THE

NORMATIVE

AND THE

NATURAL

MICHAEL P. WOLF AND

JEREMY RANDEL KOONS

The Normative and the Natural

Michael P. Wolf • Jeremy Randel Koons

The Normative and the Natural

palgrave
macmillan

Michael P. Wolf
Washington, PA, USA

Jeremy Randel Koons
Doha, Qatar

for Jocelyn
and
for Lucy

Acknowledgments

Much of our work on this book over the last four years has been done at a great distance with the help of various forms of technology. This is the way of things when the person with whom you share the strongest sense of scholarly purpose lives 11,000 km and several time zones away. Reaching a point where we could write such a work was never a solitary affair, however, and we have benefited from the guidance and collegiality of numerous philosophers, scientists, and other scholars over the years. Among those who have been wonderful colleagues who offered one or both of us support and insight in recent years are David Beisecker, Ron Bayline, Donald Bruckner, Helen Daly, Willem DeVries, Geoff Georgi, Matthias Kiesselbach, Hanna Kim, Lorin Lachs, Mark Norris Lance, Steven Levine, Chauncey Maher, Peter Olen, Michael Oliver, Gregg Osborne, Adam Podlaskowski, Michael Robinson, and Carl Sachs. The advice of Carole Sargent was invaluable in coaxing grown-up, professional behavior out of the two of us, as was the guidance offered by Matthew McAdam. Emily Golling also provided valuable editorial assistance in preparing the manuscript.

In preparing this manuscript, we benefited enormously from the support of Georgetown University and its Center for International and Regional Studies (CIRS) in Doha, Qatar. CIRS hosted a one-day conference on a very drafty draft of our manuscript on March 8, 2015. This conference brought participants to Doha and distributed the work of

commenting on our work in progress. We received enormously helpful feedback from Bana Bashour, Ray Brassier, Erhan Demircioğlu, Anjana Jacob, Daniele Mezzadri, Niklas Möller, Jim O'Shea, John Ryder, Matthew Silverstein, Lucas Thorpe, and Jack Woods. Jim O'Shea and several others from our earlier list also offered helpful comments at our session at the Summer Institute in American Philosophy in Dublin on June 8, 2015.

Michael thanks Washington and Jefferson College for its support of his research with a sabbatical during the fall semester of 2014. He would also like to thank a number of people who made it possible for him to contribute to this book during what have turned out to be several challenging years. Family members—Anne, Dan, Flynn, Liz Holtz, Sharon Stewart, and George Halleck—have been enduring source of consolation, and restoration, and his gratitude for them is immeasurable. Jocelyn Waite has been a source of love, encouragement, and inspiration, and he is grateful to have such a partner in his life. For much of the last 15 years, Debbie Mix has kept him writing, both by example and by vocal support, while Lauryn Mayer has kept him afloat in ways too numerous to recount. Jim Behuniak, Matthew Burstein, Jack Thorpe, and Margaret Weitekamp are damn fine people. Michael would run into a burning building for any of them, not that he would have any particularly brilliant ideas what to do when he got there.

Jeremy would like to thank the Georgetown University School of Foreign Service in Qatar for making his contribution to this book possible with a sabbatical during the fall semester of 2014 and also for a series of Faculty Research Grants which enabled him to travel to Washington, Pennsylvania, several times for intensive writing sessions. But most of all, Jeremy would like to thank Lucy for all of the ways she supported both the writing of this book and the process of getting it published. Whenever Jeremy turned to her for help, as this project took shape, Lucy always rolled up her sleeves, and for this, Jeremy will forever be in her debt.

Contents

List of Figures

Introduction

We live in a world of "oughts": you ought not lie, you ought to believe based on evidence, you ought to buy low and sell high. We are the sort of creatures we are because the world and our comportment to it matter to us. Whether we do well, think correctly, and do what is right are not things we can simply report with disinterest; they are affairs in which we feel the pull to engage and take responsibility for how we carry on. Our lives unfold within social spaces that are composed of lattices of norm-governed practices. We come to have stories of ourselves as individuals as we take up stances toward those practices, whether by endorsing them or resisting them. We embrace our roles as teachers, as artists, as parents (to name just a few); we take pride in work done well; we rail against injustice; we keep our promises (to name just a few). We thus find ourselves engaged in normative discourse, in which we overtly declare, question, command, and so on, how things are to be done and what is to be permitted, valued, or disvalued.

But we also inhabit a natural world, one (seemingly) devoid of our institutions, values, and norms. In our encounter with the various phenomena of the natural world, we typically seek to explain them, and our most comprehensive and explanatorily powerful accounts of this world come from the natural sciences. We find ourselves embodied as biological organisms, composed of physical parts bound together by fundamental forces. By many estimates, the scope and power of these explanatory

approaches are so great as to render suspect anything that cannot be reconciled with them. We may broadly characterize such approaches as forms of naturalism. To be a naturalist in this sense is not to subscribe to a particular set of theories, but rather to a particular conception of our overarching intellectual project. We seek the most complete description of the world possible through scientific inquiry, and philosophy thus becomes an ancillary—perhaps even disposable—form of inquiry, serving science's needs where possible, and adopting its results.

It has long been thought difficult to fit these "oughts" into a scientific world-view. For instance, moral wrongness is not something that science discovers in the world, nor is it something you can detect or measure in a laboratory. Or to borrow (and update) David Hume's example, if you examine the scene of a wicked act, such as murder, you may discover much forensic evidence—fingerprints, bloodstains, tire tracks, and so on. But nowhere will you be able to find and catalog the wickedness or wrongness. We face a problem in placing the normative within the natural world, as some philosophers have put it. This problem is most apparent when we make normative claims with apparently declarative sentences ("Murder is wrong," "This conclusion is unjustified"), which will be our primary focus in this work.

Implicit in this tension is an assumption that despite the "pull" that normative claims seem to have for us, they are, at some deeper level, descriptions of a different order of facts. Where we call something good, there would be some entity or property of goodness that made such claims true or false; where some claim or behavior is incorrect, there would be some standard to which it is held. If we assume there are such entities or properties, then we owe some scientifically respectable account of what they are and how they come to be. Many candidates have been offered—people's preferences, their desires, their methods of belief formation, or something else. In this way, it is thought, "oughts" are just statements about the physical world, like any other statement. But many find these approaches deeply unsatisfying. Those who emphasize normative dimensions of human lives often find such surrogates pale imitations of normativity, while those who emphasize the importance of scientific explanation frequently see no reason to posit robust, full-blooded normativity at all.

How do we reconcile these very different dimensions of our lives and our world? One response from many philosophers amounts to a kind of shrug at the apparent tension. There are simply many different sorts of stories (or narratives, or practices, or discourses, or some similar philosophical device) that we tell in coping with our world, and there is no perspective from which to tell a "final story of all things", nor a historical progression toward such a "final story." To the degree that we still find some value in telling any such sort of story, we are free to adopt or abandon it accordingly. The most prominent advocate of such a response in recent decades was Richard Rorty, who repudiated the very idea that we should answer the world in a way that could motivate such a problem.

We must confess a certain sympathy with more nuanced versions of this response (though not much with Rorty's in particular) in their emphasis on multiple parallel approaches to understanding ourselves and our world. But we must also confess that we feel the pull of many parts of contemporary naturalist accounts, and we are wary of solutions that insulate philosophy from scientific inquiry. Many contemporary naturalists have rightly pressed western philosophers on the presumption that there is a set of methods by which philosophy can be conducted prior to and independent of all other forms of inquiry. An important sense in which we think we ought to be naturalists is in thinking that no type of theoretical project—whether it is metaphysics, semantics, ontology, or most of all normativity—is conducted completely independently of the sort of open, world-involving engagement pursued in the natural sciences. We are especially apprehensive about approaches to normativity that add mysterious entities to the world to bear its explanatory load.

But why, we ask, can we not have both? Why not adopt approaches to philosophy that employ all the analytical sophistication and value-driven inquisitiveness we know and love, but which open their borders and weave those virtues in with the work we do in the natural sciences? Why not have both full-blooded normative discourse and some suitably moderate form of naturalism? The crucial sticking point remains the assumption that declarative assertions in normative discourse must describe (or somehow reveal) some deeper level of natural facts in addition to their action-guiding import, and this deeper level of natural facts simply cannot be placed in the world that naturalists describe. We will

argue that this assumption is fundamentally wrong. Normative claims do not aim, primarily, to describe; they urge us to do, or not do, something. Normative claims tell us that a state of affairs ought to be, or not be, a certain way.

With this in mind, we can pursue an alternative conception of normativity on which even its declarative statements are not stating facts. Indeed, they are not making descriptive claims at all. Rather, such discourse serves a fundamentally action-guiding role: it prescribes behavior (or proscribes it), or recommends a course of action (or recommends against it), and so on. We will show how normative claims are constrained by how the world, even though they do not describe this world. This constraint, combined with the availability of reasons and arguments for and against normative claims, also has the benefit of preventing a fall into relativism. If normative claims do not describe the world, then they cannot be describing the world in a way that conflicts with science. Thus, by abandoning the view that normative claims are in the business of describing the world, we will show how to reconcile the normative with natural science and a wealth of other forms of discourse.

The Road from Here

In Chaps. 1 and 2, we note that there are many competing proposals for what sort of philosophical program the term "naturalism" should entail, and just how much authority its demands should have for contemporary philosophers. Rather than trying to sort through all of these competitors, we concentrate on a number of methodological themes that run through most self-identified naturalist accounts and explain why we have some degree of allegiance with each of them. However, the normative is an ineliminable part of our lives—including our scientific practices—and so it is important to give an account of the normative that comports with these naturalist themes.

In Chap. 3, we look at the prevailing views on how to place the properties of normative discourse in the physical world, that is, reductionism and non-reductive supervenience. We argue that contemporary reductionist accounts do not fall prey to the naturalistic fallacy, but even their

more sophisticated approaches leave us with a "normative surplus" that cannot be reduced or identified with any items in the physical world. Most importantly for our purposes, we see a common fault in such accounts of looking for some non-normative correlate to do the work of determining normative matters. We will argue that the very attempt to find such a correlate is misguided, as no correlate would do. This leads us to a discussion of deflationist accounts of truth and varieties of pluralism in Chap. 4.

In Chap. 5, we introduce a theoretical notion of an interest. We argue that interests are particularly important in understanding normative discourse as truth-apt and objective without appeal to normative objects and properties. Interests are matters impressed upon us by the world, rather than matters of assent and consensus. But they are not matters of representation of some feature of that world, so there is not even a purport to posit new entities that would vitiate naturalism. In Chap. 6, we use this to initiate a novel account of action-guiding content for elements of normative discourse. We contend that claims made in normative discourse are expressive of something quite different from familiar forms of descriptive discourse, but that this content is still something that we believe and something that can be true or false. In Chap. 7, we then offer an account of normative discourse that downplays its descriptive, fact-stating role. To do this, we build on work in Chaps. 5 and 6 to demonstrate how normative claims can be non-relatively true and then present an account of how the empirical can constrain the normative even if normative discourse does not serve a fact-stating role.

In Chap. 8, we return to a theme introduced in Chap. 4. There, we argued that even if we are committed to a non-reductionist account of normative discourse, among many others, this cannot commit us to "non-overlapping magisteria" of discourse in which the commitments we make for one sort of theoretical project are insulated from those of others. Taking cues from work in the philosophy of science, we argue that there is another alternative open to us. One "region" of discourse may be said to contribute to another in various theoretically fruitful ways. This has the virtue of unifying different regions of discourse, rather than insulating them from one another, in ways that avoid reductionism while being more conducive to the naturalist themes in Chap. 1. In Chap. 9, we build

on the theme of contribution from Chap. 8. Here, we argue that there are numerous ways in which normative discourse and non-normative discourse contribute to one another's projects in fruitful ways without either one reducing to the other.

1

Naturalist Themes

1.1 Naturalist Themes: Science, Ontology, Anti-Transcendentalism

One of the great difficulties in deciding how to reconcile our fairly robust views on normativity with naturalism is that there are about as many forms of naturalism as there are naturalists. The label has been adopted by or ascribed to philosophers as disparate in their views as John Dewey, Frank Ramsey, Roy Wood Sellars, Wilfrid Sellars, Ernest Nagel, David Armstrong, W.V.O. Quine, Thomas Kuhn, Philip Kitcher, Paul and Patricia Churchland, J.L. Mackie, Philippa Foot, Aristotle, David Hume, Ludwig Wittgenstein, Richard Rorty, and even Jacques Derrida (Staten 2008); to ordinary language philosophers, to experimental philosophers, and to the entire pragmatist tradition. We do not confront a single doctrine in naturalism, but rather numerous methodologies, motivations, and projects; which of these to adopt and which to dismiss will be substantial philosophical questions.

Perhaps one unifying feature in all the various approaches that lay some claim to the title of "naturalism" would be a purport to treat humankind in all its various dimensions as part of the natural world without

privilege, priority, or enchantment. This would generally entail a rejection of any "first philosophy" in the spirit of Plato that might precede our observations, experiences, and practices. For some (but by no means all) putative naturalists, this implies an especially close relationship between the sciences and philosophy. This may grant scientific discourse a kind of authority on certain questions, and by some estimates, *good* philosophy will simply be those modes that clear obstacles for subsequent scientific inquiry. The business of philosophy will be putting philosophy out of business, either by assimilating philosophical projects to projects in the sciences or by undercutting older philosophical modes, problems, and assumptions altogether. As Quine said:

> Is this sort of thing still philosophy? Naturalism brings a salutary blurring of such boundaries… It undertakes to clarify, organize, and simplify the broadest and most basic concepts, and to analyze scientific method and evidence within the framework of science itself. The boundary between naturalistic philosophy and the rest of science is just a vague matter of degree. (1995, 256–257)

This will reek of treason to many philosophers, but this image of the sciences eclipsing philosophy over time will have its critics even among those named above. To some naturalists, the sciences are simply further sets of historically conditioned social practices for coping with experience, and they will have no ultimate authority over others. Rorty (1979, 1989) made such claims explicitly and associated many of the more important philosophical voices of the twentieth century with his project. That in turn will sound like treason to more scientifically inclined naturalists, but it suggests another way of interpreting the naturalist tenet that there is no privileged position from which inquiry can begin. At most, we can only articulate the significance of our practices from within the perspective they afford us, rather than holding them to some higher ideal such as objectivity or progress. There is a reading of Thomas Kuhn's work that lends itself to this sort of naturalism.[1] Where philosophers since the seventeenth century had generally sought to demarcate science from other forms of discourse and grant it a greater measure of epistemic legitimacy

in uncovering truths, Kuhn emphasized the importance of reading each historical phase of scientific inquiry in its own light.

> Rather than seeking the permanent contributions of an older science to our present vantage, [historians of science] attempt to display the historical integrity of that science in its own time. They ask, for example, not about the relation of Galileo's views to those of modern science, but rather about the relationship between his views and those of his group, i.e. his teachers, contemporaries, and immediate successors in the sciences… By implication, at least, these historical studies suggest the possibility of a new image of science. (1970, 3)

That "new image" would thus abandon the view of science as a progressive, cumulative, world-representing enterprise rising above its time and place, and instead treat it as a further set of problem-solving practices best understood in light of their social and historical context. Science itself would thereby become one more item to situate in the history of natural world, rather than a means to step outside it and look back in.

Why should all of this concern us? Very little is at stake for the future of any given philosophical project or method in the name that we assign it. Yet we do think that there are issues of genuine philosophical concern that we must address here, even in the absence of a single, unified doctrine of naturalism. While the many strains of naturalism at hand differ greatly from one another, there is still a sense in which they are responses to a common set of concerns that have some traction for us. Rather than offering a unified doctrine of naturalism, we can approach the matter thematically, by making explicit a number of these animating concerns and illustrating how different forms of naturalism incorporate them. Not all who call themselves naturalists will endorse all of these themes, but there will be considerable overlap, and in our affinity to these themes, we are closer to the naturalists than to most non-naturalists in contemporary philosophy. We see three major themes that cut across most of those lumped together under the banner of naturalism: (a) a priority assigned to scientific practices; (b) ontological and explanatory conservatism; and (c) anti-transcendentalism. Once these themes are on the table, it will be possible to elaborate why we see each one concerns us as we offer our account.

1.1.1 Naturalism and the Priority of Scientific Methods

Naturalists often grant some form of priority to the methods of the natural sciences and their results. With a nod to Protagoras, Sellars said "in the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that it is not" (1956/1997, §42). In some cases, this priority is granted in light of the predictive and instrumental successes of scientific practices. Scientific methods and practices have proven themselves reliable sources of knowledge, we might say, and there are no grounds but tradition and dogmatism to assign them second-class status. This scientific orientation overturns the Platonistic assumption that only the methods of philosophy lay bare for us the real structure of the world, the nature of knowledge, and the purpose of our lives in it. According to the most robust forms of naturalism, there would be no special role left over for philosophy but to clear the clutter and confusion that might inhibit scientific progress. "[It] is within science itself, and not in some prior philosophy, that reality is to be identified and described," as Quine put the point, and this implies the "abandonment of the goal of a first philosophy prior to natural science" (1981, 21, 67). Scientific practices would thus be at least the peers, perhaps even the more able and enlightening successors, to the canonical methods of philosophy.

How radically would this priority of scientific methods and results reshape the landscape of our knowledge? Much of the debate on this point has been shaped by Sellars's "Philosophy and the Scientific Image of Man" (1962), in which he claimed that modern philosophy generates the "manifest image" of the world, "a [philosophical] refinement or sophistication of what might be called the 'original' image... which makes it relevant to the contemporary intellectual scene" (§17). But scientific practices will generate the "scientific image," which at least appears to supersede the "manifest image" from which it emerged. Sellars suggested that we may hope to see our world through both images in the future, but many naturalists have been less concerned with preserving our philosophical past. On such strongly science-friendly forms of

naturalism, all of our theoretical understanding (including the crude theories we call "common sense") will give way to scientific successors in the future (e.g., Churchland 1981, 1996; Stich 1983; Rosenberg 2014). Some would say that philosophy serves only to clarify and organize the work then undertaken in scientific practices (Quine 1969, 1995), while others have focused on naturalizing particular regions of traditional philosophy such as language, mind, and knowledge (e.g., Millikan 1984; A. I. Goldman 1992; Rosenberg 1999; Kornblith 2002). More recently, some philosophers have applied methods from social psychology to the critical examination of central philosophical intuitions under the banner of "experimental philosophy" (e.g., Knobe and Nichols 2007).

