James, 165, 173 Manchester, Keith L., 46, 50-51 Mandelbrot, Benoît, 139, 152 Markosian, Ned, 319, 321, 337, 340-341 Martin, C. B., 98-99, 109-111 Martin, Thomas E., 247, 252 Mass, 15, 54-55, 62, 64, 69, 71-74, 76-80, 82, 89, 97-101, 103-106, 109-110, 118, 243, 252, 270 Matthen, Mohan, 175, 179, 188, 193, 196-198

Maxwell, James Clerk, 66, 83 Mayr, Ernst, 1, 23, 30, 156, 157, 163, 169, 173, 180-183, 191-192, 197, 260, 290, 293, 299, 313 McCartney, James J., 231 McKenna, Michael, 319, 321, 340-342 Mechanisms, 17, 39, 41, 46-47, 50, 117, 163, 205, 210 causal, 39, 210, 261, 269 homeostatic, 17, 261, 269 (see also Kinds, homeostatic property cluster [HPC]) Medical sciences, 122, 202, 205-207, 215-216, 225-226 Medin, D., 42, 51 Mellor, David H., 27, 30, 112, 171, 184, 195, 197, 338, 341 Mereology, 145 Metaphor, 1-2, 17, 116, 141, 143-145, 300 Metaphysics, 2-3, 6-10, 21-22, 97, 101, 119, 126, 144-148, 181, 202, 210, 269, 286, 311, 318, 339 prescriptive, 147 Meyers, Marcia K., 252 Mill, John Stuart, 65, 226, 232 Millikan, Robert, 44-45, 48, 50-51 Millikan, Ruth G., 49, 51, 179, 188, 197 Mills, Susan K., 233, 252 Mishler, Brent D., 16, 30, 167-168, 170, 173, 261, 289, 290, 294, 302-303, 313 Mitchell, Sandra, 4, 30 Models, 5, 9-10, 15, 19, 23-24, 34, 37-38, 44, 47-49, 55, 90, 119, 133, 135, 144, 189-190, 207, ch. 11 passim, 311 auto-regression, 236 demographic, 235, 242 dynamic structural equation, 236 individual-level, 193, 238, 240-241, 244, 248-249, 251
Models (cont.) instantial (of confirmation), 5 Lotka-Voltera, 181, 235 population-level, 23-24, 193, 235-251 reductive, 236-237, 240-241, 244, 248-251 Moderate realism. See Platonism, moderate realism Monism, 31, 147, 157, 293-294, 296, 306 Monmonier, Mark, 133, 152 Morality, 19, 28, 291-292, 308, 310-313, 316-322, 326, 333, 336-339 and freedom/responsibility, 316-322, 326, 333, 336-339 Morgan, B. J. T., 251 Morrison, Benjamin, 133, 152 Mumford, Stephen, 109-111 Mutation, 188, 210-211, 214 Nagel, Ernest, 3, 30 Nahin, Paul J., 80, 83 Nanay, Bence, 22-23, 178, 184-185, 188, 190, 193, 197-198 Natural goodness, 24-25, 291-293, 300-302, 304, 306, 308-313 Naturalism, 291, 312-313, 324-325, 337 Natural kinds. See Kinds, natural Natural laws. See Laws Naturalness, 4, 19, 34-39, 41, 56, 74, 78-79, 159, 161, 165, 203-205, 219, 226-227, 270 Nature's joints. See "Carving nature at its joints" Navarro, Carlos, 251 Neander, Karen, 188-191, 198-199 Necessitarians. See Laws, necessitarian view of Necessity, 10, 12, 14, 20-21, 58, 61-65, 69-73, 80, 85, 88-97, 240, 245, 282 a posteriori, 101, 110