Outside the confines of academic philosophy, there is widespread suspicion that even the most science-friendly proposals here are too little, too late for philosophy as a discipline. Stephen Hawking, perhaps the preeminent public figure in the natural sciences in recent decades, declares philosophy "dead" and that "scientists have become the bearers of torch of discovery in our quest for knowledge" (Warman 2011). This has brought charges of "scientism," or a hegemony of scientific methods born of dogmatism rather than innovation or insight. McDowell (1994) and Margolis (2003) both resist philosophical trends they see as scientism, for instance. To those more committed to "perennial philosophy," this *overtaking* of philosophy by scientific methods will seem unmotivated. Science succeeds on its own terms, but scientific practices do not generate the grounds for their own legitimacy. That we should value the knowledge or instrumental possibilities they produce and privilege them over other methods and practices is not a matter that scientific practices even purport to settle. Know-nothing declarations of disinterest in philosophy do not eliminate these problems, so it will be premature to say that naturalism can set aside all philosophical reflection. Dan Dennett, a staunch proponent of approaches to the mind deeply informed by the sciences, struck a gently defiant note on the subject recently: "There is no such thing as philosophy-free science, just science that has been conducted without any consideration of its underlying philosophical assumptions" (2013, 20).

We share with most naturalists the sense that scientific practices should be granted some priority in "describing and identifying" the contents of the world, as Quine put it, although we do not share the more radical ambitions of some forms of scientism to supplant all philosophical (and other) discourse with some scientific successor. We see reason to be catholic in our methods and practices, embracing whatever contributes positively to our interests and goals in inquiry, and our scientific practices have more than earned a place at the table at this point. Many philosophers are apprehensive about the crude myopia of dogmatic scientism, and we share such apprehensions, but our response to this should not be to circle the wagons around some idealized conception of philosophy that insulates it from other forms of discourse.

Here, it may be helpful to bear in mind a distinction offered by Huw Price between what he calls *object* naturalism and *subject* naturalism. An object naturalist takes it that ontologically "all there *is* is the world studied by science," and epistemically "all genuine knowledge is scientific knowledge." Subject naturalists take it that "humans are natural creatures, and if the claims and ambitions of philosophy conflict with this view, then philosophy will have to give way" (2004/2011, 185–186). Subject naturalists would incorporate scientific discourse into our philosophical self-reflection as a partner, even if this discourse does not replace those reflective modes wholesale. Adopting such an approach will require novelty and ingenuity to avoid either collapsing into scientism or diluting the role of scientific practices to window-dressing. We see this as a tremendously difficult challenge, but just the one this naturalist theme rightly prompts. Our goal should be a better fusion of those parts of science, philosophy, and many other practices that best inform our understanding and guide the pursuit of our interests. How all of those parts should hang together, whether we should adopt new ones and abandon old ones, and how we should conceive of our goals in this self-correcting enterprise are all matters that we will find ourselves compelled to revisit; and any conclusions we reach are provisional and remain open to future challenges. But note that any such reflexive examination of our practices will be a form of philosophical inquiry. We have every reason to assume that philosophy will always be with us, even if it looks very different from

its past and present forms. Count us among those in whom such a prospect inspires awe and excitement, rather than dread.

Our primary concern in this book remains normativity, however, and many naturalists have been particularly suspicious about it. Reconciling normative discourse with our various forms of scientific discourse is no trivial matter, and many would prefer to see some reduction of the normative to non-normative, or perhaps a successor that was somewhat more anemic but more readily placed in the physical world. We will argue that no substitute for normativity drawn from biology, psychology, or the social sciences will suffice, and we must commit to robust notions of normativity that do not have a home in any scientific theory. A key move in this project will be to demonstrate that there are expressive modes in addition to describing and reporting that are contentful in their own right, not merely expressive of our attitudes.

1.1.2 Naturalism as Ontological and Explanatory Conservatism

A second theme running through many naturalist accounts is something we will call *ontological conservatism*. Entities should not be multiplied beyond the necessities of our best causal accounts of the world, to pinch a little from Ockham. This does not necessarily imply a strict physicalism; we might be ontologically conservative in the present sense while admitting distinctive layers of biological, psychological, and social entities. So long as those entities were composed of parts and processes at least token-identical with parts of the physical world, the urge is satisfied. This is a strongly held guideline counseling us to avoid adding items and ontological categories that we cannot locate in the physical world, or which do not play some ineliminable role in explaining phenomena within it. By contrast, to assert that there were *only* physical entities and properties (and perhaps supervening layers) would be to adopt a metaphysical stance that belied many other critical commitments that naturalists would make: ontological conservatism emerges not from *certainties* about the contents of the universe, but rather from a *suspicion* that further

categories and realms of entities are being posited to shoulder explanatory burdens in misleading ways. In the absence of accounts of how the physical and non-physical could interact with or compose one another, ontological conservatism suggests that many non-physical posits will not be able to serve genuinely explanatory roles, and so we should shun them.

This conservatism is often cashed out in nominalist terms. Again, Quine and Sellars are prominent examples of this. But the approach we are highlighting is not simply nominalism under another name. David Lewis, perhaps the twentieth century's preeminent advocate of nominalism (1983, 1986), did not explicitly endorse naturalism and pursued metaphysics in a manner that did not answer to the natural sciences in the ways we described in Sect. 1.1.1. Armstrong (1978) defends the reality of universals as not merely consistent with naturalism and physicalism, but as essential parts of laws of nature; universals on his account are not abstract objects in the sense of being outside the spatio-temporal confines of the physical world, even though they are not particulars. Tooley (1977) and Dretske (1977) make similar moves, woven into a body of work committed to various naturalist positions (though Dretske does so in a more conditional fashion).[2] However, not all naturalists are unified in their opposition to *all* abstract objects. Quine famously relented and admitted some classes and some mathematical objects into his account in his later years:

> Limited to physical objects though our interests be, an appeal to classes can thus be instrumental in pursuing those interests. I look upon mathematics in general in the same way, in its relation to natural science. But to view classes, numbers, and the rest in this instrumental way is not to deny having reified them; it is only to explain why. (1981, 15)

Quine's admission of them to his "desert landscapes" arises from his sense of their indispensability in our best theories. There appears to be no way to do science without mathematics in his view, and no way to do mathematics without admitting some abstract objects to be real. One could dispute either of these things, of course. Most nominalists would dispute the latter as would fictionalists, (e.g., Field 1989; Balaguer 1996; Yablo 2001).[3]

So ontological conservatism does not categorically deny the existence of non-physical entities, but it suggests there is a considerable cost, best avoided, in incurring such ontological commitments. We can trace such concerns to an interest in causal explanation. We wonder first about those things we can observe, and we presume these have causes that we could discover, describe, and relate to one another. Whatever exists outside the physical world would presumably need some causal pathway to enter the physical world to have these kinds of observable effects. Attempts to articulate this causal story have struck scientists and philosophers alike as implausible. This interest in causal explanation and its centrality to our grounds to make ontological commitments have led many philosophers to formulate physicalism in terms of *causal closure* or the *completeness* of physics (See Melnyk 2003; Papineau 2001, 2009; Vicente 2005 for some recent discussions of causal closure.) Roughly, the physical world would have causal closure if every physical effect had sufficient physical causes:

> I think that physicalism is best formulated, not as the claim that everything is physical, but as the significantly weaker claim that everything that interacts causally with the physical world is physical. (Papineau 2001, 11)

This is a claim for which some have argued we have or could have empirical evidence. Finding it would not disprove the existence of things outside the physical, but the burden would be on those who want to posit further realms to show some other grounds for them. Such grounds are in short supply once the causal loop is closed.

Our perspective on this theme is somewhat unorthodox. Many philosophers as committed to robust, objective accounts of normativity as we are tend to see their naturalistic options as anemic and adopt some form of non-naturalism instead (e.g., Korsgaard 1996; McDowell 1988; Bedke 2012; Parfit 2006, 2011, part 6). Most naturalists who make a point of ontological conservatism prefer to reconcile the normative and the natural in reductionist or supervenience terms, or even outright non-cognitivist ones. We share with ontologically conservative naturalists an aversion to posits of non-physical items as explainers of normativity. The correctness of moral judgments and the content of thoughts and language that use

overtly normative terms should not be understood or explained by positing a further non-physical, non-causal, non-natural realm of objects, properties, relations, or any other such categories. But we will ultimately reject ontologically conservative approaches that correlate moral terms with familiar non-normative terms and properties or suggest that they supervene on non-normative ones in explanatorily fruitful ways. In this sense, we are as skeptical of positive claims about the existence of moral *stuff*, whether in the physical world or in some other realm, as are error theorists like Mackie (1977) and Joyce (2001).

Thus, we take ourselves to be ontologically conservative in the same sense as many naturalists. Our account will not correlate elements of normative discourse with a special class of supernatural objects or properties, nor even with a class of natural objects or properties. In our view, there is something misguided about trying to *place* normativity in the world at all. Although we speak metaphorically of it having a "pull" or a "grip" on us, we have to make normative judgments and claims in order to adjust one another's behavior and attitudes precisely because rightness, correctness, and so on *do not exert causal influence* over things as physical properties do. If we said that supernatural normative entities do not exert causal powers directly, but that we are aware of them through some special intellectual faculty, we would be at a loss to explain such a faculty or how we could interact with such items.

Moreover, we are sympathetic to Mackie's "queerness" object to a certain degree. Assigning normative force to an object seems mysterious, and adding a further realm of supernatural items adds to the mystery, rather than resolving it. But in rejecting these ontological commitments, we will not offer an error theory about moral judgments and discourse. A large part of the drive to resolve the ontological implications of normative judgment stems from the widely held doctrine in the philosophy of language that a normative term like "goodness" must designate something if the expression is to mean something in a truth-evaluable way. Otherwise, it will have to be explained away as a cloaked expression of our emotions, an order given to others, or some other expression of our attitudes. We view this as a choice forced upon us by an unduly limited palette of options in accounting for normative discourse. The very fact that we will account for normativity without reducing its terms to those

of some scientific discourse may lead ardent naturalists to say that we are not of their tribe. If so, so be it, but this strikes us as restrictive without warrant. To the degree that we concern ourselves with ontology in this book, we will not make additional commitments that would be incompatible with causal closure and token physicalism for most domains. In that respect, we take some of the naturalist traditions to have placed appropriate restrictions on our theorizing, and we will strive to operate within those bounds.

1.1.3   Naturalism as Anti-Transcendental

Our third naturalist theme concerns what we will call *anti-transcendental* approaches to philosophical questions. It is widely held, but also subject to a wider range of interpretations, and thus less distinctive as a philosophical position in its own right. But it exposes an interesting seam in the history of philosophy not tracked by the analytic/continental split or the usual botany of subdisciplines. Much of the western tradition has treated philosophical inquiry as a search for eternal truths delivered by modes of reflection epistemically prior to our contingent, historically located practices—a search that takes precedence over empirical methods and results. In this way, philosophy starts from a privileged intellectual position and incorporates a priori intuitions outside the scope of empirical, practical, and historical scrutiny. Its goal is to transcend the limitations of other modes of thinking to deliver knowledge of a more enduring, enlightening sort. Plato looms over western philosophy in this regard, and we can read much of the rationalist and idealist traditions, as well as large parts of contemporary analytic metaphysics, epistemology, and normative ethics in a similar way. For a wide variety of reasons, various philosophers have rejected that perspective in favor of approaches that emphasize empirical evidence and proceed from the practical and historical positions in which we find ourselves. We would include empiricists of all stripes here, but also pragmatists, post-structuralists (and most Continental philosophers, for that matter), and naturalists of many other varieties. While we will situate the details of our account in language familiar to contemporary analytical philosophers at many points, we believe that the departures we

take from some of those positions suggest interesting intersections with philosophers not often brought into debates over naturalism (Heidegger, Foucault, Merleau-Ponty, Haugeland, etc.).

Anti-transcendental approaches may be articulated either negatively or positively. (And many philosophers offer both, often in the same work.) The negative formulation would be to deny that there is any other grounding for a philosophical perspective (methods, intuitions, etc.) from which to determine and reflect upon eternal truths. There is no "view from nowhere" and hence no "perennial philosophy" (as Quine put it above) that we could conduct from it. This negative position has emerged out of despair in some cases and a perceived deficit in others. To those who despair, the western philosophical tradition had over two millennia to take its best shot, and it produced nothing like the results it promised. Rorty (1972, 1979, 1989) has been perhaps the most prominent advocate of this position in recent decades, much to the consternation of analytic philosophers. As Rorty has also noted, the early pragmatists are among those who perceived a deficit in transcendental approaches. These pragmatists lamented that in the quest for fixed certainty, western philosophy had expressed derision for practical concerns and injected hollow and inaccessible abstractions in their place. The result was a view of subjects isolated from their world. As Dewey said:

> Indeed, according to some thinkers the case is even in worse plight: Experience to them is not only something extraneous which is occasionally superimposed upon nature, but it forms a veil or screen which shuts us off from nature, unless in some way it can be "transcended." So something non-natural by way of reason or intuition is introduced, something supra-empirical. (1925/1958, 1a)

Rorty (1982, xvii–xix) connects these same doubts about the transcendental strains of nineteenth-century western philosophy to Nietzsche and later Heidegger and other Continental philosophers.

> James and Dewey were not only waiting at the end of the dialectical road which analytic philosophy travelled, but are waiting at the end of the road which, for example, Foucault and Deleuze are currently travelling. (1982, xviii)

We can also compare Dewey's language here with Wittgenstein's warning that we must avoid "bewitchment of our intelligence by language" (1953, §111).

A positive formulation of anti-transcendentalism would emphasize that philosophical reflection can and should begin with our own embodied experiences and practices, and that these do not discredit our reflection, but rather vivify and motivate the inquiry. Dewey's *Experience and Nature*, quoted above, offers both positive and negative formulations of anti-transcendentalism in its opening chapter, and positive formulations are perhaps most common in the pragmatist tradition. If one interprets Wittgenstein's later work as a substantial departure from his early work (as most do), the primary difference may be read as a transition to an anti-transcendentalist stance. In the later work, there is simply no single, eternal, underlying foundation of thought and language to be uncovered by philosophers, but rather a great tapestry of interlocking shared modes of coping and cooperating ("language games" and "forms of life") in the world. Dreyfus (1990) and Haugeland (1982) have also been instrumental in producing readings of Heidegger (1927/1962) in this vein.

One issue that dominates many anti-transcendental approaches is the perilous status of the objectivity of our normative claims. (For some anti-transcendentalists, this spreads even further, but we will set that aside for the moment.) Without non-natural facts or perennial methods to ground the correctness of our normative claims and judgments, many anti-transcendentalists see no basis for treating them as objective truths. Alternatives vary widely, but some adopt forms of overt relativism (e.g. Harman 1977), others take more anthropological approaches (e.g., Foucault 1977), while others naturalize fields such as ethics more ruthlessly by treating moral psychology as the successor to genuinely normative accounts. Rorty endorsed the abandonment of objectivity more puckishly and generated as much interest outside the confines of academic philosophy as within them. He eschewed the mantle of a "relativist," insisting that that distinction makes sense only within the confines of the very sort of transcendental and representationalist view that he rejected. His view is often flattened out to a simplistic appeal to accepted practice and social assent by his critics, but his motivations and commitments are more subtle than this. For him, the question is how to move

past a philosophical framework in which such concerns have a grip on us at all:

> [T]he question is not about how to define words like "truth" or "rationality" or "knowledge" or "philosophy," but about what self-image our society should have of itself. The ritual invocation of the "need to avoid relativism" is more comprehensible as an expression of the need to preserve certain habits of contemporary European life. These are habits nurtured by the Enlightenment, and justified by it in terms of an appeal of Reason, conceived as a transcultural human ability to correspond to reality, a faculty whose possession and use is demonstrated by obedience to explicit criteria. So the real question about relativism is whether these same habits of intellectual, social, and political life can be justified by a conception of rationality as criterionless muddling through, and by a pragmatist conception of truth. (1991, 28)

We will embrace a kind of anti-transcendentalism, particularly indebted to the pragmatist tradition. We share the enthusiasm of other anti-transcendentalists in rejecting the conception of philosophical inquiry as a shedding of practical, historically informed perspectives in favor of eternal, abstract verities. However, we do not share the view that there is no role to be played by notions of objectivity and answering to the world in our inquiry. We feel that there are a number of conflations in the paths that lead anti-transcendentalists in these directions. First, while we agree that it makes no sense for an anti-transcendentalist to speak of normative facts as practice-independent features of the world waiting for representation, the ways in which some normative claims and judgments are made do not undermine their objectivity in the pertinent sense, which we will elaborate more fully in Chap. 7. They are relative to a set of practices in that only those practices will provide the access that make those judgments intelligible. But the bogeyman of moral and conceptual relativism does not turn on this. It is the further consequence of rational immunity from challenges emanating from outside those practices that should haunt us, and the necessity of being embedded in social practices to make normative claims and judgments does not entail this sort of immunity.[4] There is a role to be played by distinctions that outstrip

the approval of any community—even ideal communities at "the end of inquiry"—that has been played by traditional notions of objectivity. It will not suffice to permit speakers to continue using objectivity-talk as they do with an ironic nudge and a wink from philosophers, either. If we are to be naturalists (and particularly if we adopt the sort of pragmatism being offered here), then we must offer an account with the same sort of world-involving potency without committing the same sorts of mistakes. We elaborate this point at much greater length in Chaps. 4 and 5.

1.2 Why Should We Care about Whether We Are Naturalists?

As will become clear as we continue, we are committed to the viability and objectivity of various types of normative discourse—moral, epistemic, prudential, semantic, and so on. And we have noted a number of important metatheoretical themes that run through the many different philosophical projects that get called naturalism in one way or another. Still, readers might wonder why we should feel any urgency to reconcile an account of normativity with them. Why is it important to be a naturalist, and what sort of naturalism should we adopt?

Like some pragmatists before us, our motivations derive in part from an aversion to elements of abstraction and idealization that stretch back deeply into the western tradition. Plato's approach to philosophy was an attempt to explain our engagement with *this* world by positing *another* world. To explain our cognition of an object in this world by appealing to something existing in another, unseen world seems to us obfuscatory rather than explanatory. Inquiry (including philosophical inquiry) begins in this world and must be located in this world; it cannot begin by stepping out of this world. This is a common theme running through Dewey, Heidegger, Wittgenstein, and Sellars. They are naturalists in the sense that endeavor to explain (or explain away) philosophically interesting notions via our practical engagement with the world. From this perspective, then, the scientific reductionist route is mistaken, but it is in an important sense less mistaken than the Platonist route, in that the

former at least tries to locate all human inquiry in the natural world where we reside, and with which we are practically engaged, rather than trying to explain this engagement by appealing to something mysterious and other-worldly.

The approach we offer in this book is naturalist in the same way that the classical pragmatist approach and its close cousins—that of Dewey, Wittgenstein, and Sellars—are naturalist. We are embodied creatures, whose lives are carried out immersed in a natural world—a world of slabs and hammers, of token linguistic utterances and brain events, and of objects with various sensible qualities. If something is going to matter to us as an account, it has to explain our engagement qua the kind of embodied creatures we are, with the world we inhabit. But this modest naturalism will not require a commitment to reductionism or scientism.

1.2.1 Fallibilism and Methodological Modesty

Another advantage many forms of naturalism have as philosophical approaches is that it embodies the philosophical virtues of methodological modesty and a commitment to fallibilism. We can explain each of these virtues in turn, why it is a virtue, and why a modest naturalism embodies this virtue.

We can say that a philosophical approach is modest in the relevant sense here if it properly acknowledges the limits of our capacities to investigate and comprehend the matters we investigate. Unchecked appeal to a priori intuition or prodigal expansions of unexplained explainers of crucial distinctions might be failures of modesty in this sense. Even some naturalistic approaches fail to exhibit this virtue. Some naturalists seem to believe (a priori, in many cases) that science (or perhaps a future completed science) will have sufficient methodological resources to account for every pertinent concern, and that everything will either be reduced to a scientific field or subfield, or eliminated altogether. Such an expansive scientific reductionism is actually methodologically bold, perhaps even arrogant, and not modest at all. A more modest approach is certainly willing to revise central features of our practice, but not without sufficient philosophical reason; and it certainly is not going to declare

a priori that scientific discourse will suffice for every possible purpose. Methodological modesty would involve, among other things, taking central human practices as the starting place of inquiry and proceeding with at least a faint presumption of their legitimacy. It may emerge, after philosophical inquiry, that practices lacking scientific bona fides cannot sustain their own legitimacy, but again, there is a difference between discovering this after inquiry and assuming this at the outset.

Implicit in the above approach is a thoroughgoing fallibilism. That is, we take our epistemic position to be one in which error is a perennial possibility, and so a proper respect for the limits of our capacities must be woven into our practices at every turn. Epistemic vigilance and self-correction lie at the heart of any such approach. The methodological modesty advocated above would take some set of practices (including scientific, moral, and epistemic) as a starting point, but treat these practices as revisable in the light of compelling reasons. Indeed, we think not only that fallibilism is key to understanding our position as rational agents in the world, but also that naturalism is better than rival approaches at taking fallibilist themes to heart.

Some might consider fallibilism to be a pessimistic position, which treats our knowledge of the world as inherently tenuous and unreliable. But fallibilism only appears so if one begins with a set of inflated expectations about the possibility of certainty, or incorrigible foundations. In the spirit of Peirce, we think Descartes is to blame for these unrealistic expectations, but such a conception of knowledge is neither inevitable nor desirable. An infallibilist view of knowledge encourages dogmatism and undermines the features of our discourse which give us any grounds at all for confidence in our beliefs. As Mill once said:

> In the case of any person whose judgment is really deserving of confidence, how has it become so? Because he has kept his mind open to criticism of his opinions and conduct…*The steady habit of correcting and completing his own opinion by collating it with those of others, so far from causing doubt and hesitation in carrying it into practice, is the only stable foundation for a just reliance on it* (Mill 1869/1978, II.7, emphasis added)

Read optimistically, infallibilism would privilege some intuitions ("*cogito ergo sum…*") as immune to revision in ways that ossify inquiry;

read pessimistically, it would set an unbearably high standard for knowledge that devalued our actual practices of inquiry (see the dreary history of epistemic skepticism). Both of these would be gross errors born of undue allegiance to one epistemic intuition.

By contrast, fallibilism helps us understand how we, as real embodied agents, cope rationally with the world. For example, only by embracing fallibilism can we capture the insight that rationality is not in the first instance a feature of the structure of a belief system at a given time (who has the epistemic resources to evaluate his or her own entire belief system?), but is rather a feature of diachronic processes of revision in response to evidence and other sorts of authoritative inputs. But to accept that *rationality* requires *revisability* is to accept *fallibilism* of a thoroughgoing sort.

This insight was seen (sometimes dimly, sometimes clearly) by early pragmatists, many of whom were explicitly reacting against the Cartesian conception of knowledge. Indeed, one can (without committing too much Procrustean sin) distinguish between two different epistemological methodologies[5]: the Cartesian method of radical doubt and the pragmatist method of fallibilist conservatism.[6] The Cartesian method starts by doubting everything and then attempts to rebuild all of knowledge on a foundation of beliefs which are both certain and incorrigible. The pragmatist method suggests that this method is mistaken along several dimensions. First, it simply is not possible to begin by doubting everything. Peirce, in some of his earliest writings establishing pragmatism as an independent school of thought, writes, "We cannot begin with complete doubt. We must begin with all the prejudices which we actually have when we enter upon the study of philosophy" (1992, 28). As Wittgenstein noted in a similar spirit, one can only doubt some things if other things are not doubted:

> [T]he questions that we raise and our doubts depend on the fact that some propositions are exempt from doubt, are as it were like hinges on which those turn. That is to say, it belongs to the logic of our scientific investigations that certain things are *indeed* not doubted. But it isn't that the situation is like this: We just can't investigate everything, and for that reason we

are forced to rest content with assumption. If I want the door to turn, the hinges must stay put. (1969, §341–343)

The second error in the Cartesian approach is related to the first, and this is evident in how the two methodologies conceive of rationality. The pragmatists' understanding that all inquiry must have a starting point of things that are not doubted shifts the focus of our investigation into the nature of rationality from the *structure* of belief systems toward an investigation of how these belief systems are *revised* over time. We cannot raze our belief system to the ground and rebuild it from scratch. Thus, rationality is not a feature of the structure of a belief system in a given time-slice; rather, it is a diachronic property of belief systems.[7] Sellars wrote that the problem with the traditional view of empirical knowledge is

> its static character. One seems forced to choose between the picture of an elephant which rests on a tortoise (What supports the tortoise?) and the picture of a great Hegelian serpent of knowledge with its tail in its mouth (Where does it begin?). Neither will do. For empirical knowledge, like its sophisticated extension, science, is rational, not because it has a *foundation* but because it is a self-correcting enterprise which can put *any* claim in jeopardy, though not *all* at once. (1956/1997, §38)

A crucial thing to notice about this model of rationality (and here is where we begin to see the second flaw in the Cartesian method) is that it requires a fallibilist conception of knowledge. "We must begin with all the prejudices which we actually have," and each of these "prejudices" is subject to revision, and hence must be treated as fallible.[8] Even Peirce and Wittgenstein, who regarded certain propositions as de facto immune to challenge, conceded that such propositions can be subject to revision over time. That is, they distinguished (as Bernstein 2010, 34 notes) between *indubitability* and *incorrigibility*. For example, Wittgenstein famously argues that it is impossible (not merely psychologically, but perhaps also logically) to doubt some propositions—doubt simply cannot get a toe-hold—because there is no context apparent to us in which the doubt makes sense.[9] But even these propositions could be rejected in time:

"the same proposition may get treated at one time as something to test by experience, at another as a rule of testing" (1969, §98). Amusingly, Wittgenstein (1969, §§106–12) offers "No one has ever been to the moon" as an example of a proposition so basic that one would not know how to engage with someone who asserted its negation. Nowadays, only people who are completely ignorant (or members of the lunatic fringe) sincerely assert this proposition.

The second error in the Cartesian methodology is now thrown into sharp relief: agents are not rational to the extent that they identify a set of beliefs that are certain and incorrigible, and then build an edifice of knowledge on these unchanging foundations. Rather, rationality is essentially tied to a willingness to acknowledge one's fallibility, and to revise one's beliefs (even one's core beliefs) should there be sufficiently compelling reason to do so. The pragmatists' fallibilism, not Descartes' infallibilism, is the key to understanding rationality.

An examination of the actual process of theory choice, and the reasons why we have confidence in certain classes of theories, should cement the place of fallibilism and revision at the heart of our conception of rationality. Our great advances in knowledge (particularly since the start of the Scientific Revolution) have come in large part by modifying or discarding old ways of conceiving of the world. Indeed, the fallibilism which is implicit in much of scientific practice is a large part of why we like scientific practices and remain open to them in ways that many philosophers will not be. Physicist Carlo Rovelli writes

> Science is not about certainty…In fact, not only is it not certain, but it's the lack of certainty that grounds it. Scientific ideas are credible not because they are sure but because they're the ones that have survived all the possible past critiques, and they're the most credible because they were put on the table for everybody's criticism. (Rovelli 2014, n.p.)

Treating beliefs as *de jure* unrevisable makes it in fact impossible to undergo the kind of processes that would confer any kind of rational justification on them in the first place. A core pragmatist commitment is that philosophy begins with our practices as they are actually structured. Infallibilism cannot capture the sense in which a theory is treated

as credible, but subject to challenge and revision. It advances a family of epistemic concepts that are irrelevant to actual epistemic practice that we should therefore reject. By contrast, we think that most naturalist approaches do a better job of taking fallibilist themes to heart. A methodological commitment common to naturalists is that one cannot simply sit in one's armchair and acquire transcendental knowledge. The business of acquiring knowledge is messy, requires engagement with the world, and is inevitably one of trial and error.

1.2.2 Naturalism and Ontological Parsimony

Commitment to naturalism carries with it a certain sort of ontological parsimony: We will be suspicious of supernatural posits as ways of accounting for various phenomena of philosophical interest (such as values, universals, and semantic contents). We will thus be committed to a kind of token physicalism, according to which we will endeavor to explain our core practices without appeal to properties or entities whose existence is not sanctioned by our best scientific understanding of the world. We say that we are committed to a token physicalism in that while we think our ontology should be constrained by our scientific theorizing, we are in no way committed to the idea that all types of discourse should be reducible to or translatable into physics.

Our suspicion of such extra-physical posits does not arise out of a dogmatic attachment to physicalism. Instead, our suspicion is that such posits are generally ones that philosophers often feel compelled to make because of theoretical limits imposed by assumptions they hold elsewhere, many of which we will challenge. For example, consider a typical argument in favor of realism about universals, which starts from the premise that in a subject–predicate sentence, both the subject and the predicate must play their roles by designating some object.[10] Take the sentence, "Socrates is courageous." This sentence can only be true if there is some object in the world corresponding to the word "Socrates," but it can also only be true if there is something in the world corresponding to the word "courageous" (and if the former exemplifies the latter). Now, "courageous" is a general term; it can apply not just to Socrates, but to

Plato, Aristotle, and so on. Thus, "courageous" has the same designatum in the sentences "Socrates is courageous" and "Plato is courageous." From the above considerations, "the realist concludes that 'courageous' picks out a single entity in [these two sentences], a single entity such that in virtue of being related to it, both Socrates and Plato count as courageous" (Loux 2001, 26). The realist about universals concludes that abstract objects must have real existence.

Note, though, that this argument depends upon a *representationalist* view of language. On this view, subjects and predicates have content in virtue of bearing a relation of *designation* with the world; and we can only understand other semantic notions (such as truth) in terms of this semantic primitive (designation). However, we see good reasons to question representationalism, especially since it seems to carry with it such problematic collateral commitments. Language does represent the world, and any account of language must do justice to this feature of our linguistic lives. But this does not require us to take representation (and representational notions such as reference) as the starting point for building up other semantic notions (such as meaning and truth). In this work, we will challenge these assumptions, following in the path of other philosophers (such as Sellars and Brandom) who have argued against the representationalist consensus.

In emphasizing ontological parsimony, we do not locate the philosophical danger in an ontology with a large number of items. If one is a nominalist and a naturalist, there may still presumably be innumerably many items in one's ontology. Indeed, if all one cares about is number of items, supernaturalism might be more parsimonious. The real danger lies not in the number of objects, but in their *type*. When the non-naturalist or the supernaturalist posits a property or an entity (such as an abstract universal) to shoulder a particular explanatory burden, we find ourselves agreeing with naturalists that this is problematic on several levels. First, such posits seem to us to be just-so stories, and a theory of what such posits would be like looks like an explanatory island with respect to the rest of our world-view. That is, there is a lack of explanatory links between this and our other accounts, tying them together as part of a comprehensive theory of the world. When a theory is explanatorily isolated like this, that makes a strong case that the theory is an ad hoc explainer, rather than

one in which we should invest real confidence. Second, such posits are almost unavoidably "queer" in Mackie's sense of the word. If these entities and properties are not the sort of things that fit in a naturalistic view of the world, how can we, as embodied cognizers, interact with them and incorporate them into our practices? Again, these posits seem to mystify rather than genuinely explain.

Third, these abstract properties and objects are bad at explaining our normative engagement with the world. To appropriate Richard Kraut's example (2011), consider a sentence like, "Smoking is bad," where badness is taken to be an objective property, independent of our attitudes and interests. The sentence just means that smoking is bad, *simpliciter*. What are we to make of such a claim? Kraut argues that such sentences, while meaningful, are false; talk of absolute goodness or badness is comparable to talk of phlogiston, a relic of an outdated and faulty world-view. To be sure, smoking is bad for the health of one who smokes, but in saying this, we need not invoke badness-as-such. Anything that is good (or bad) is good (or bad) for someone, or good (or bad) of a kind. We would say that all normativity arises in the context of our practical engagement with the world, and hence must relate to our interests as they are embodied in and affected by this engagement. Thus, smoking is bad because it is bad for our health; lying is bad because it is detrimental to our interest in truth, communication, and community; double-blind studies are good because they advance the cause of inquiry and discovering, say, effective medical treatments. None of these explanations appeal to good or bad *simpliciter*. To appeal to interests in explaining normative practices is not to commit ourselves to consequentialism. We will argue later there is no straightforward way to reify and quantify interests or interest satisfaction; and even if there were, it does not follow that our obligation would be to maximize this quantity. (By analogy, neo-Aristotelianism appeals to a notion of human flourishing in explaining the content of morality without in any way being committed to some kind of utilitarian calculus or maxim of greatest overall happiness.)

The way in which goodness or badness is related to our interests and our practical engagements brings out another aspect of the inadequacy of the role of abstract properties or objects in explaining normativity. We find ourselves much more sympathetic to Scanlon's "buck-passing"

account—"Being good, or valuable is not a property that itself provides a reason to respond to a thing in certain ways. Rather, to be good or valuable is to have other properties that constitute such reasons" (Scanlon 2000, 97).[11] If something had an abstract property of goodness, but did not engage any further interest of ours—did not promote health, or community, or knowledge—we would frankly be at a loss to explain why we should *care* about it at all. And thus, we can only explain goodness or badness (or any property that is supposed to have a normative pull, be *reason-giving*) through actual involvement with our concerns and practical engagements, not with respect to some kind of abstract property or object. Consider again Kraut's example, the idea that smoking is bad, *simpliciter*. Kraut supposes that such sentences are meaningful, but false. But suppose, for a moment, that this sentence is true. Smoking is bad, *simpliciter*, not bad in virtue of causing cancer, or emphysema, or chronic obstructive pulmonary disease (COPD), or in virtue of tending to the detriment of any human interest, just simply *bad*. How can this abstract property of badness give us any reason not to smoke? To be sure, the fact that smoking is bad for our health is a reason not to do so, but this detrimental effect on our interests is what *makes* smoking bad. Thus, we can only understand normative properties, *and their reason-giving force*, by reference to our interests, and not in terms of abstract properties and objects.

For all of the above reasons—our commitment to token physicalism, our suspicion of abstract properties, our pragmatist conviction that representationalist theories of language complicate our ontology without providing sufficient explanatory payoff—we think that greater insight on the nature of normativity is to be gained by shunning commitment to substantive normative properties. Rather, we find it more useful to articulate the content of normative expressions in pragmatic terms. In this sense, we are more the heirs of Sellars on abstract objects; and indeed, it is perhaps by looking at what he says on this issue that one can get a flavor for the pragmatist approach we favor to normative expressions.[12] For Sellars, meanings are simply functional roles that words play in a language (1963, 1974). He uses the device of "dot quotes" to name these roles. Thus, ●red● denotes the role played by "red," "rouge," "rot," and so on. In Sellars's parlance, "red," "rouge," "rot," and so on are all ●red●s. Thus, Sellars explains meaning not by relating words to some extra-linguistic

item, but to their role in actual linguistic practice. His answer is thoroughly *pragmatist*. Further, this device allows him to give an account of universals that does not commit him to abstract objects in the way the representationalist we discussed above was.

Without endorsing Sellars's account in every last detail,[13] we think that it provides an excellent model for how to think about, generally, how language works once one abandons the representationalist model, and specifically, for how to think about normative discourse. Once one abandons the idea that normative terms must have their content in virtue of designating something in the world external to our practices, one is able to spell out an account of normative discourse that is congenial to token physicalism, and at the same time pragmatist in character. On the view we inherit from Wittgenstein, Sellars, and Brandom, norms are implicit in our practices. When one speaks, one is implicitly following semantical norms; when one interacts with others, one is implicitly following various norms of etiquette and social interaction; when one plies a craft or trade, one is implicitly (at least if one is an expert) following various norms of that craft or trade. Overt expression of these norms (say, in the form of a linguistic utterance) serves merely to express a norm that is already implicit in practice (or, alternately, to call for the revision of a practice by claiming that the members of this practice are committed, by various considerations, to this alternate norm). On such a view, normative utterances serve a primarily *expressive* or *legislative* role, not in the first instance a *fact-stating* or *causal-explanatory* one. Thus, just as Sellars sought to explain abstract entities by appeal to the linguistic role of general terms (rather than by appealing to a relation between such terms and an extra-linguistic reality), we will explain the role of normative utterances in terms of their expressive function, and not in terms of their representational role.

To say this is not to say that the role of such terms is cut off from practice or from empirical reality. If one is to interpret the role words play in various languages broadly as their inferential role, then this will include (to use Sellars's terminology) not just intralinguistic transitions (inferences, e.g., from "The light is red" to "The light is colored"), but also language-entry moves (e.g., moving from an experience of redness to a perceptual judgment, "The light is red") and language-exit transitions

(e.g., moving from the judgment, "The light is red" to applying one's brakes).[14] His account and ours are no linguistic idealism; there is still a fundamental connection between his theory of abstract entities and how linguistic practice is embedded in empirical reality. Similarly, on our account, when one uses a normative utterance to express a norm which is implicit in practice, there is very much an engagement with practice, and with the world: the norms themselves are fundamentally engaged with empirical reality. Thus, one is normatively bound to read the word "normative" in a particular way, or to stop upon arriving at a red traffic light, or form a particular belief when finding Smith's fingerprints at the crime scene, and so on. Norms deal, in large measure, with our engagement with the world, and normative expressions only make explicit an existing engagement with the world—an engagement that is freighted with rules, appropriatenesses, "oughts," "mays," "ought-nots," and "may-nots." Thus, our expressivist account of normative discount is no more detached from reality than is Sellars's account of abstract entities, despite its rejection of representationalist assumptions.

Many questions remain to be answered about such an account. For example, what is it for a norm to be implicit in a practice? Brandom holds that "the norms that govern our discursive conduct [are] instituted ultimately by our *attitudes*" (1994, 280), primarily by our attitudes of attributing and undertaking commitments and entitlements. A key question that arises for such an account is whether it can handle charges of relativism. If all normativity arises out of a social practice, then is there any way to make room for objectivity? In his distinction between different attitudes we take (e.g., in attributing to a speaker commitment to P and entitlement to P), and in his distinction between *de dicto* locutions (which make explicit the commitments the person we are attributing to would acknowledge) versus *de re* locutions (which make explicit the commitments we ourselves undertake), Brandom thinks he has the tools to generate an account of normativity that allows for objectivity. That is, he thinks he can give a social practice account, but one within which

genuine, and therefore *objective*, conceptual norms can be elaborated. These bind the community of concept-users in such a way that it is possible

not only for individuals but for the whole community to be mistaken in its assessments of what they require in particular cases. (1994, 54)

We will offer our own account, rooted in the same pragmatist tradition, of how normative statuses arise out of social practices, and yet still can be objective and truth-apt. But let us not get ahead of ourselves. For the moment, let us continue our discussion of naturalism.