grades/levels of, 21, 61-65, 69-73, 80, 85, 88-96, 97 logical, 58, 65, 245 metaphysical, 89, 97 natural, 14, 20, 57, 61-65, 69-7, 102, 105, 107-109 (see also Laws) Newman, Stanton, 202, 232 New riddle of induction. See Induction, new riddle of Newton, Isaac, 62, 76-77, 83, 89, 100, 110 Nominalism, 23, 141, 158, 171, 178, 180, 182-183, 186, 191-192, 194, 258, 268, 285 explanatory trope, 183, 186, 194 trope, 178, 180, 182-183, 186, 191-192, 194 Norton, John, 49, 51 O'Connor, Timothy, 321, 336, 339 O'Donoghue, Mark, 251 O'Hara, Robert J., 156, 173 Okasha, Samir, 16-17, 30, 156-157, 173, 175, 179-180, 186, 191-192, 195 Oliver, Alex, 112, 171, 192-193, 196

Oppenheim, Paul, 74, 77, 83

Para-natural kinds. See Kinds,

Palmer, Todd M., 238, 242, 249-250

Para-reflections, 113-114, 126-127

Particles, 15, 17-20, 53-57, 64, 69-75,

77-79, 81, 89-91, 96, 98, 101, 105,

Ownership, 133-136

Pal, Palash B., 80, 83

Parfit, Derek, 315, 339

Pargetter, Robert, 36, 51

Parsimony, 105, 120, 183

Parmenides, 8, 124, 257, 283

Paul, L. A., 127, 336, 338-339

Peberton, Josephine, 249

para-natural

136

Pascal, Blaise, 124

Peirce, Charles S., 133, 152 Penrose, Roger, 63, 83 Pereboom, Derk, 316-317, 319, 335, 339 Perez-Contreras, Tomas, 249 Phenotypes, 24, 156, 193, 232, 240, 242, 246, 258-259, 263, 265, 293 Phenylketonuria (PKU), 212-213, 227 Philbrick, Nathaniel, 124, 127 Pierson, Richard N., 250 Pigliucci, Massimo, 18, 30, 248, 250 Place, U. T., 109-111 Plato, 1-2, 8-9, 17, 21, 116, 142, 152, 285.290 Phaedrus, 1, 116, 142, 152 Platonism, 258-260, 286 immoderate/extreme realism, 24, 258, 260-264, 266, 268, 272 moderate realism, 24, 258-259, 264, 266, 270, 272, 286 Platts, Mark, 186, 196 Pluralism, 17, 147, 155-157, 159, 161-163, 166, 168, 272-273, 291-298, 304, 309 Pollock, J., 79, 84 Pope John XXI, 124 Popper, Karl, 8, 30 Population genetics, 231, 233, 243, 246, 249 Population thinking, 23, 180-183, 192, 195 Populations, 23-24, 37-39, 47, 156, 170, 180-183, 192, 195, ch. 11 passim, 254, 258-261, 263, 278, 291-294, 296, 298-299, 303 Post, John F., 306, 310 Powers. See Properties, dispositional Praise, 315-318 Predicates, 3-6, 8-9, 12, 14, 19, 34-36, 38, 41, 47, 55, 57, 115, 131, 161, 170, 201, 290 non-projectible, 38, 55, 131 (see also Induction, new riddle of) projectible, 5-6, 19, 57, 115

Prevention, 91, 120, 201, 203, 212, 220, 224, 317, 331-333 Principle of alternative possibilities, 319, 331 Pringle, Robert M., 250 Probability, 24-25, 40, 42, 47, 79, 82, 236, 244-245, 321-333, 336 Problem of moral responsibility, the, 314 Processes. See Kinds, process Projectability, 56-57, 115, 117, 204, 224 Properties, 7, 10, 12-17, 20-21, 26-27, 45, 54-57, 69-73, 77, 79, 81, 90, 97-100, 102-110, 117, 131, 141, 144, 170, 175-180, 188, 206-208, 212, 218-220, 223, 231, 259, 267, 270, 272, 285, 293, 295-297, 333, 306 accidental, 7, 10, 15, 207 categorical, 97-98, 104, 206, 208, 219-220, 331 derivative, 20, 54-57, 69-73, 79, 81, 90 disjunctive, 55, 73, 77, 104, 106, 131, 170 dispositional, 97, 99-100, 102-108 essential, 7, 10, 12-14, 16, 21, 26, 175-177, 179-180, 206-207 (see also Essences) fundamental, 15, 20, 53-57, 69-73, 77, 79, 81-82, 85, 89-90, 92, 97-103, 105-109, 271 genetic, 16-17, 27, 45, 141, 144, 188, 212, 218, 223, 259, 267, 270, 272, 293, 295-297, 306 (see also DNA) homeostatic clusters of (see Kinds, homeostatic property cluster [HPC]) intrinsic, 17, 97-100, 102-110, 117, 175, 179-180, 206, 208, 270, 332-333 natural, 35, 39, 41, 55-57, 69, 73-75, 77-79, 141-142 phenomenal, 98, 103-105