1.2.3   Naturalism or Scientism?

"Most contemporary philosophers," write Darwall, Gibbard, and Railton, "would agree that our going view treats empirical science as the paradigm of synthetic knowledge," our "going view of the world" (1992, 126). We agree that science provides us with our best account of the world in a certain sense, but would also caution that this admission should not encourage an embrace of a scientism which tells us that all legitimate forms of discourse must, in the end, be *assimilated* to science (or be eliminated, if such assimilation proves impossible). We will discuss in Chap. 2 some specific accounts which are guilty (by their own admission!) of scientism. Certainly, various advocates of scientism have attempted to "scientize" some of the more philosophically popular types of discourse (such as ethics and epistemology), about which we will say more later. But we find it genuinely implausible that a scientific account can be given of all other types of discourse. Consider the category of imperatives. Is there some physical commonality shared by, say, imperatives issued by a parent to a child, a lieutenant to a private, a manager to a salesman, a person playing "Simon Says" to the participants, and so forth? To be sure, a linguist might be able to identify various syntactic features of imperative performances that characterize them as such, but how (for example) to give a scientific rendering of the authority of some utterances and the absence of authority in others (say, an imperative uttered by a child to a parent in a specific circumstance, or one uttered by a private to a lieutenant)? In many cases, this authority is conventional, but even if conventional, we doubt that a recognizably scientific account can be given of it. A theme

we will return to is that it is even less plausible that a scientific account can be given of, for example, moral and epistemic authority.

Denying scientism does not preclude a commitment to some form of scientific realism, and it does not preclude the stance that science might be the final word on the kinds of questions it purports to address. We do not claim that some discipline other than physics might have the final word on the ultimate constitution of matter, or the relation between matter and energy. Nor do we deny that some range of scientific disciplines (such as biology and biochemistry) will be the ultimate arbiters of claims about evolution. This is not to say that the truth of any such question is a matter of what the practitioners of that field say or believe at any given time; it is merely to privilege the claims of such practitioners, and the methods of this practice, over those of other practices when it comes to investigating truths in a specific domain. It is a question of which claims we find epistemically authoritative, not a question of whose mere utterances are truth-constituting.

But acknowledging science as our best causal-explanatory account of the world does not require us to embrace scientism. It is one thing to say that science is authoritative on certain questions; it is another thing to say that every area of discourse must either be restatable in scientific terms or be eliminated on the grounds that it cannot establish its scientific bona fides. We should instead embrace naturalism, but of a more modest variety. One way in which we could realize such a modest naturalism would be to recognize the authority of science within its sphere, but to recognize that giving science this type of authority does not entail delegitimizing other sorts of discourse, in particular discourse which does not (and is not meant to) serve a causal-explanatory role. Of particular interest to us in this work are types of discourse which are primarily evaluative, such as moral and epistemological discourse.[15] Our view is that normative discourse does not serve the purpose of causal explanation, and so is not a theoretical rival to scientific discourse. It serves another purpose that cannot be duplicated from within scientific discourse. The modest naturalist will insist that such non-explanatory forms of discourse can be legitimate, provided that in giving our account of them, we are not compelled to commit ourselves to any properties or powers that are themselves incompatible with a scientific world-view. Thus, for example, if having moral

knowledge required that agents possessed a special, non-natural faculty of moral intuition, one that could not be reconciled with a scientific account of our cognitive faculties, then that would be a serious difficulty for our account of moral knowledge. But as long as we can offer an account of normative discourse which does not offend against the commitments we have made in our scientific discourse, it is mere parochialism to deny that such forms of discourse can be legitimate, or to insist that only scientific or causal-explanatory forms of discourse can serve any important human purpose. Thus, we align ourselves with Sellars, who writes,

> [O]nce the tautology 'The world is described by descriptive concepts' is freed from the idea that the business of all non-logical concepts is to describe, the way is clear to an *ungrudging* recognition that many expressions which empiricists have relegated to second-class citizenship in discourse, are not *inferior*, just *different*. (1957, 282/§79)

In avoiding scientism, it is important not to commit the opposing error of simply allowing discourse areas to proliferate without making any attempt to reconcile the commitments made within each one. Thus, in saying that, for instance, moral and epistemic discourses are legitimate, we do not endorse a doctrine of "non-overlapping magisteria" (NOM), which would serve to insulate other kinds of discourse from the reach of science. (Most commonly, NOM is held with respect to science and religion.) The problem with NOM is that it creates an artificial distinction between different kinds of discourse; it holds that we can make a set of commitments in discourse A, and a set of commitments in discourse B, but that we are never obligated to reconcile these different sets of commitments. Again, this strikes us as wildly implausible, and its implausibility stems from a view of discourses in which discourses are like suburban yards, neatly fenced off from each other, with well-demarcated boundary lines. We will have a great deal more to say about this, but for now, let us just say that while the notion of a discourse area might be of pragmatic use, the notion really only marks off a set of beliefs, questions, and concerns which serve a particular set of interests. A different discourse area might well cover much of the same territory, contain much of the same theoretical apparatus, and so forth. (Consider, for example, the overlap

between the separate discourse areas of chemistry and biology, or geology and petroleum engineering.) The idea that discourse areas are entirely separate and distinct is an artificial view of how languages work, and of how theories relate to each other. Indeed, a discourse area that was completely isolated from the rest of our language—which had no explanatory links to the rest of our language, or to other well-established theories we hold—would not properly be regarded as thereby immune to criticism from other discourse areas, but would better be regarded with significant suspicion as an ad hoc theoretical appendage to the body of our world-view.

Notes

1. We emphasize that this is "a reading" of Kuhn because Kuhn himself spent his later years revising his views and in some cases distancing himself from many more radical interpretations of his approach: cf. Kuhn (1970/2000, 1977).
2. "I have not argued that there are universal properties. I have been concerned to establish something weaker… universal properties exist, and there exists a definite relationship between those universal properties, if there are any laws of nature" (Dretske 1977, 267).
3. Field (1980) even took a swing at disputing the necessity of numbers for physics.
4. This point is also discussed at greater length in Wolf (2008) and (2012).
5. We do not claim that these two methodologies exhaust the epistemological landscape or fill the logical space of epistemology.
6. Versions of this method are found in Brandom (his "default and challenge" model of entitlement), James (his "credit system"), and other pragmatists.
7. And since inquiry is carried on over generations, and by many people even within a single generation, rationality also becomes essentially *intersubjective* on the pragmatist viewpoint. This is also a crucial departure from the epistemological individualism of the Cartesian methodology. Appreciating the intersubjective nature of the epistemic is crucial to understanding knowledge, but we will not address this issue here.
8. This conception of knowledge—fallibilist, but not skeptical—has deep roots in the pragmatist tradition. Hilary Putnam goes so far as to claim,

"that one can be both fallibilistic and antiskeptical is perhaps the unique insight of American pragmatism" (Putnam 1994, 152).

9. "Doubting has certain characteristic manifestations, but they are only characteristic of it in particular circumstances. If someone said that he doubted the existence of his hands, kept looking at them from all sides, tried to make sure it wasn't 'all done by mirrors,' etc., we should not be sure whether we ought to call this doubting. We might describe his way of behaving as like the behaviour of doubt, but this game would be not be ours" (1969, §255).

10. We borrow the following examples from Loux (2001).

11. McDowell makes a similar claim about obligation: He notes that it is empty (or at best a placeholder) to say that you have reason to do A because you *should* do A. Rather, "the reason [one has to do something] must involve some appropriate specific consideration which could in principle be cited in support of the 'should' statement" (1978, 14).

12. As Jim O'Shea has reminded us, for someone committed to this sort of pragmatist approach to linguistic universals, talk of normative properties need not entail anything that violates the naturalist approaches we have described. We address this concern more fully in Chap. 4.

13. See Wolf (2002) for an account of some universals (particularly natural kind terms) that avoids some of the pitfalls of earlier inferentialist accounts.

14. It is disputable whether Sellars thought that meaning was constituted only by intralinguistic transitions, or also by language-entry and language-exit transitions. We favor the latter, in any case.

15. Many (perhaps most) contemporary philosophers will wish to analyze such normative sentences into a prescriptive element and a factual element. We think this move is mistaken, primarily because it results from the impulse to understand all languages from a representationalist paradigm; but we will have much more to say about this as we go along.

References

Armstrong, David. 1978. *Universals and Scientific Realism*. Cambridge: Cambridge University Press.

Balaguer, Mark. 1996. A Fictionalist Account of the Indispensible Applications of Mathematics. *Philosophical Studies* 83: 291–314.

Bedke, Matthew. 2012. Against Normative Naturalism. *Australasian Journal of Philosophy* 90(1): 111–129.

Bernstein, Richard J. 2010. *The Pragmatic Turn*. Cambridge: Polity Press.

Churchland, Paul. 1981. Eliminative Materialism and the Propositional Attitudes. *Journal of Philosophy* 78(2): 67–90.

Churchland, Paul. 1996. *The Engine of Reason, The Seat of the Soul: A Philosophical Journey Into the Brain*. Cambridge, MA: The MIT Press.

Dretske, Fred. 1977. Laws of Nature. *Philosophy of Science* 44(2): 248–268.

Field, Hartry. 1980. *Science Without Numbers: A Defence of Nominalism*. Princeton, NJ: Princeton University Press.

Field, Hartry. 1989. *Realism, Mathematics and Modality*. Oxford: Basil Blackwell.

Foucault, Michel. 1977. *Discipline and Punish*. Trans. Alan Sheridan. New York: Vintage Books.

Goldman, Alvin I. 1992. Epistemic Folkways and Scientific Epistemology. In *Liaisons: Philosophy Meets the Cognitive and Social Sciences*, ed. Alvin I. Goldman, 155–175. Cambridge, MA: The MIT Press.

Harman, Gilbert. 1977. *The Nature of Morality: An Introduction to Ethics*. Oxford: Oxford University Press.

Haugeland, John. 1982. Heidegger on Being a Person. *Noûs* 16(1): 15–26.

Heidegger, Martin. 1927/1962. *Being and Time*. Trans. John Macquarrie, and Edward Robinson. London: SCM Press.

Joyce, Richard. 2001. *The Myth of Morality*. Cambridge: Cambridge University Press.

Knobe, Joshua, and Shaun Nichols. (Eds.). 2007. An experimental philosophy manifsto. In *Experimental Philosophy*, 3–14. Oxford: Oxford University Press.

Kornblith, Hilary. 2002. *Knowledge and Its Place in Nature*. Oxford: Oxford University Press.

Korsgaard, Christine. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.

Kraut, Richard. 2011. *Against Absolute Goodness*. Oxford: Oxford University Press.

Kuhn, Thomas S. 1970. *The Structure of Scientific Revolutions*, 2nd ed. Chicago: University of Chicago Press.

Kuhn, Thomas S. 1977. Objectivity, Value Judgement and Theory Choice. In *The Essential Tension*, 320–329.

Lewis, David. 1983. New Work for a Theory of Universals. *Australasian Journal of Philosophy* 61: 343–377.

Lewis, David. 1986. Against Structural Universals. *Australasian Journal of Philosophy* 64: 25–46.

Loux, Michael. 2001. *Metaphysics: A Contemporary Introduction*, 2nd ed. New York: Routledge.

Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. New York, NY: Penguin Press.

Margolis, Joseph. 2003. *The Unraveling of Scientism*. Ithaca, NY: Cornell University Press.

McDowell, John. 1988. *Projection and Truth in Ethics (Lindsey Lectures)*. Lawrence, KS: University of Kansas Press.

Melnyk, Andrew. 2003. Some Evidence for Physicalism. In *Physicalism and Mental Causation*, ed. Sven Walter and Heinz-Dieter Heckmann, 155–172. Exeter: Imprint Academic.

Mill, John Stuart. 1869/1978. *On Liberty,* ed. Elizabeth Rapaport. Indianapolis: Hackett Publishing.

Millikan, Ruth Garrett. 1984. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, MA: The MIT Press.

Papineau, David. 2001. The Rise of Physicalism. In *Physicalism and Its Discontents*, ed. Carl Gillett and Barry M. Loewer. Cambridge: Cambridge University Press.

Papineau, David. 2009. The Causal Closure of the Physical and Naturalism. In *The Oxford Handbook of Philosophy of Mind*, 53–65. Oxford: Oxford University Press.

Parfit, Derek. 2011. *On What Matters*. Oxford: Oxford University Press.

Parfit, Derek. 2006. Normativity. Vol. 1, in *Oxford Studies in Metaethics*, 325–380. Oxford: Oxford University Press.

Putnam, Hilary. 1994. *Words and Life*. Cambridge, MA: Harvard University Press.

Quine, W.V.O. 1969. Epistemology Naturalized. In *Ontological Relativity and Other Essays*, 69–90. New York: Columbia University Press.

Quine, W.V.O. 1995. Naturalism; or, Living Within One's Means. *Dialectica* 49(2–4): 251–261.

Rorty, Richard. 1972. The World Well Lost. *The Journal of Philosophy* 69(19): 649–665.

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

Rorty, Richard. 1982. *Consequences of Pragmatism*. Minneapolis, MN: University of Minnesota Press.

Rorty, Richard. 1989. *Contingency, Irony and Solidarity*. Cambridge: Cambridge University Press.

Rosenberg, Alexander. 1999. Naturalistic Epistemology for Eliminativist Materialists. *Philosophy and Phenomenological Research* 59(2): 335–358.

Rosenberg, Alexander. 2014. Disenchanted Naturalism. In *Contemporary Philosophical Naturalism and Its Implications*, ed. B. Bashour and H. Muller, 17–36. London: Routledge.

Rovelli, Carlo. 2014. Science Is Not About Certainty. *The New Republic.* July 11. https://newrepublic.com/article/118655/theoretical-phyisicist-explains-why-science-not-about-certainty. Accessed 17 June 2015.

Scanlon, Thomas M. 2000. *What We Owe to Each Other*. Cambridge, MA: Belknap Press.

Sellars, Wilfrid. 1956/1997. *Empiricism and the Philosophy of Mind.* Cambridge, MA: Harvard University Press.

Sellars, Wilfrid. 1962. Philosophy and the Scientific Image of Man. In *Frontiers of Science and Philosophy*, ed. Robert Colodny, 35–78. Pittsburgh: University of Pittsburgh Press. Reprinted in Science, Perception, and Reality, pp. 1–40.

Staten, Henry. 2008. Derrida, Dennett, and the Ethico-Political Project of Naturalism. *Derrida Today* 5(1): 19–41.

Stich, Stephen. 1983. *From Folk Psychology to Cognitive Science: The Case Against Belief*. Cambridge, MA: The MIT Press.

Tooley, Michael. 1977. The Nature of Laws. *Canadian Journal of Philosophy* 7(4): 667–698.

Vicente, Augustin. 2005. On the Causal Completeness of Physics. *International Studies in the Philosophy of Science* 20(2): 149–171.

Warman, Matt. 2011. Stephen Hawking Tells Google 'Philosophy is Dead'. *The Telegraph*, May 17.

Wolf, Michael P. 2002. The Curious Role of Natural Kind Terms. *Pacific Philosophical Quarterly* 83: 81–101.

Wolf, Michael P. 2008. Language, Mind and World: Can't We All Just Get Along? *Metaphilosophy* 39(3): 363–380.

Wolf, Michael P. 2012. Boundaries, Reasons and Relativism. *Journal of Philosophical Research* 37: 205–220.

Yablo, Stephen. 2001. Go Figure: A Path Through Fictionalism. In *Midwest Studies in Philosophy (Volume XXV: Figurative Language),* eds. Peter French and Howard Wettstein, 25: 72–102.

2

Why Do We Need Normativity?

We are willing to grant that science has a place of authority when it comes to causal-explanatory projects. However, we share an equally strong commitment to the claim that human endeavor within this world cannot be understood without making room for the normative. Let us examine some of the reasons why we think the normative is ineliminable from our self-understanding, and of our understanding of the place of persons in the world.

2.1 Science, Instrumental and Non-Instrumental Normativity

First, those who want to give science pride of place among the sources of knowledge must also understand that scientific practice itself relies on various kinds of normative appraisals and cannot occur without them. One kind of appraisal that takes place is that scientific research is directed at questions whose answers are deemed *important*. This importance can be of various kinds. For example, answers to some scientific questions are instrumentally important, and these instrumental goals are determined

by various interests, needs, political arrangements, and so forth. Many of our most familiar products of today (such as duct tape, GPS devices, jet engines, microwave ovens, and superglue) were either developed by the military or from technologies developed by the military. Investigation into a better navigation system, a better way of sealing ammunition boxes, or a way of making fighters and bombers faster obviously involves a set of judgments setting various priorities and allocating resources based on goals that are taken to be important. Thus, the development of radar technology during WWII was made with specific purposes in mind (e.g., to better detect enemy aircraft), and this research embodied a practical commitment to the judgment that winning the war was important, that these means were appropriate and effective, and so on. When substantial resources were devoted to the development of a polio vaccine in the 1950s, this was based on the judgment that polio was a significant evil, and curbing the polio epidemic should be an important priority. These judgments of instrumental priority are often mistaken: an oft-repeated criticism of medical research funding is that research into diseases that claim fewer lives is often better-funded than research into diseases that are more deadly (e.g., cardiovascular disease). But in making such a criticism, one is doing just that: *criticizing*. One is making an assessment that one type of activity should be given a higher priority than another. Crucially, medical research is governed by well-known ethical standards that are in place largely to protect individuals from exploitation. Thus, while it might be efficient to use prisoners for medical experimentation, it would not be ethical to do so, and no ethical researcher would be allowed to do this. We should thus avoid any temptation to reduce the point about funding priorities to a question of mere efficiency.

Thus, scientific inquiry is guided by this sort of instrumental rationality. This instrumental rationality is still thoroughly normative, though (we will say more about this important issue shortly). But science is also guided by non-instrumental normativity. When Stephen Hawking endeavors to understand black holes, or the scientists at *Conseil Européen pour la Recherche Nucléaire* (CERN) endeavor to find the Higgs boson, or Andrei Linde endeavors to understand the early universe, they are not doing so to solve an immediate practical problem (such as winning a war or curing a disease). These questions are investigated because the

answers to them are deemed *important*—not instrumentally important, but important merely because of the understanding they provide: understanding of the universe, how it functions, and our place within it. The 21 states funding CERN could have decided to devote these massive resources to another question (say, a complete cataloging of the species of beetles in Europe), but their decision reflected a set of priorities, an understanding of which sorts of scientific questions are more important to investigate. We will discuss this issue of how questions of what is important, or worthy of investigation, essentially structure our investigation of the world at greater length in Chap. 9, where we discuss in more detail the ways in which normative and non-normative discourse areas contribute to each other in mutually fruitful (and ineliminable) ways.