Properties (cont.) relational, 99-102, 176, 179-180, 206, 208, 285, 333 role (see Properties, dispositional) surface, 206-208, 218-219 types, 23, 177-179, 182-192, 217 unnatural, 20, 56, 73, 79 (see also Properties, disjunctive) Propositions, 85, 87-90, 94, 97, 290 Purugganan, Michael D., 251 Pust, Joel, 188, 196 Putnam, Hilary, 6, 12-15, 22, 27, 30-31, 121-122, 127, 145, 147, 152, 175, 179-180, 191, 196, 199, 218, 222-223, 230, 257, 288 Putnam tests, 121-122 Qualities, 7, 11, 104, 117. See also Properties primary, 11, 117 secondary, 11, 117 Questions, 4, 6-7, 14, 26, 33, 35, 44, 46-48, 92, 102-103, 149-156-157, 175, 184, 226, 240, 298, 314, 316-320, 335-336 compatibility, 318-319, 335 existential, 318-319, 335 Quidditism, 105, 109, 111 Quine, Willard V. O., 2, 5-7, 15, 19, 29, 31, 35, 51, 117, 127, 147, 152, 158, 286, 288 Quinn, Thomas P., 250 Rand, Ayn, 283, 288 Raven paradox, 5-6, 26, 33, 117, 120 Rawls, John, 34, 50 Realism, 19, 21-22, 24-25, 100, 102, 141-145, 155, 160-165, 168-170, 204-205, 208-209, 226, 256-262, 264-270, 273-276, 281, 284-285, 291 about boundaries (see Boundaries) immoderate (see Platonism, immoderate/extreme realism)

metaphysical, 143, 155, 158, 168, 285 moderate (see Platonism, moderate realism) moral, 19, 28 about natural kinds (see Kinds, realism about) ontic structural, 100, 102 pluralistic, 22, 155, 160, 162-163, 167-168, 273, 168, 272-273, 291-298, 304, 309 scientific, 143, 295 about species (see Species, realism about) Reductive strategy, 234-235, 238-239, 242, 247 Reference, 13, 15, 27, 103-104, 110, 121-122, 148, 159, 226, 276, 280 Reflections, 21, 113, 126, 256 Reichenbach, Hans, 35, 42, 48, 51-52, 57 Relativity theory, 63-64, 66 Reliability, 11, 34, 37-38, 41-42, 44, 46-47, 49, 208, 217, 228, 233, 238-239, 241-244, 247, 272 Reproductive success, 24, 232, 242-247, 276, 293 Revenson, Tracey A., 230 Reynolds, J. D., 150 Reznek, Lawrie, 222, 226, 230 Rheins, Jason G., 24 Ritter, William, 113 Robb, David, 185, 196 Robertson, Teresa, 175, 196 Robinson, David A., 251 Rodd, Helen F., 250 Rorty, Richard, 145, 149, 152 Rosenberg, Alex, 2, 30-31, 196, 231, 249-250 Ross, Don, 111 Ross, Katherin E., 250 Ross, Patrick, 250 Rueness, Eli Knispel, 251