2.2   Scientific Progress and Epistemic Appraisal

There is another way in which epistemic assessment is crucial for scientific progress. Let us approach this topic by discussing an important virtue of the successful scientist. Thomas Kuhn (1959) notes that we often think that successful science comes from "divergent thinking"—the ability to look at obvious facts without necessarily accepting them and the ability to actively imagine unique possibilities, instead of merely interpreting facts in the way taught in textbooks and by scientific authorities. Surely divergent thinking is important: without it, there would be no scientific revolutions. But very little of science is revolutionary; "normal" science is more akin to "puzzle solving" than revolution (1970, Chaps. 3 and 4). And in normal science, Kuhn emphasizes that "convergent" thinking is more useful than its divergent counterpart. Kuhn writes,

> [N]ormal research, even the best of it, is a highly convergent activity based firmly upon a settled consensus acquired from scientific education and reinforced by subsequent life in the profession. Typically, to be sure, this convergent or consensus-bound research ultimately results in revolution… But revolutionary shifts of a scientific tradition are relatively rare, and

extended periods of convergent research are the necessary preliminary to them. (1959, 227)

Kuhn argues that the history of virtually every scientific discipline, from physical optics to physiology, demonstrates that during times when research in a field is not characterized by convergent thinking (i.e., when researchers are confronted with a variety of different theories, allowed to examine the evidence for each, and choose among the theories on their merits), research in this field makes "very little progress" (1959, p. 231). Study of the pre-paradigm periods of these other disciplines strongly suggests that "without a firm consensus, this more flexible practice will not produce the pattern of rapid consequential scientific advance to which recent centuries have accustomed us" (1959, 232). Without a theory to which one is committed, one does not know what phenomena are significant, and what problems are worth solving (or perhaps indeed what the problems are in the first place). Further, one must be committed to a theory to undertake the serious work which is required to extend, deepen (*and ultimately overthrow*) the theory:

> Who, for example, would have developed the elaborate mathematical techniques required for the study of the effects of interplanetary attractions upon basic Keplerian orbits if he had not assumed that Newtonian dynamics, applied to the planets then known, would explain the last details of astronomical observation? But without this assurance, how would Neptune have been discovered and the list of planets changed? (Kuhn 1959, 235)

Appraisal along the epistemic dimension—the practice common in every field from chemistry to philosophy, of evaluating beliefs as justified or unjustified, scientific methods as rational or irrational, and so on, and of evaluating and revising not only those beliefs, but also the standards of evaluation themselves—is one of the tools that enables a consensus to arise and be maintained, and thereby speeds scientific progress. For a physicist working in 1800, it was rational to believe that Newtonian physics was correct. It was rational to carry out research projects using Newtonian techniques, and it would have been irrational to attempt to use Aristotelian physics to account for planetary motion. A physicist who

incorporated unjustified techniques (such as those of Aristotelian physics) would have been sanctioned by partial or complete exclusion from the profession of physics. Those who were adept at applying Newtonian techniques to the study of planetary motion were affirmed by inclusion.

Willem deVries (writing about how even our scientific engagement with the world is inherently normative) argues,

> One might think that the scientific image or framework is simply devoid of practical claims, because it aims only at description or explanation. But the scientific framework must also contain methods and canons, and these consist in part of prescriptive claims. There are ways experiments ought to be done and ways they ought not to be done; ways data ought to be interpreted and ways it should never be interpreted. There are inferences one may make and inferences one is forbidden to make. Every conceptual framework necessarily has a prescriptive or normative dimension. (2005, 272)

Jean Hampton makes a similar point. She notes that the pronouncements of scientists are only accorded (prima facie) authority to the extent that scientists have followed the relevant epistemic norms. Thus,

> we accept the results of experiments performed by, say, medical researchers only insofar as we are sure they accept the hypothetical imperative: 'If you want to generate evidence relevant to the truth of a certain hypothesis, you ought to construct double-blind experiments.' But here we have an ought of (practical) reason, one among many generated by a norm of (practical) reason on which science is based… And our acceptance of the results of scientific experiments is not only based on our belief that scientists have followed such imperatives, but also on our belief that they ought to have done so. (1998, 209)

Putnam, in arguing against a sharp fact/value dichotomy, notes that while every philosopher of science will acknowledge that criteria such as simplicity, coherence, and so on are relevant to theory choice in science, such criteria can only be understood as invoking certain values and making normative judgments. One does not make a neutral, merely factual statement in describing one theory as more simple or coherent or fruitful

than another. Such judgments are meant to be *action-guiding*; they are meant to guide theory choice (Cf. Putnam 1990, Chap. 9).

Epistemic evaluation can be discursive or practical. In other words, we can *call* a methodology irrational or a belief unjustified, or we can *carry out* our scientific practice in a way which favors this methodology, or which presupposes the falsity of that belief. And when our eighteenth-century professional physicists applied Newtonian techniques to the study of planetary motion, they were practically (though not discursively) endorsing these techniques. Similarly, the exclusion of the Aristotelian physicists from their ranks was a form of practical evaluation, a *repudiation* of these methods. This practical evaluation is crucial to scientific progress. As Kuhn pointed out, if we do not practically favor a particular paradigm over all others, science does not progress rapidly. Rapid scientific progress requires practical epistemic evaluation, which can help enforce uniformity of methodology—the crucial element in scientific advancement, according to Kuhn.

It might be objected that epistemic evaluation is not the sort of thing that can be implicit in scientific practice. Epistemic evaluation—for example, the assessment of techniques as rational or irrational—must be explicit (spoken, written, etc.). But evaluation—moral, epistemic, and other varieties—can surely be implicit in practice. For example, a person may firmly believe that lying is wrong, even if she never utters the sentence, "Lying is wrong." Her belief can be expressed by her constant truth-telling, especially in those cases where lying would clearly be more expedient for her. Alternately, her belief in the wrongness of lying could be manifested by her feelings of guilt when she does tell a lie. Similarly, Smith's belief that the dishwasher will wash his dishes is implicit in his actions (placing dirty dishes in the washer and turning it on) even if he does not utter the sentence, "My dishwasher will wash my dishes" and does not use the belief in a bit of explicit internal practical reasoning. We can evaluate or believe on particular occasions without saying anything, even *in foro interno*. Our evaluation and belief are internal in some cases, and implicit in others. In fact, the majority of our practice occurs at the level of the implicit. We only proceed to the level of explicit discourse when we need to inform, educate, or argue. So it is with epistemological evaluation. Our belief that double-blind studies are more reliable

than chicken bones is implicit in our decision to perform a double-blind study instead of visiting a practitioner of bone-reading. One of the chief insights to emerge from the pragmatist tradition is that normative language serves an *expressive* rather than a *causal-explanatory* role. That is, rather than serving to refer to natural facts, or explain causal regularities in behavior, normative language serves to make explicit norms which are implicit in actual practice (so that these norms may be subjected to rational scrutiny), or to endorse (or call for the revision of) such norms (which first requires that these norms be made explicit).

Indeed, we can only explain normative rules as a making explicit of norms which are already implicit in social practice. As Brandom writes, "Wittgenstein's pragmatism about norms—his insistence that norms made explicit in principles are intelligible only against a background of norms implicit in practice" (1994, 591) implies that we cannot explain the correctness of a performance by appeal to an explicit rule, because such a picture would then have to explain correctly following this rule by appeal to another explicit rule, leading to an infinite regress. Thus, the existence of explicitly stated normative principles presupposes the existence of normative proprieties which are implicit in practice. This picture of the normative will be drawn more firmly as the book progresses.

We have been arguing that science necessarily involves normative assessments about which types of questions are instrumentally or intrinsically valuable, assessments of which theories and methods are rational, and so on. But (goes one objection), cannot this point be accommodated by a sociology of science? Surely, scientists *take* themselves to be guided by certain norms, and to understand scientific practice, we must understand the various ways in which scientific practice embodies a commitment to this or that way of doing things (much as studying, say, a remote tribe involves uncovering the commitments they have to various rules, purportedly factual propositions, etc.). But (continues the objection), displaying these commitments in no way involves endorsing them, or committing ourselves to the truth of these commitments. Thus, how does the necessity of normative assessments within scientific practice underwrite the importance of the normative, as opposed to merely underwriting the importance of practitioners' *beliefs* about the normative?

We have two responses to this objection. First, while one can no doubt do the sociology of science, the above discussion demonstrates that the participants in a scientific inquiry *must* take themselves to be guided by norms in their scientific practice. These norms will privilege some questions as more important than others; these norms will privilege some methodologies as rational and justification-conferring; and these norms will privilege some theories over others. So while a sociology of science will certainly show that scientists are committed to certain norms, one cannot fully account for this normativity via a sociological explanation: commitment to norms of inquiry is not merely contingently, in fact, a part of scientific practice (much in the way that Oktoberfest is contingently, in fact, part of German culture). Commitment to norms of inquiry is a crucial, ineliminable part of scientific inquiry.

Second, that scientific inquiry is itself a goal worth pursuing is a normative judgment, and the scientist cannot explain why it is worth doing in the first place without making an evaluative judgment. Let us assume that science is guided by a norm of truth. (Surely, there are more norms than this—explanation, understanding, instrumental control over nature, political power, control of disease, etc. But let us make this simplifying assumption for now.) The point that must be emphasized is this: that a theoretical understanding of the world is even worth pursuing is a normative judgment. That scientific truth is *a goal we ought to pursue in the first place* is a normative judgment. It is not an unproblematic assumption at all. Thus, a scientist who eschews normativity altogether strips herself of the resources to say why it is that she does what she does in the first place.

Many philosophers have the intuition that the pursuit of truth has some kind of built-in necessity to it. (We suspect this is why fewer philosophers are skeptical of epistemic normativity than are skeptical of moral normativity.) Many would say, if one were not truth oriented, then one could not be an *agent* in the first place. Agency requires autonomy, the capacity to form beliefs, and so on; and these capacities can only be possessed by a truth-oriented being.

We may grant that one must be at least minimally truth-oriented in order to be an agent. But this in no way vindicates the scientific practices so valued by naturalists. A commitment to being truth oriented does not

entail a commitment to being *globally* truth-seeking. For example, while Calvin thought that humans could and should acquire certain worldly truths, and could and should acquire various religious truths by means of Scripture, he thought that human reason was so corrupted by sin that any attempt to go beyond Scripture and reason about the divine was misguided to the point of being sinful. More radical is the Boko Haram movement in Nigeria; the name translates roughly as "Western education is forbidden." Presumably, members of this group are truth oriented insofar as day-to-day living is concerned, but believe that certain methods of truth-seeking (in particular the ones most admired by scientifically oriented naturalists) are sinful.

Thus, one could have a "gappy" commitment to the truth. One could be truth oriented when it comes to acquiring food, shelter, and the like, but totally indifferent to truth (or even *actively opposed* to its pursuit) when it comes to forming beliefs about the shape of the earth, the author of *The Tempest*, the atomic weight of iridium, the color of the sky, the truth of evolutionary biology, the existence of black holes, and any other number of other types of beliefs. Only when we import normativity can we explain why anyone *ought* to care about truth (beyond what is required for agency). And this point includes scientific truth.

The clear implication of this discussion is that even if some minimum level of concern for the truth is a necessary condition on agency or otherwise rationally mandatory, scientific inquiry is itself a contestable value, and something we do because we are convinced that there exist sufficient reasons for doing so. But to say that the importance of scientific inquiry is contestable is not to say that it can or will be successfully contested. We hold that those who think it is not valuable are wrong. But the point is you cannot *say* that they are wrong without using normative discourse, and being committed to their wrongness is to undertake a particular kind of normative commitment.[1]

Jean Hampton makes a similar point, arguing that "any science-based argument against the idea of objective normative authority is self-refuting" (1998, 207). The idea that science should be taken as authoritative on factual questions is a normative stance one takes. The idea that science tells us our best account of the world is norm-free undermines the very

authority of science to make such pronouncements. As Hampton notes, a scientific story which denies the authority of the normative

> can explain why any person takes her methodology for understanding the world to be 'right,' but it cannot establish that one of them *is* right. And if this is so, science is only authoritative for those who accept it, and (as Feyerabend says) the enterprise of science as we know it cannot really undercut ethics, *because it has no objective authority to undercut anything*. (1998, 210)

Thus, we are in a position now to reemphasize a conclusion we argued for earlier: by eschewing normativity altogether, the advocate of scientific inquiry renders herself unable to advocate for scientific inquiry, or explain why it is valuable, or worthy of pursuit. The position is, ultimately, self-defeating. Thus, even scientific inquiry depends in a very intimate way on the normative. The normative is ineliminable.[2]

2.3   The Normativity of Instrumental Reason

Some philosophers seem to hold the view that the best route to naturalizing normativity runs through instrumental reason. The thought behind this strategy seems to be that instrumental reason represents an unproblematic kind of rationality, and that some strategy can be offered for naturalizing whatever normativity instrumental reason might have by grounding it in the ends pursued by this reasoning. Thus, suppose we take the goal of truth (even scientific truth) as given. Some would argue that given this goal, we can develop a naturalist account of normativity (employing purely means-end reasoning) to justify the various elements of scientific practice (as conducive to the end of truth, or to the multiple ends of truth, explanation, etc.). Or one might start with a purely descriptive account of desire and argue that the fulfillment of desire gives a naturalistic account of the normativity of instrumental reason. Implicit in this line of reasoning is that all reasoning is instrumental, all imperatives are hypothetical, and the normativity of these imperatives is given a naturalistic grounding in the pursuit of some unproblematic end (truth,

desire satisfaction, etc.). It might be thought that this offers a promising route to a naturalistic reduction of normativity: all normativity is instrumental, and the normativity of instrumental reason is all reductively generated by whatever naturalistically described end is pursued by this instrumental reason.

We reject this line of reasoning. We have argued that science presupposes not only means-end reasoning, but also reasoning about the value of final ends. Thus, *that scientific knowledge is valuable and worthy of pursuit* is itself a conclusion we must reach (either explicitly, or as a norm embodied in our practice) before we employ means-end reasoning to determine the best way of achieving scientific knowledge. But even if we bracket this point, it can still be established that instrumental reason is thoroughly normative, and this normativity is not simply generated by our desires or by some natural features of our biology. A naturalistically described end cannot in and of itself generate the normativity of instrumental reason, and the normativity of instrumental reason cannot be reductively accounted for in this way. Hence, normativity *tout court* cannot be reductively accounted for through this route.

2.3.1 Reductive Accounts of Instrumental Normativity

It is instructive to look at one of the most thorough presentations of the sort of view we are rejecting here. One exemplar is Alex Rosenberg's attempt to develop a purely naturalistic epistemology, grounded in the human organism's evolutionary goals. Rosenberg happily pleads guilty to the charge of "scientism," and argues that "epistemology must be continuous with psychology because philosophy is continuous with science" (1999, 335). Carrying on Quine's project in "Epistemology Naturalized," Rosenberg sees no problem in naturalizing epistemology because he thinks epistemology has no *specifically epistemic* normativity in the first place. It is "no more normative than, say, statistical methodology or engineering is" (1999, 336).

While there is no specifically epistemic normativity in his view, inquiry is guided by instrumental normativity. But Rosenberg is quick to claim that "[e]pistemology is prescriptive in only a relatively unproblematic

way. Its 'normativity' is exhausted by the prudential force of the imperative of instrumental rationality" (Rosenberg 1999, p. 337). He clearly identifies the goal served by instrumental rationality:

> QNE [Quinean naturalized epistemology] holds that there is at most only one intrinsic good, value or goal in nature: fitness maximization… That biological creatures are fitness maximizers is not explained functionally or teleologically. It is explained causally. If Darwin is right then in the end all our functional traits are shaped to attain this end, fitness maximization, and it in turn has no further end. That is what makes fitness maximization an intrinsic goal of organisms, in a purely naturalistic sense. Fitness maximization's status as an intrinsic goal of organisms enables us to grade the means they employ to attain it for efficiency—i.e. instrumental rationality. (1999, 337–8)

Rosenberg is clear that even if the goal of the organism is fitness maximization, "our cognitive economy is not [itself] directly designed to maximize fitness" (1999, 338). However, the distinction is one with barely a difference: he argues that our epistemic goals "are selected because they are conducive to fitness maximization [and] give epistemology its purely natural, prudent, instrumental normative force" (1999, 338). This rules out the formation of true beliefs as the goal of epistemology, as forming true beliefs will often not be conducive to fitness maximization (and sometimes the formation of false beliefs will be).

We find Rosenberg's account unsatisfying on several counts. One of our chief concerns is with Rosenberg's attempt to derive instrumental normativity from his account of our evolutionary goal. First, notice that Rosenberg explicitly eschews any functional or teleological account of biology. Rosenberg gives a purely causal account of our role as fitness maximizers. Presumably, fitness maximization *causally* explains why we possess the traits we do (and why other traits were not selected for) and so forth. But we cannot extract an account of prudential rationality out of this. Fitness maximization might explain (again, causally) why trait X was selected for, but the causal selection of a trait for a purpose does not generate a *prudential requirement* that we use the trait for this (and not another) purpose. More fundamentally, what is so special about the

causal selective story that privileges it over (say) what an agent desires or values? For example, a chief measure of biological fitness is the number of offspring one leaves behind; but how is it *prudent* for one to have as many children as one can support? If it is prudent to do so, it is so for extrinsic reasons (e.g., many children can support you in your old age), and not for any reason related to biological function. The inference from biological function to prudential imperative is simply a *non sequitur*. What if one's deepest aspiration is to a life of prayer and contemplation, or to a life of travel and adventure, and one finds such a life incompatible with having children? Or what if one simply does not *want* to have as many children as one can support? To say that it is prudent for us to pursue the goals evolution has selected for us, even if they conflict with our deepest values and desires, is to wave off all of our other goals and aspirations as irrelevant evolutionary epiphenomena, rather than actually accounting for them.

It is open to Rosenberg to concede that these evolutionary "aims" do not really generate any normativity (even of an instrumental, prudential sort). Thus, not only would there be no specifically epistemic normativity, but our cognitive lives would also not be governed by instrumental normativity, either. This would, in effect, make Rosenberg an eliminativist about epistemic normativity. But as we argued above, the cost of this move is high, particularly for a self-described advocate of scientism like Rosenberg. If we jettison normativity altogether, then we can no longer say what is distinctively important about doing science, or advocate for it. Of course, *we* agree that science is an important project, and we (as in the community of cognizers) *should* be in the business of advancing scientific inquiry. But one cannot *say* this without the language of normativity. Thus, Rosenberg's view would become self-defeating: he cannot advocate for the very scientific project that he says should be the basis for all further inquiry into the basic structure of the world and our place within the world.