SAARDs, 204 Sampling, 19, 38-39, 41, 44, 46-47, 79, 225 Sandler, Ronald, 24-25, 301, 307, 309, 311 Sanford, David H., 177, 196 Sawyer, Sarah, 116, 127 Schaffer, Jonathan, 111, 177, 196 Schall, Ira, 42 Scheffler, Israel, 152 Schroeter, Laura, 42 Scientific essentialism. See Essentialism, scientific Scurvy, 212-213, 227 Searle, G. F. C., 67, 80-81, 84 Searle, John, 137, 146, 152 Selection, 19, 23-24, 37, 44, 146, 188-190, 192, 194, 231, 233, 239-241, 244-247, 263, 276, 293 Sepulveda, Dennis, 250 Shadows, 21, 113, 116-118, 121-126, 131 Shimizu, Kentaro K., 240, 251 Shoemaker, Sydney, 56, 66, 71, 73, 79, 84, 109, 112 Sidelle, Alan, 148-149, 152 Similarity, 2-3, 6, 15, 24, 26, 43-44, 46, 156, 160, 201, 204, 223, ch. 12 passim Simons, Peter, 136, 152, 177, 196 Slater, Matthew H., 31 Smith, Barry, 113, 127, 132, 149-150, 152, 196, 226 Snell's law, 113 Sobel, David, 302, 309-310 Sober, Elliott, 169-170, 173, 179, 181-182, 188-189, 196-197, 231, 250, 292, 311 Sokal, Robert, 142, 152, 156, 173, 292, 311 Soler, Juan J., 249 Sorenson, Roy, 21, 113, 117, 124, 127 Sortals, 12, 26, 131-132

Species, 1-2, 4, 16-17, 24, 138, 146, 253-157, 159-170, 176, 179-182, 192, 205, 208-209, 224, 242-246, 253-255, 258-267, 269-281, 284-287, 289-307, 309-310 as individuals, 17, 169, 253-255, 266, 277-280 as kinds, 208, 224, 254, 267 monism about, 155, 157, 272, 291-292, 294, 298, 304 pluralism about, 24, 157, 291, 294-295 realism about, 155, 157-158, 160-163, 254, 267, 291, 294-295 Species category problem, the, 156-157 Species concept, 24-25, 155-162, 169, 181, chs. 12-13 passim axiological, 25, 299-306, 309-310 biological/interbreding, 24, 156-157, 181, 253-254, 258-261, 263-264, 266, 279, 284, 286, 291-293, 298, 303 ecological, 24, 156, 160, 253-254, 260-263, 266, 275-277, 279, 292-297, 303 evolutionary, 24, 156, 253-254, 260-261, 266-267, 275, 277-279, 292-294 phenetic, 156, 292, 295 phylogenetic/cladistic, 156-157, 160, 254, 267, 281, 292-294, 298-299, 302 Species problem, the, 2, 22, 24, 156, 291, 294 Species taxon problem, the, 156-157 Spurrett, David, 111 Stability, 16-17, 41, 54, 57-77, 80-81, 86-90, 92-96, 116, 201, 259, 297 sub-nomic, 59-70, 80, 86-88, 92-96 Stalker, Douglas, 29, 35, 52 Stalnaker, Robert, 79, 84, 331, 339 Stamos, David, 142, 152, 291-292, 311 Stanford, P. Kyle, 27, 31, 157, 169-170, 173 Stanton, Maureen L., 250