There are deeper reasons for thinking that this strategy of trying to naturalize all reasons by reducing it to instrumental reason cannot succeed. First, it is plausible to think that the normativity of instrumental reason is no less objective than the normativity of non-instrumental reason. Second, it is also plausible to think that the normativity of instrumen-

tal reason presupposes an account of constitutive rationality, an account which itself involves a notion of non-instrumental normativity. Let us consider these reasons one at a time.

First, the normativity embodied in a hypothetical imperative is just as robust, and hence just as resistant to naturalization, as that embodied in a categorical imperative. Jean Hampton argues that "[h]owever contingent the hypothetical 'ought' is on a desire, it is still *not* the same as a desire" (Hampton 1998, 162); it has a normative force over and above the desire. As Hampton points out, there is a difference between merely *wanting* an end (in which case there is no irrationality in not willing the means), and *willing* an end (something only instrumentally rational agents can do). One who wills an end is normatively bound to will the means; failure to do so is to be practically irrational. It is difficult to see how this normative force could be captured in a purely reductive fashion. As Hampton notes,

> A Humean could, if he liked, admit the possibility of something like a psychological state of 'commitment' to an end [analogous to 'willing']…
> [B]ut he could not say that if this commitment failed—so that this psychological state somehow changed—he had in any way made a mistake. The charge that such a person made a mistake relies on a norm dictating the persistence of this state until the end is achieved. (1998, 164–165)

One who wills an end does not merely will the means as a matter of statistical regularity, or causal necessity; one does so as a matter of rational necessity. That is, one is bound by a norm of practical rationality to do so, and this norm is no less robustly *normative* than any norm of morality.

Second, we noted above that it is plausible to think that the normativity of instrumental reason presupposes an account of constitutive rationality. No agent can count as instrumentally rational in the absence of some substantive account of his or her own good. Suppose (to consider Korsgaard's example) that Howard, a man in his thirties, must have a course of injections or he will die soon. Howard does not want to die young, and but for his fear of injections, he would gladly submit to the treatment. But he truly is in horror of injections, and it is this horror of injections that motivates him to decline treatment (1997, 227). Korsgaard points out that the simple Humean, who holds that all rea-

son is instrumental, is faced with a series of dilemmas that show that instrumental reason cannot stand alone as an account of reason. On the one hand, we could go with Hume (who famously argues that "Tis not contrary to reason to prefer even my own acknowledg'd lesser good to my greater" (1740/1978, 416)) and claim that it is not a requirement of reason that we pursue our greater good. Thus, Howard's action is perfectly rational. But on this reading, it seems difficult to see how one could be *ir*rational. One can be mistaken, in that (for example) one's "passions can be provoked by non-existent objects" (1997, 228). But to make such a mistake is not, properly speaking, to be irrational. If there is no requirement to seek our greater good, then one cannot, properly speak, act irrationally. And if one cannot act irrationally, then instrumental rationality can impose no normative requirement on us. It is not an account of *rationality*, per se.

Thus, consider Bill, who says he really *wants* to lose weight (and you might think he really *ought* to lose weight) and then eats an entire gallon of ice cream in a single sitting. He has pursued a means (eating a gallon of ice cream) to an end (namely, immediate pleasure and gratification). He has only acted irrationally if you think he acted on the *wrong* end. But that is something you can only say if you have an account of constitutive rationality. Without this, an account of means-end rationality ceases to be an account of rationality and merely becomes a descriptive statement that a person who in fact has a particular end will in fact pursue a means to that end. All normativity or rationality will drop out of the account.

Thus, according to Korsgaard, the way to reinstate the instrumental principle as a principle of rationality is to distinguish between what a person actually wants and what she has a reason to want, or between what a person thinks she wants and what she really wants. But this just presents the simple Humean with another dilemma. Either of these options means going beyond mere instrumental rationality, and embracing reasoning about final ends. This should be relatively clear with the first option, which distinguishes between what a person actually wants and what she has reason to want. But it is also true with the second option, for it involves not only distinguishing between what a person thinks she wants and what she actually wants. It also involves the claim that "a person *ought* to pursue what he *really* wants rather than what he is in fact

going to pursue. That is, we will have to accord these 'real' desires some normative force" (Korsgaard 1997, 230).

Hampton makes a similar point, arguing that an account of instrumental reason cannot count as such without some account of the person's good, without some account of constitutive rationality. Consider the degree to which we should privilege the immediate over the distant future in the satisfaction of our preferences. Any theory of rational choice will say that this is permissible, but the question of how and to what degree we should do so turns out to be a theory about one's substantive good. Consider what Elster calls a "local maximizer," one who chooses actions based on how they will satisfy one's current preferences in the immediate future (without regard to how this action will affect one's preference satisfaction in the long-term). Local maximizers strike us as "lamentably stupid," in Hampton's words, because they are not "willing to invest" (1998, 181). That is, they are not willing to endure short-term costs (e.g., a trip to the dentist, or a college education) in exchange for even substantial long-term benefits. But again (and here we hear echoes of Korsgaard), this means that a rational agent cannot just act on whatever preferences she has, but must have some conception of her own good. That is, she must have some conception of what subset of her overall preferences is her good-defining set. And so a theory of instrumental reason requires at least "a very minimal substantive stand on the content of an agent's good-defining preferences, such that she can be declared rational" (1998, 181).

One possible reply due to Gauthier is that rational agents have second-order preferences (e.g., to maximize one's long-term utility rather than merely one's short-term utility). But as Hampton points out, such second-order preferences amount to a conception of the agent's good:

> Such a second-order preference is not like a preference for cheetos over corn chips; it's a preference that comes from reflection about how to understand and pursue her good… So here the second-order preference is being driven by a belief about what she *ought* to consider part of her *Source Set* of preferences if she is to act rationally, so that it is this belief, and not the second-order preference it generates, that is fundamental to her thinking about what she has reason to consider. (Hampton 1998, 197)

(The Source Set is a subset of our Total Preference Set and is defined by Hampton as the "Set of all preferences that are the source of our good-defining preferences" [Hampton 1998, 173, emphasis removed].)

2.3.2   Teleology and Instrumental Normativity

More sophisticated Humeans have built accounts that are supposed to be less susceptible to the sorts of criticism levied by Hampton and Korsgaard. For example, Alan H. Goldman (2009) tries to build an account of instrumental rationality on naturalist foundations. Goldman is not an orthodox Humean, in that he denies that reasons are constellations of beliefs and desires. Instead, reasons are states of affairs (such as a tennis racquet's being on sale being a reason to visit a shop). Nevertheless, he thinks that all reasons are constituted as such by our desires and concerns. For example, these states of affairs are not intrinsically reasons: if I did not enjoy tennis, or did not need a racquet, then this state of affairs would not constitute a reason for me. Thus, Goldman grounds all reasons in imperatives that are resolutely hypothetical: a state of affairs can only constitute an F reason for a person if one has the relevant set of concerns (if one is "F-minded"), where "F" can be moral, prudential, religious, aesthetic, or whatever.

For Goldman, isolated desires do not constitute states of affairs as reasons. Rather, "multiple coherent but sometimes conflicting sets of desires anchored by deeper concerns constitute as practical reasons various states of affairs that indicate how to satisfy those concerns" (2009, 108). Although Goldman does not specifically reply to Korsgaard, he seems to think that his more sophisticated Humean account can avoid the sorts of objections she has raised against more modest accounts of the instrumental principle. For example, Goldman argues for coherence and information requirements on our set of desires and concerns which constitute states of affairs as reasons. Darwall argues, however, that to require coherence is to require that we subscribe to a categorically valid norm of practical inference, one that is binding regardless of our desires and concerns. According to Darwall, we are "not automatically motivated to avoid self-defeat… [T]o be rational we must accept norms of

inference in both the doxastic and the practical spheres that take us from premises to logically or practically implied conclusions" (Goldman 2009, 59). Goldman rejects this contention. He argues that to be guided by considerations of coherence is intrinsic to the notion of being motivated to act (or believe) in the first place. He writes, "To be motivated is in part to be disposed to act on that motivation. We therefore do not need to accept an independent norm in order to act so as to avoid self-defeat or achieve practical coherence" (2009, 60).

All of this may give Goldman the tools to respond to another of Korsgaard's objections: for Korsgaard, the Humean account of instrumental reason failed to be an account of *reason* (i.e., an account of rationality) in large part because it failed to make prudence a rational requirement. Goldman asks us to distinguish between broad self-interest ("satisfaction of [my] informed and coherent desires, [my] rational concerns" [2009, 134]) and narrow self-interest, which "makes essential reference to oneself" (2009, 136). It is no requirement of reason that one act in one's narrow self-interest. One may put the welfare of another ahead of one's own welfare without behaving irrationally. But, Goldman argues, we are necessarily motivated to act in our broad self-interest, because

> we cannot be unmotivated to act on our own desires, since desiring is just being motivated. But this is a conceptual truth, hence trivial, not a requirement of rationality that one could fail to meet. Thus there is no point in claiming that we are rationally required to be concerned about our welfare in this sense, except to lend plausibility to a more controversial claim by equivocation (2009, 134).

Thus, we are necessarily motivated (and necessarily have reason) to act on our deepest coherent set of desires. But this is all just instrumental rationality; the requirement of coherence (as is noted above) just follows from the nature of motivation itself and the avoidance of defeat. It is not a separate, non-instrumental norm of rationality, according to Goldman.

Goldman offers a sophisticated account that is Humean in spirit. But we doubt that a reductive account such as Goldman's, which ultimately reduces all reasoning to instrumental reason, can succeed. Goldman, like many other reductive naturalists, is relying on a descriptive substrate

to generate normativity in a way that is ultimately unsatisfactory. For Goldman, states of affairs are constituted as reasons by our rational desires. He denies that values are out there in the world, independent of our desiring and valuing; the normativity comes from our desiring. But how? Some of the work is done by the rationality requirement: as noted above, an isolated desire does not generate a reason to act. Only our coherent, informed set of desires generates reasons for action. So rationality is part of the explanation, but rationality ultimately "reduces to the non-normative concepts of coherence and information" (2009, 35). But normativity is ultimately grounded in the aim or purpose of desiring. Goldman argues that the building blocks for his theory are naturalistic and free from objective values: he argues that an account of practical and epistemic normativity follows from the "natural aim" of desire and belief, which are purely psychological states:

> Rationality makes the normative demand that we follow them, but again, there is no non-naturalness here. Desires, like beliefs, aim at their own satisfaction, and their natural function is to prompt actions in accord with the reasons that indicate how to satisfy them… Given aims and the possibility of succeeding or failing in them, we have normativity; given natural aims, we have natural normativity. What determines what counts as a reason is a basic normative fact, but natural and of internal derivation, not irreducible and external. (2009, 183)

Thus, Goldman's strategy is to ground normativity in teleology (in this case, a teleological account of belief and desire). But this will not work, as Goldman hopes it will, as an attempt to ground the normative in the natural. For any discussion of teleology, and of the purposes and aims of organisms or their constituent parts or states, is unavoidably in the realm of the normative, not the purely descriptive. There is considerable debate in the philosophy of biology whether there is a type of natural normativity built into certain teleological features of organisms or species. We will not stake out a position on that question here, but we will note that even if there is such robust teleology in nature, it is not sufficient to reduce normativity in practical reason (or morality, or epistemology) to descriptive terms. This fact should be familiar to Goldman, who has

himself written extensively on sexual ethics. (Indeed, Goldman does not seem to think that other biological functionings necessarily create normative imperatives. He writes in "Plain Sex" (1977) that reproduction is sex's "primary biological function," but that "[w]hile this may be nature's purpose, it need not be ours" [1977, 271].) Many debates in sexual ethics founder on discussions of the "natural aim" of reproduction or the "natural function" of the sexual organs. These are questions not settled merely by figuring out which activities or purposes were favored by evolution, but are questions of our own moral values. As Goldman himself recognizes, natural teleology and value do come apart, and settling on the natural aim or function of an organ or biological component does not generate a normative imperative for us. Thus, when Goldman asserts that a psychological state has a "natural aim," that this aim is the same for everyone (and a priori so), and that this generates prudential imperatives for *everyone*, he is making what is (by his own 1970s-era lights) a *non sequitur*. Thus, even if we grant Goldman his other points above (that coherence represents an internal requirement on motivation and not an objective norm of action, that rationality can be reduced to nonnormative notions, etc.), he cannot ground an account of instrumental rationality purely in the natural.[3]

Ultimately, we align ourselves on this issue with McNaughton and Rawling, who claim that the normative is *sui generis* with respect to the natural. We can complicate the natural story as much as we like, but the normative "ought" will not be fully captured by whatever natural facts we have designated as relevant to the normative characterization of the situation. (This is not to abandon naturalism, as we will argue at the end of Chap. 7.) To use McNaughton and Rawling's example, "[S]uppose Eve has a headache, and Al has an aspirin that will relieve it. Because of this circumstance, Al has a reason to give Eve an aspirin: the circumstance is this reason (as distinct from the fact that Al has this reason)" (2003, 43).[4] Making one's account of the natural more sophisticated (e.g., giving a purely descriptive account of desire, but then requiring that one's desires be coherent, where one also gives a purely descriptive account of coherence) is not going to somehow bridge the is-ought gap, or make it disappear.

There are several lessons to be gleaned from the discussion so far in Sect. 2.3. First, not only is science ineliminably governed by normative methodology, and not only are the aims of science set by a process that is thoroughly normative, but even to accord science the authority to make pronouncements on what is and is not, is a normative stance; and in the absence of normativity, science cannot have this privileged position. Thus, the scientific argument for normative eliminativism is self-refuting. Second, any attempt to eliminate or reduce normativity by reconstructing all normativity in terms of instrumental reasoning will not succeed, because (a) the normativity of instrumental reason is as robustly objective as the normativity of any categorical imperative and (b) instrumental reason is not an account of *reason* unless it is accompanied by some account of constitutive rationality, some account of the good for persons. Thus, it is not possible to reconstruct all normativity in terms of instrumental reason. Normativity, both instrumental and non-instrumental, is ineliminable.

2.3.3   Normativity and Agency

Alex Rosenberg writes,

> [Since] the brain cannot be the locus of original intentionality, then original intentionality just doesn't exist. But without intentionality, we have to recognize that most of our conceptions about ourselves are also illusions. If plans, projects, purposes, plots, stories, narratives, and the other ways we organizes our lives and explain ourselves to others and ourselves all require intentionality, then they too are all illusions. (2014, 28)

As noted above, Rosenberg gladly pleads guilty to the charge of "scientism" and is happy to jettison the manifest image and all that it contains. For him, "Naturalism… bids us doubt that there are facts about reality that science cannot grasp" (2014, 17). A thoroughgoing naturalist like Rosenberg would argue that we can understand persons and their engagement with the world, without appeal to the language of values, norms, or intentions.

We disagree. To start small, consider the example of a hammer, as discussed by Willem deVries (2005, 275): no purely physical description of a hammer would be complete. First, it would be "wildly disjunctive." More importantly, it would not capture the most important element of its subject. A hammer is not just something that can be used to pound nails. (One can use a rock, or a metal ingot for that.) A hammer is something that *is to be used* for pounding nails. Prescriptivity bound up in its very identity, with our engagement with it, and indeed, with our very understanding of what the object is in the first place.[5] Someone who does not understand these normative proprieties—what doors are for, or windows, or cars, or pens or paper, or books, or any other artifacts whose identity is in large part determined by their appropriate use—*does not understand the world of persons*, the world in which we live and interact. More fundamentally, our interaction with each other presupposes not just that we are complex physical systems, but that we are agents who are responsive to reasons, and bound by norms. An agent who was fundamentally unresponsive to reasons would cease, in our eyes, to count as an agent at all. We can surely imagine a human acting on instinct (eating when hungry, fighting, or fleeing when afraid, etc.) but a human who did not respond to reasons *qua* reasons (e.g., who did not, for example, see something's being dangerous as a *reason* not to do it, or something's being poisonous as a *reason* not to eat it, or something's being pleasant or kind as a *reason* to do it) would not be an agent at all. The description of such a person sounds more like a feral being than an agent.

Our engagement with others *qua* agents presupposes not just that they are in the space of reasons, but that they manifest certain intentional states: they *intend* to meet us for dinner tomorrow at 6, just as they *promised*; they *believe* that Jesus died for their sins, which is why they attend church every Sunday; they *expect* the newspaper to be on their front porch this morning; and so on. And these intentional states (including mental states) are themselves normative in character.

Our commitment to the ineliminability of the normative leaves us with an important promissory note. It is one thing to argue that the normative is ineliminable, and that the science has a kind of priority in explaining the world and what it contains. But can these two positions be

reconciled? This was the key question confronting Sellars in "Philosophy and the Scientific Image of Man." Sellars argued for a "synoptic vision," in which (while retaining the primacy of the scientific image) the manifest image containing persons, norms, and the other conceptual apparatus of agency is fused to the scientific image. An important goal of this work is to show how we can give a robustly objective account of the normative without ruffling any naturalist feathers.

Notes

1. One might try to evade this objection by arguing that in advocating for the use of science, one is merely making an instrumental claim that *if* one wants to acquire truths, then one should privilege scientific methods and canons. However, this reply is caught on the horns of a dilemma. If one is merely making an instrumental claim, then one is not in fact arguing that scientific methods and canons ought to be privileged. One is merely saying that if one wants to acquire truth in certain fields of inquiry, then one should privilege the methods and canons of science; to say, this is not to argue that one should afford science any kind of privilege. (And in the next section, we will argue that these sorts of "hypothetical claims" are in any case normative, so that even such a minimal instrumental claim cannot be made without a commitment to normativity.) On the other horn of the dilemma, if one claims that one ought to pursue truth in these disciplines (and so one therefore ought to privilege the methods and canons of science), then obviously one is committed to both constitutive normativity (normativity about ends) and instrumental normativity.

2. The problem for the eliminativist may go even further. As Sellars (1957, 283) argued, "It is a logical truth that [a description of the world], however many modal expressions might properly be used in *arriving* at it, or in *justifying* it, or in showing the *relevance* of one of its components to another, could *contain* no modal expressions." And these are the modal expressions needed in the formulation of scientific laws, that are needed to describe the behavior or ordinary goods in our folk vocabulary ("If this match is struck, it will light")—indeed, Sellars argues that understanding such counterfactuals is part and parcel of understanding what such ordinary goods *are*, of understanding talk about them. Any attempt to ana-

lyze counterfactuals or modals linguistically or metalinguistically will run afoul of the problem of the ineliminably normative nature of the semantic, an issue we will touch upon in Chap. 3. But we will not pursue the issue of modality any further here.

3. Much of our thinking on this point was spurred by helpful comments from Lucas Thorpe.

4. We are not entirely satisfied with McNaughton and Rawling's account. One of the chief purposes of their essay is to reconcile the *sui generis* nature of the normative with respect to the natural with the supervenience of the normative on the natural. But in our view, their account does nothing to alleviate our deeper concern, namely, that such accounts do not explain how the natural is supposed to account for this normativity.