States of affairs, 18, 108, 184, 228 Stenseth, Nils Christian, 234, 240, 251 Stephens, Christopher, 231, 251 Sterelny, Kim, 27-28, 155-158, 161, 165-167, 169, 171, 173, 231, 251, 258, 260, 288 Stereotype, 122, 199, 206, 218, 257. See also Essences, nominal Stoljar, Daniel, 109, 112 Stove, David C., 47, 50, 52 Strawson, Galen, 315-316, 318, 320, 334-335, 339 Strawson, P. F., 31, 314, 316, 318, 334, 336, 340 Strevens, Michael, 48, 52 Structure, 2, 6, 11, 15-16, 26, 45-46, 140-141, 146-147, 159-161, 164, 176, 184–187, 200, 204, 206–209, 239, 241, 244-247, 254, 270, 295 atomic, 11, 15, 176, 257 causal, 41, 159, 200, 204, 208, 239, 241, 244-247, 295 chemical, 6, 15, 45-46 genetic (see Properties, genetic) intrinsic, 39, 138, 140-141, 223 Suarez, Francisco, 132, 152 Subjunctive facts, 20, 26, 57, 69, 71-79, 86, 90, 330 Substances, 3, 6, 8-9, 11, 21, 115, 116, 120, 123, 132, 147, 201, 206-208, 214-215, 225, 228 Sullivan, Lisa M., 249 Sulmasy, Daniel, 215, 228, 230 Supervenience, 97, 102-109, 215 Surfaces, 22, 113, 132-138 Survival, 24, 122, 132, 137-140, 189-190, 232, 242-247, 263, 269-272, 276, 304, 307 Swartz, Norman, 79, 84 Swoyer, Chris, 109, 112 Symptoms, 23, 199-201, 203-204, 212-214, 216-221, 227-229 Syndrome, 200, 228

Taxa, 22-23, 141-142, 157, 159-160, 162, 164-169, 270-274, 270-274, 285, 291-292, 294, 296-297, 302, 305 higher, 155, 165, 270-274, 285, 302 species, 23, 157, 159-160, 162, 164-167, 291-292, 294, 296-297, 305 Taxonomy, 2, 7, 113, 141-143, 146, 156, 163-169, 254, 273-274, 281, 287, 295, 298, 302 Theory of evolution. See Evolutionary theory Thompson, Allen, 308, 311 Thompson, Michael, 290, 299-300, 311 Tilley, Stephen G., 242, 251 Tipler, Frank J., 75, 82 Todd, Peter M., 42, 50 Tooley, Michael, 26, 31, 109, 112 Traits, 24-25, 156, 176, 188-190, 241, 244-245, 273, 286, 293, 308-309 genotypic, 24, 246, 276, 293 phenotypic, 24, 156, 241, 246, 293 Travisano, Michael, 242, 250 Tropes, 177-179, 181-187, 191 Truth, 10, 58-62, 64-67, 70-71, 74-78, 80, 85-89, 91, 93, 95, 100-102, 243, 280, 290, 313, 328 contingent (see Contingency) necessary (see Necessity) Tuberculosis, 199-200, 217, 222-223 Typological thinking, 181-182, 186, 278 Universals. See Properties Unwin, Martin J., 250 Üstün, Tevfik, 228, 230

Vacuum, 119, 124 van Cleve, James, 144, 153 van Fraassen, Bas, 26, 31, 187, 197 van Inwagen, Peter, 79, 84, 325-328, 330, 336-337, 340 Van Valen, Leigh, 156, 173, 292, 296, 311 Varzi, Achille, 21-22, 113, 124, 127, 149, 152-153, 196 Vihvelin, Kadri, 25, 320, 335, 337-338, 340 Viljugrein, Hildegunn, 251 Visser, Marjolein, 250 Voids, 115-116, 124, 127 von Pettenkofer, Max Joseph, 74 Wallace, John, 26-27, 31 Wallace, R. Jay, 331, 340 Walsh, Denis, 16, 31, 179-180, 182, 188, 191, 197, 231, 251 Watson, Gary, 334, 338, 340 Watson and Crick, 45-46 Widerker, David, 338, 340 Wigner, Eugene, 63, 78, 84 Wiley, E. O., 292, 311 Wilkerson, Terence E., 179, 197 Wilkes, Charles, 123-124 Wilkins, Maurice, 46 Williams, D. C., 50, 52, 177, 197 Williams, Neil, 17, 23, 226-228, 230 Wilson, Jack, 140, 153 Wilson, Peter, 249 Wilson, Robert A., 16, 27, 31, 157, 169, 173, 186, 197, 224, 230 Wimsatt, William C., 197 Wolf, Susan, 314, 316, 318, 334-335, 338, 340 Wolterstorff, Nicholas, 197 Woodward, James, 4, 31 World Health Organization (WHO), 218, 230 Young, Truman P., 250 Zimmerman, Dean, 149, 153, 340