5. The most famous usage of this example is probably due to Heidegger, who intended precisely the same point: objects like hammers exhibit "readiness-to-hand," meaning that they have a practical significance in virtue of being caught up in a system of practices, which are themselves *practices* in virtue of the norms they embody. Thus, Axel Honneth writes that, "according to Heidegger, we do not encounter reality in the stance of a cognitive subject, but rather we practically cope with the world in such a way that it is given to us as a field of practical significance" (2008, 30).

References

deVries, W.A. 2005. *Wilfrid Sellars*. Montreal: McGill-Queens University Press.

Goldman, Alan H. 2009. *Reasons from Within: Desires and Values*. Oxford: Oxford University Press.

Hampton, Jean E. 1998. *The Authority of Reason*. Cambridge: Cambridge University Press.

Korsgaard, Christine. 1997. The Normativity of Instrumental Reason. In *Ethics and Practical Reason*, ed. G. Cullity and B. Gaut, 215–254. Oxford: Clarendon Press.

Kuhn, Thomas S. 1959. The Essential Tension. In *The Third University of Utah Research Conference on the Identification of Scientific Talent*, 162–174. Salt

Lake City: University of Utah Press. Reprinted in Kuhn (1977), pp. 225–239.

Putnam, Hilary. 1990. *Realism with a Human Face*. Cambridge, MA: Harvard University Press.

Rosenberg, Alexander. 1999. Naturalistic Epistemology for Eliminativist Materialists. *Philosophy and Phenomenological Research* 59(2): 335–358.

3

Against Supervenience and Reductionist Accounts of Normativity

In Chap. 2, we argued that no form of naturalism could do away with normativity altogether. In this chapter, we address a large swath of the naturalist literature that attempts to *place* properties of normative discourse within the terrain of entities posited by the natural sciences. There are numerous strategies to place them indirectly in the natural world by showing that they reduce to or supervene upon reputable categories of properties and entities from other theories. For instance, they might reduce to or supervene upon our attitudes, which we might take to be some subset of our psychological states; those might then be shown to reduce to or supervene on some further physical substrate. An implicit assumption in any placement strategy—reductionism or non-reductive supervenience—is that it offers us an *explanatory* account. To identify a reduction or supervenience base successfully would be to specify some range of objects and properties that would explain why there are just the true normative claims that there are, and why they have the particular sort of grip on us that they do.

As the previous two chapters might suggest, we take a pessimistic view of the prospects for placing normative properties in the natural world. For now, our approach will be to assess the prospects for more conventional ways of dealing with normative properties for naturalists: reductionism

and non-reductive supervenience. Reductionism itself introduces a kind of supervenience relation: if A-properties reduce to B-properties, then there can be no change in A-properties without a change in B-properties. For brevity's sake, we will generally refer to all non-reductionist supervenience accounts as "supervenience accounts" and treat the reductionistic ones as a separate category. Our strategies for addressing them will be quite different. First, we show that reductionist approaches to normative discourse fail to do justice to normative discourse in numerous ways. We show this to be a problem in principle, not merely a failure of some accounts. Second, we show that non-reductive supervenience accounts either (a) collapse back into reductionist accounts or (b) fail to explain normative discourse in the ways necessary for a placement strategy. We also take this second line to be a problem in principle for non-reductive supervenience, rather than a problem for particular accounts. This is somewhat different from past debates about the "reduction or mystery" status of supervenience accounts: we do not think that non-reductive supervenience accounts have to devolve into mystery. But we will argue that for those not embracing reductionism for the normative, their means for avoiding it do not suit a placement strategy—they will either presume (on any plausible reading) some of the normativity that they purport to explain, or else fail to explain it at all. Our aims in this chapter will thus be wholly negative, with an eye toward later rejecting the call to place normativity in the natural world at all.

We should also say in passing here that much of this chapter will also provisionally entertain many ways of thinking about properties, sentences, and normativity that depart significantly from those we ultimately endorse. For those picking up on some of the Sellarsian and pragmatist themes in Chaps. 1 and 2 and wondering where they have gone, bear with us. They will be back.

3.1   A Quick Stage-Setting Note: "Mere Aggregate" Properties

The term "property" is treated very differently by different philosophers. Perhaps the weakest possible view of what properties are treats them as

merely sets of things, or sets of sets of things in different possible worlds. There is a long nominalist tradition in this vein, defended in recent decades by nominalists such as Lewis (1986a, b) and Rodriguez-Pereya (2002). In part, such strategies have been devised to avoid making the introduction of a property a matter of commitment to abstract objects or real universals, only extensions. For our present purposes, call these *mere aggregate properties*.

If a property is nothing more than sets of sets of things aggregated together, then there are no abstract entities to countenance, and the nominalist itch might be scratched. Commitment to mere aggregate properties would thus be minimal, perhaps as weak an ontological commitment as we could make assuming we accepted particulars at all. Many possible aggregates would be arbitrary or monstrous to us, but not all need be. Some will exhibit interesting commonalities for theoretical purposes (e.g., the positively charged particles, the mammals, etc.), and these will be a focus of different forms of inquiry. The pertinent question for a placement strategy is whether there is some real quality that recurs in the members of some of the aggregates, uniting them in a more substantial way. It is that sort of additional ontological commitment—to a further, more robust notion of what a property is, over and above an aggregate of particulars—that concerns us when it comes to placing the normative in the natural world. Some naturalists take normative properties and the normativity of our judgments and actions to be *real*, expressing support for something more than mere aggregate properties. Such properties would then need to be placed in the natural world, either by reduction or non-reductive supervenience.

We note this distinction with the thought that someone might argue that normative properties can be placed in the natural world by selecting some subset of the mere aggregate properties. For instance, we could imagine looking at the complete history of the actual world (with godlike perception, omniscience, etc.) and picking out some set of events that we could then call "morally right actions." We could then do the same for other possible worlds, bundle up the sets into a class, and have a normative property. As many have noted, this would leave us with a hideously gerrymandered class across possible worlds from other theoretical perspectives. But if we only commit ourselves

to mere aggregate properties in general, then this is not an *ontological* problem. We have added nothing to the world, even if we still have lots of work to do.

Any attempt to account for all properties as mere aggregates faces substantial technical challenges. More pertinently for present purposes, they offer too thin a basis for any claim to have placed normative properties in the natural world. One could argue that there are mere aggregate properties that capture all the concrete things that figure in our thinking about normativity. But even if we concede that, we might still make the case that the pertinent question is whether we need some more-than-mere-aggregate properties to underwrite normative discourse. While all the members in the aggregates might be suitable for naturalistic accounts, the real work for any such approach would be to select *which* aggregates were the normative properties. Presumably the selection would either have to appeal to a regularity describable in some theoretical terms (in which case, it is no *mere* aggregate), or else make the selection of the aggregate brute and arbitrary (in which case, it is not explanatory and not a suitable placement strategy). When we challenge reductionist and supervenience accounts of normative properties in this chapter, we challenge the view that there are more-than-mere-aggregate normative properties that either sort of account could point us toward. If placement succeeds, then normative properties (whatever they may be) should explain the features of normative discourse as features of the natural world, and mere aggregate properties do not explain. So we set them aside for now and turn our attention to reductionist and supervenience accounts for the next two sections.

3.2   Reductionism[1]

One purported method of resolving the placement problem described above would be reductionism. Such an account would try to achieve what Darwall et al. (1992, 126) term a "substantive assimilation" of normative discourse to science. Thus, a moral property such as *morally right* might be claimed to be equivalent to the natural property *maximizes utility*; an

epistemic property such as *epistemically justified* might be claimed equivalent to the natural property *formed by a reliable belief-causing mechanism*; and so forth. These natural properties may be very complex properties: For example, *maximizes utility* may involve many psychological and other physical states. But they can be investigated by the empirical sciences, they can legitimately be used in causal explanations, and so forth. Normative properties, on this line of thought, can thereby inherit their legitimacy from these natural properties.

For naturalists of this stripe, reduction is accomplished "in the form of…a synthetic identity statement" (Darwall et al. 1992, 174) via a statement about the co-extensiveness of moral (or epistemic) and natural terms. James Rachels writes that the ethical naturalist will "say that moral properties (such as goodness or rightness) are identical with 'natural' properties, that is, properties that figure into scientific descriptions or explanations of things" (2000, 75).[2] A similar move can be made with respect to other types of normative discourse. Thus, Railton writes that "the naturalist in ethics can be a naturalist in epistemology" (1993, 281). For example, evolutionary and moral psychologists have made efforts to explain normative judgment in terms of the exercise of capacities that are conducive to an organism's reproductive fitness.[3]

For many naturalists, failure to reduce normative properties to natural ones casts into doubt the prospects for success of these normative projects. Thus, Harman argues that "there is no way to test moral claims empirically, unless they are reducible to naturalistic claims" (1986, 58). Harman is famous for arguing that to be legitimate, moral claims must be subject to empirical testing. Sturgeon, discussing the idea that there are irreducibly moral truths, argues that "this retreat [would] certainly make it more difficult to fit moral knowledge into anything like a causal theory of knowledge, which seems plausible for many other cases, or to follow Hilary Putnam's suggestion that we 'apply a generally causal account of reference…to moral terms'" (1988, 236–236).

G.E. Moore (1903) famously doubted the coherence of identifying the normative with, or reducing it to, the natural. In its simplest form, Moore's Open Question argument states that for any natural property (such as *being desired* or *causing pleasure*) and any normative property (such as *good*), one can, without contradiction, question whether some-

thing that possesses that natural property also possesses the normative property. Thus, according to Moore, these properties cannot be identical. Sellars, who was a champion of the idea that normative concepts have a normative conceptual surplus which defies outright naturalistic reduction, writes:

> Now the idea that epistemic facts can be analyzed without remainder—even "in principle"—into non-epistemic facts, whether phenomenal or behavioral, public or private, with no matter how lavish a sprinkling of subjunctives and hypotheticals is, I believe, a radical mistake—a mistake of a piece with the so-called "naturalistic fallacy" in ethics (1956/1997, 19/§5).

Discussing ethical naturalism, Sellars (1953) considers the following equivalence:

x is right = x maximizes happiness

Sellars argues that this equivalence cannot be the whole story: whatever "x is morally right" might *refer* to, it *says* that x is morally right, and that cannot be said using purely descriptive vocabulary. As Sellars writes, "Whatever users of normative discourse may be *conveying* about themselves and their community when they use normative discourse, what they are *saying* cannot be said without using normative discourse" (1953, 82).

Contemporary reductionists respond to this worry by distinguishing between the co-extensiveness of normative and non-normative terms and their synonymy. A phrase such as "morally good" might be extensionally equivalent with "maximizes utility," but there is a normative surplus that is not a feature of the extension of the phrase "morally good," but is instead a feature of the phrase's meaning. Thus, the well-known distinction between meaning and reference is pressed into service for the reductionist agenda. This approach is reflected in the earlier quote from Rachels, who asserts the identity of moral and natural *properties*, and the quote from Darwall, Gibbard, and Railton, who speak of *synthetic* (as opposed to analytic) identity statements. This specific

move is particularly well-developed in the ethics literature: again and again, we see this strategy and motivation. The naturalist begins with the worry that moral and naturalistic concepts play a different "cognitive role" that moral concepts have a normative content which naturalistic concepts lack. But, the naturalist argues that this fact only rules out identifying moral and natural concepts, while allowing identification of moral and natural properties. We find this move being made by Brink (1989, 144–170), Fenske (1997, 301–304), Rachels (2000, 75), Wedgwood (1999, 206), and others. On these theories, goodness turns out to be maximization of utility, a particular homeostatic property cluster,[4] or some similar naturalistically specifiable property or entity. These naturalistic philosophers often then take advantage of direct theories of reference pioneered by Kripke and Putnam to assimilate moral kinds to natural kinds, and thus bring moral entities under the causal theory of reference.[5] Fenske has even gone so far as to say that "Kripke-Putnam semantics renders Moore's open-question argument obsolete" (1997, 301).

A similar approach can be seen in the work of various naturalized epistemologists. Hilary Kornblith, for example, writes, "The goal of a naturalistic theory of knowledge, as I see it, is not to provide an account of our concept of knowledge, but instead to provide an account of a certain natural phenomenon, namely, knowledge itself" (1999, 161). Jaegwon Kim claims that while "justification is a normative *concept*," it is still the case that "*criteria* of justification…must be stated in descriptive terms" (1988, 383, 397 emphasis added). Kim draws an explicit parallel between naturalized epistemology and naturalistic ethics, writing that

> Normative ethics can serve as a useful model here. To claim that basic ethical terms, like "good" and "right", are definable on the basis of descriptive or naturalistic terms is one thing; to insist that it is the business of normative ethics to provide conditions or criteria for "good" and "right" in descriptive or naturalistic terms is another. One may properly reject the former, the so-called "ethical naturalism", as many moral philosophers have done, and hold the latter; there is no obvious inconsistency here. (1988, 397–398)

So in reducing the normative to the natural, reductionists are reducing only the extension of normative terms. On top of the extension of terms such as "right," "justified," and so on, there is a normative surplus which is not reduced. This surplus is supposed to be a feature of the normative terms' meaning, not their extension. So epistemologists and moral philosophers hand off this normative conceptual surplus to philosophers of language, to be dealt with by a theory of meaning.

We doubt that a theory of meaning can cope with this normative conceptual surplus while remaining true to the motivation that drove philosophers in the first place to restrict their reduction to a reduction of extension, not of meaning. The same arguments that lead philosophers to reject the notion that moral or epistemic terms could be *synonymous* with natural terms seem to *forbid* a reductionist account of meaning, whereas the motivation that drove philosophers to seek reductive theories of morality and epistemology seems to *require* a reductionist account of meaning. Reductionism is torn between two incompatible requirements.

3.2.1   Normativity and the Meaning of Normative Expressions

Like most philosophers, we accept the view that meaning is normative. Someone who deviates from the proper meaning in the use of an expression has violated a norm: she has done something *wrong*, something *incorrect*. This sort of normativity is often understood as a point about the possibility of error. To misuse a word is to have done something wrong, whereas someone who merely deviates from a disposition, or from the most common behavior in their community, does not necessarily commit an error. Wittgenstein (1953) and Kripke (1982) discussed this point extensively, and it is widely accepted by contemporary philosophers of language.

This would imply that the meaning of an overtly normative piece of discourse (e.g., a moral or epistemological claim) has a *doubly* normative character. Consider:

M: "X is morally right for S" means (in part) "S ought to do X"

"A means B" is normative in the way that all meaning claims are normative; in virtue of containing the term "means," it indicates how the term A ought to be used. But a claim like M has a second normative element. The moral concept has a normative surplus, the one that we could not reduce to some natural fact. It expresses (via the phrase embedded on the right) the normative component of the moral claim "X is morally right for S." We might call these M's *semantic normativity* and its *moral normative content*, respectively. (Sentences about the meaning of normative *epistemic* terms would involve *epistemic normative content* and so on for other types of normative discourse.)

In our view, this double normativity of sentences like M presents an intractable set of problems for those who pursue reductionist strategies. How would we pursue a naturalistic reduction of the moral normative content of M? To assert M is to claim that "X is morally right for S" means (in part) "S ought to do X." What would be a suitable candidate for the meaning of "S ought to do X"? That question resists any readily available reductionist answer. The reductionist has to deny that "X is morally right" *could* mean the same thing as "X maximizes utility"; to assert this is just to commit the naturalistic fallacy and to fail to articulate the sentence's normative content. The same line of reasoning will forbid a reductive account of what it is for a sentence to *mean* that S ought to do X. Suppose that our reductionist endorses a simple dispositional theory of meaning and claims that M means that, in the long run, the majority of people would be disposed to condemn S for failing to do X. But whatever that description of what the majority will do in the long run may convey, it does not *say* what "S ought to do X" does. For instance, we regularly note patterns in people's dispositions that run counter to what they ought to do, and there is clearly no contradiction in saying this. Where we want to articulate the moral (epistemic, etc.) normative content a piece of normative discourse, a description of circumstances—even one with a counterfactual element—simply will not do.

This is precisely the sort of reasoning that led the reductionists we mentioned earlier to concede that their strategies could only equate the extension of normative and descriptive terms, not their meaning. If the normative content of the moral expressions M cannot be reduced to a set of natural facts (which was the central concession that led to a splitting

of questions of the extension and meaning of the expressions), then the philosopher of language will be in no better position than the moral philosopher in reducing normative content to natural facts and properties. (Nor for epistemic claims and epistemic normativity, etc.) The normative conceptual surplus that could not be reduced, due to the *sui generis* nature of the normative with respect to the natural, was then pushed off onto the theory of meaning. But the same *sui generis* status stands in the way of a naturalistic reduction of this same content when it is handled by our theory of meaning. Note that we are not yet denying that a dispositionalist account of meaning will work for the meaning of non-normative sentences, though as our readers might imagine, we probably would. All we are denying at this point is that such a reduction strategy will work with expressions in normative discourse like M. Such a reduction strategy either commits the naturalistic fallacy or leaves the most important features of normative discourse unexplained and unplaced in the natural world.

Thus, it seems that the reductionist must tolerate a non-reductive theory of meaning. But this undercuts the entire motivation for adopting reductionism in the first place. That approach presumed that only expressions designating natural properties are legitimate elements of our best accounts, and so if normative vocabulary is to be respectable, it must be co-extensive with (i.e., reducible to) naturalistic vocabulary. But if the reductionist concedes that semantic vocabulary is legitimate but irreducible, then the objections to the normative dimensions of morality and epistemology are no longer convincing.

There is no reason why, if the reducibility of moral and epistemic claims ("Murder is wrong," "S knows that P," etc.) is a condition on their legitimacy, the reducibility of semantic claims ("A means B") should not also be a condition on the legitimacy of these claims. Indeed, the arguments in favor of reductionism in morality and epistemology cited at the beginning of this section can just as easily be rephrased to encompass semantics. What Harman said earlier applies again, *mutatis mutandis*: "There is no way to test [*semantic*] claims empirically, unless they are reducible to naturalistic claims." And so for Sturgeon's earlier claim: "If we concede that there are [*irreducibly semantic*] truths, this retreat would certainly make it more difficult to fit [semantic] knowledge into anything

like a causal theory of knowledge, or… 'apply a generally causal account of reference…to [semantic] terms.'" If these are conditions that a reductionist account must meet, it seems that reductionism cannot meet its own demands when it comes to normativity.

Thus, moral (or epistemic, etc.) normative content presents a barrier to a reductive account of the meaning of normative claims. But matters get even worse when we look more closely at the semantic normativity of meaning claims. For instance, Eric H. Gampel argues that meaning is normative in that it is *essential* to meaning to be able to justify use. He calls this the Essential Justificatory Role of Meaning (EJRM). But a serious problem for reductionist theories of meaning arises from this. "The general problem," Gampel argues, is that,

> While it seems essential to a rule, or to a fact about a rule, that it have a capacity to justify, it does not seem essential to natural objects or facts to be able to do so…Physical facts about the meter bar certainly could justify various measurement claims, but it was not essential to them to be able to do so…So EJRM, if right, would provide the basis for a *prima facie* case against identificatory reductions of meaning facts. (1997, 231-232)

So if a fact is to be a *meaning* fact, it must be essential to this fact that it be able to justify. No natural fact *essentially* possesses this justificatory efficacy. Thus, semantic normativity stands in the way of any reductive account of meaning.

As we noted above, the thesis that meaning is normative is often cast in terms of the possibility of error. But it is on this point that naturalized semantic accounts tend to fail, and their failure on this count is *systematic*. Barry Loewer writes,

> None of the naturalization proposals currently on offer are successful. We have seen a pattern to their failure. Theories that are clearly naturalistic… fail to account for essential features of semantic properties, especially the possibility of error and the fine-grainedness of content. (1997, 121).

Paul Boghossian writes "Reductionist versions of [semantic] realism appear to be false…Meaning properties appear to be neither eliminable,

nor reducible" (1989, 547–548). Peter Godfrey-Smith (1989) also reluctantly concludes that the problem of error may be insurmountable for naturalized semantics. Thus, attention to the possibility of error lies at the heart of semantic normativity and also represents a major stumbling block to naturalized semantic theories.

Of course, the solution to the vexing problem of normativity pursued by the moral theorist and the epistemologist was to argue that they were only reducing the factual element of moral and epistemic claims, the normative element they were pushing off on the semantic theory. But a similar move is not available to the would-be reductionist in semantics: you cannot reduce the factual element of a semantic claim and then push the normative element off onto your semantic theory, for it is precisely our semantic theory we are trying to reduce! Normativity proves to be a troublesome bump in the rug, but the reductionist can only move the bump around without ever smoothing it out. Even if one denies that semantic normativity is a barrier to a reductive account of meaning (indeed, even if one denies that there *is* semantic normativity), the thesis of this section still remains intact, since moral and epistemic normative contents remain a barrier to a fully reductive theory of the meaning of normative expressions.

3.2.2   Non-Semantic Normative Content

It might be argued that the normative surplus, which was not reduced, is not a feature of the moral term's meaning, but is accounted for in some other way. Suppose we claim that the phrase "morally good" refers to the maximization of utility. This would be to identify moral goodness with the maximization of utility via a synthetic identity statement. We might then claim that the normativity of moral goodness is located somewhere other than meaning of the phrase "morally good." Instead, let us imagine that the reductionist claims that the normativity of moral goodness consists in the ability of moral goodness to motivate perceivers of the good. (No reductionist we know of endorses this particular alternative, but it is simple enough to serve as an illustration for the moment.) On this proposal, *goodness*—which is co-extensive with the natural property of *util-*

ity maximization—has a certain "to be doneness" which causes normal moral agents to act in order to achieve the good. This would eliminate the problem with which we began Sect. 3.2: it would be open to the reductionist (though not mandatory) to say that moral and epistemic terms are synonymous with natural terms. After all, if there is nothing normative to the meaning of moral and epistemic terms, then there seems to be no problem with saying that such terms are synonymous with natural terms or phrases such as "maximizes utility" or "produced by a reliable belief-forming mechanism."

There are, however, good reasons for rejecting this move on the part of the reductionist. On this view, the normativity of moral and epistemic terms (like "good," "right," "rational," "justified," and so on) is not a feature of these terms' meaning. Thus, to say, "Torturing cats is immoral" or "Belief in astrology is irrational" is not to make any sort of normative or evaluative claim. It is to categorize those action types in some way, but not one that has any evaluative or action-guiding import. Again, though, it seems prima facie absurd to deny that terms of moral and epistemic appraisal have normative meaning. (Certainly, the reductionist seems to have been assuming that they do; otherwise, why go to all the trouble to distinguish between meaning and extension in order to respond to Moore's Open Question Argument and avoid the naturalistic fallacy?) Moral and epistemic claims clearly *mean* something normative, which distinguishes them from non-normative claims.

Second, by identifying the "to-be-doneness" of normativity with some natural property, this solution runs directly afoul of Mackie's queerness objection. If a natural fact is intrinsically motivating, or reason-giving, or in some other way normative, then we are countenancing the existence of "queer" properties or facts. Analogous problems will emerge when we consider epistemic properties. Some facts or properties will have an intrinsic to-be-believedness—one that compelled belief, regardless of the agent's collateral commitments. But the natural sciences simply do not have a place for inherently prescriptive facts and properties, and so the reductionists have no grounds to add them to their account. Perhaps, taking a hint from Hume, this queerness could be diffused by locating the intrinsically motivating character in the interaction between the natural property (e.g., maximizing utility) and our psychological make-up. That

is, we might be naturally disposed to find maximizing utility appealing, or murder revolting, and those states could motivate us in ways explained by psychology. This would go some way toward explaining some the phenomenology of normativity (i.e., that we *feel* a certain way when a norm "has a grip on us"). But in that case, this would no longer be a reduction of, say, moral goodness to utility maximization; it would be an account of the psychology of motivation appended to that synthetic identity statement. While that might prevent queer properties and facts from entering the picture, it would also seem to jeopardize the identity statement itself; not everyone is attracted to what is morally good, what we are attracted to/repulsed by shifts from culture to culture and historical epoch to historical epoch, whereas theoretical identities should not make such shifts, and so on.

The final problem with this proposed solution is that it does not respond to the original problem posited in this section. We noted there and in Sect. 3.2.1 that the normative content of moral and epistemic utterances cannot be definitionally reduced, and that this stems from the *sui generis* nature of the normative with respect to the natural. To avoid making the normativity of morality a feature of the meaning of moral terms, we have been considering the proposal that it consists in the ability of moral goodness to motivate normal perceivers. But this fails as an attempt to reduce the *normative* content of morality, even where noting motivational structures offers a partial explanation of our behavior. To say that humans are necessarily motivated to do X is generally a very good predictor of their doing X, all things being equal, but it is not the same as saying that X is good, or that humans ought to do X. Christine Korsgaard states this rejection of motivation as an explanation of normativity in her *Sources of Normativity*:

> One possibility…in connection with that theory [is] that our moral instincts would be so strong that they could move us, or at least make us miserable, even if we decided that their claims on us were illegitimate. The theory might then explain moral conduct, including the conduct of people who know the theory. But it would not be normative, because the people themselves would not think that their conduct was justified. If they could cure themselves of their instincts, they would. (1996, 87–88)

Although she is speaking specifically of theories that explain moral behavior in terms of evolutionarily selected behavioral traits here, the point applies to the proposal we have considered in this section, and we would argue, to all similar attempts.

While such theories might be *explanatory*, they are not *justificatory*. They can explain broad (perhaps even universal) patterns in our behavior, but such descriptive regularities of behavior cannot by themselves suffice to justify our behavior. The problem is thus general: it is not as though a different or better reduction might somehow reduce this normative content while preserving its normativity. Attempts to reduce normativity without remainder to natural facts (as the attempt to reduce moral normativity to the ability to motivate) simply strip the normativity out entirely.

3.2.3   Practical Reason and Normativity

Our purpose in Sect. 3.2.2 was largely illustrative, rather than genuinely exegetical. Queerness objections are familiar to metaethicists and other philosophers concerned with normativity, and no philosopher who identifies morality with some natural property (*via* synthetic identity statements) will say that the normativity of morality consists in the ability of moral facts to motivate. We offered that proposal only because it brings out clearly the problem the reductionist is facing: the impossibility of reducing the normative surplus to a physical fact.

On the contrary, many ethical naturalists deny that moral facts are, in and of themselves, reason-giving at all. Externalism about moral reasons is the norm (no pun intended) for most reductionists. To be an externalist about moral reasons in this sense would be to deny that being under a moral obligation to Φ entails that you have any reason to do Φ. Such externalists might say that, given their interests and desires, *most* people would have reasons to do Φ, but this is not a matter of entailment or necessity, and there is nothing intrinsically reason-giving about moral obligations. David Brink (1989) has argued that even correctly judging that you ought to Φ need not give you any reason whatsoever to Φ. Reductionism of the usual sort would permit the positing of moral

facts (they simply reduce to other kinds of facts), but they would not be intrinsically reason-giving, either. Moral facts and entities would not be the source of moral normativity, assuming there is any. Instead, moral normativity will arise from our theory of reasons or rationality. Epistemic reasons will run along parallel lines. No epistemic fact is reason-giving, and the normativity of truth-seeking in our beliefs arises from our theory of reasons or rationality. If S is justified in believing that P, this does not entail that S has a reason to John knows that P. Given their interests and desires, most agents will have some reason for believing P (in cases where it is justified) given that they have an interest in the truth in general. One might argue that having truth-oriented reasons—being one who is reliable about forming at least some kinds of true beliefs and who cares in general about whether the beliefs they form are true or false—is a necessary condition for counting as an agent at all. However, this sort of externalism about reasons leaves epistemic justification, like moral justification before it, devoid of any power to confer reasons. Epistemic John knows that P is no reason to John knows that P. We find this a troubling outcome for any purported account of justification.

So we have come to the reductionist's final gambit. Since the normative surplus is not a feature of the meaning of normative terms, nor some other set of naturalistic facts, it must be shifted over to a theory of practical reason. This turns on a well-developed, comprehensive theory of rationality showing that agents with characteristically human concerns do *for the most part* have a reason to perform morally right actions, regardless of their subjective motivational set, even if those actions' being right offer us no reason to perform them. This would leave us with no specifically moral or epistemic normativity—no normative or evaluative content of any sort resides in them—but normativity would be constituted by practical reason itself.

We find this proposal deeply implausible on its face. It denies any normative content to moral and epistemic concepts, which would render them unrecognizable as such in our view. But there are deeper problems than our incredulity.

Let us grant to the externalist for a moment that a person does not always have a reason to perform morally obligatory actions, or to adopt justified beliefs. However, on the view under discussion, someone who

did not do either of these things would not be subject to any sort of *normative* criticism; there is no normativity to morality and epistemology on externalist views, after all, so there would be no basis for any such criticism. With all of the normative works shifted to a theory of practical reason, there is not necessarily anything *normatively* amiss with someone's failure to do these things. For moral and epistemic commitments, this would be a remarkably implausible position to hold. Brink notes that even if we embrace externalism about moral reasons, "We could still charge people who violate their moral obligations with immorality, even if we could not always charge them with irrationality" (1989, 75). But if there is no moral normativity as such, then *to charge a person with immorality is not to evaluate the person at all, or to make any normative judgment of any kind*. (We could make an analogous point about those who violate epistemic norms or standards.) We should make clear that Brink is not committed to this implausible position (nor is any other philosopher we know of). But the reductionist has run out of options, and the work of constituting normativity has to be done somehow, if it is not to be eliminated altogether. The most serious difficulty for this last reductionist proposal is before us, though: the normative surplus must be reduced by the theory of practical reason, if the reductionist is to remain consistent and place normativity in the natural world.

Even in a project of this length, we cannot hope to address everything that might arise in an attempt to reduce practical reason to a set of natural facts, but we do not think this is necessary. There are very strong reasons for pessimism about the prospects for such a project. As we have seen, we cannot simply identify normative and natural facts in domains such as morality, epistemology, and semantics. Practical reason does not have any additional features that would obviously make it a better candidate for such identities. Suppose that an agent has a decisive reason to maximize utility. Could this be reduced to some natural fact such as her wanting to want to maximize utility? Suppose the reductionist might say that "Jane ought to do X" and "Jane wants to want to do X" are co-extensive, but not synonymous, but it should be clear by now that this move only passes the buck without really solving any of the reductionist's problems. If Jane wants to maximize utility, it still makes sense to ask whether this entails that Jane ought to maximize utility.

Attention to the matter of *objective* authority provides further reasons for supposing that practical reason cannot be naturalized. In particular, Hampton (1998) has argued that this feature of reason cannot be naturalized. According to Hampton, both instrumental and non-instrumental reasons rely on the notion of objective authority—the notion that a person is bound by reason regardless of his or her desires or interests.

For instance, the Kantian notion of a categorical imperative exemplifies this idea of objective authority. Ostensibly, they bind agents regardless of the particular sets of desires, interests, and preferences we might have (e.g., *do not lie*). Hypothetical imperatives bind only those agents who possess the end or desire embedded in the antecedent of the imperative (e.g., if *you do not wish to go to jail*, then do not lie to the police). But Hampton argues that even hypothetical imperatives implicitly rely on the notion of objective authority. If you have a desire (or some end which you concede is valuable), then you have a reason to take the means to achieve that end, *regardless of whether or not you believe or acknowledge that you have such a reason*. The objective authority of a reason thus seems to be a fact not reducible to some feature of human psychology. As Hampton writes,

> However contingent the hypothetical 'ought' is on a desire, it is still not the same as a desire; to say, therefore, that its objective normative authority is what moves us to act rationally is to analyze the 'prescriptive force' of hypothetical imperatives such that it is identical to the prescriptive force of categorical imperatives. (1998, 163)

So even hypothetical imperatives—which seemed ripe for grounding in natural facts—exemplify objective authority. But objective authority is not a viable part of an account from the naturalist who favors reduction. Hampton writes,

> Is the idea of objective authority acceptable from a naturalist standpoint?… [T]he idea of an authority that is objective is ineffable—that is, impossible to pin down in a way that seems to make sense. From a naturalist point of view, this ineffability militates against its being a real phenomenon in the

world, and is instead a good indication that the authority of reasons is a psycho-social phenomenon. (1998, 99–100)

So we can now see how difficult and implausible the remaining path for the reductionist would be. In order to successfully reduce morality and epistemology, she must show (a) that moral and epistemic concepts have no normative content, (b) that there is no moral or epistemic normativity, and (c) that the normativity of practical rationality can be naturalistically reduced. All three hurdles seem insurmountable, and so the prospects for fully reductive theories of morality and epistemology do not look very good. This seems to leave the naturalist with two options: eliminativism or supervenience. We argued in Chap. 2 that eliminativism is not a good option, though we will revisit this claim in the last section of this chapter. In a moment, we will turn our attention to the second option, supervenience.

But first, let us consider one final rebuttal on behalf of the naturalist. The naturalist might argue as follows:

> We don't need to give a reductive account of the normative surplus. All we need, to establish the naturalistic *bona fides* of normative discourse, is to establish the natural *facts* to which normative terms refer. The reductionist does this (and shows that such facts are scientifically respectable), thereby securing the legitimacy of such discourse.[6]

This response is ultimately unsatisfying. As we have emphasized (and will continue to emphasize), what is characteristic of normative discourse, and what a theory of normative discourse must capture, is the prescriptive character of normative discourse. Thus, any theory of normative discourse must account for the normative surplus of normative discourse. As we have seen, the reductionist must shuffle this normative surplus off onto either the theory of meaning or the theory of practical reasons. However, by the very reasoning expressed in the above rebuttal, for these two additional types of discourse to establish their bona fides, they must be given a reductionist account. That is, we must give a reductionist account of the *fact-stating role* of these types of discourse. And since we have shuffled the normative surplus off onto one of these

two types of discourse (semantic or practical reason), that means we must give a reductive account of this normative surplus; we must explain what semantic fact (or practical reasoning fact) constitutes this normative surplus, or to which this normative surplus is reduced in the theory of meaning or practical reason. But we have argued that no such account can be given in a way that is consistent with the reductionist's commitment to the *sui generis* nature of the normative with respect to the natural. Thus, this rebuttal fails.

3.3   Against Supervenience Accounts of Normativity

We proceed on the assumption that reductionist accounts do not offer satisfactory accounts of normativity. The most serious concerns for reductionist accounts of normativity are problems in principle for such accounts, not simply the fine-grained details of different variations on the theme. If we are to place normativity within the natural world, then the remaining option for the naturalist would be some form of supervenience.

Supervenience captures a set of expectations about the correlation and determination relations that hold between sets of properties. "Higher order" properties and facts may be composed, constituted, or otherwise determined by more "fundamental" ones. For two sets of properties, A and B, A-properties and facts supervene on B-properties and facts iff all changes in A-properties occur as a result in changes in B-properties. Aesthetic properties and facts might thus supervene on physical ones, economic properties and facts might supervene on social or behavioral ones, and so on. Every fact and property will trivially supervene on itself, and every reduced type will supervene on its reduction base (e.g., *being water* will supervene on *being H_2O*). More potentially interesting appeals to supervenience involve separate, well-demarcated domains that purportedly do not reduce to one another, such as the mental and the physical. (So we are actually turning to *non-reductive* supervenience relations and accounts predicated on them.) More precise formulations of supervenience conditions with due attention to modal force are necessary to

capture the elusive sense that the relation might still hold where reduction fails between two domains of facts and properties.

Many philosophers concerned with normativity, particularly those in metaethics and the philosophy of mind, have embraced non-reductive supervenience. The hope in many such accounts is that supervenience is a metaphysical thesis that allows us to have our cake and eat it, too. The theoretical and explanatory autonomy of a supervening level will be preserved, but it will entail no ontological commitments incompatible with naturalism. But if this is to be a metaphysical thesis that places normativity in the natural world, the relation between supervening properties and subvening ones must *somehow* explain this supervenience. It will be facile to simply work out an account of moral or epistemic normative content, and when pressed on the metaphysics of those distinctions say, "Oh, it's fine. Everything just supervenes." If some properties do in fact supervene, then *how* and *why* they do so must be established.

To demonstrate the problems with such approaches to placing normative properties non-reductively, we consider strong, weak, and global supervenience in turn. As we shall see, stronger readings of these supervenience strategies undo the hope for non-reductive accounts, while weaker readings undercut explanatory requirements for any placement strategy, and there is no sweet spot in the middle for non-reductive supervenience accounts of normativity to occupy. We thus offer a "collapse" argument for the failure of supervenience to account for normativity: all such accounts will either fail on their own terms as explanatory projects, or collapse into a form of reductionism that we have discounted here.

3.3.1   Strong Supervenience and Normativity

Following Kim (1987, 316),[7] we may define a strong supervenience relation this way:

> [A strongly *supervenes* on B just in case] for any worlds w_j and w_k, and for any objects x and y, if x has in w_j the same B-properties that y has in w_k, then x has in w_j the same A-properties that y has in w_k.

This implies that supervenience holds across all possible worlds, and this will figure prominently in subsequent arguments. (Strong supervenience accounts of normativity have been defended by Walton (1994) and Ridge (2007), among others). The most straightforward manner of fulfilling this sort of relation—one that will figure in our discussion throughout this section—is property identity. Table salt (suitably idealized, impurities aside, etc.) just *is* sodium chloride, and so table salt properties strongly supervene on sodium chloride properties. If some sample in w_j is sodium chloride by virtue of its chemical microstructure and another sample in w_k is sodium chloride by virtue of the same microstructure, then both samples, each in its respective world, will be table salt. If both samples fail to be so, neither will be table salt.

The strongly modal character of this form of supervenience commits us to trans-world correlations between properties, and those identities will have to be single properties or subsets of all possible properties. (Complete sets of properties and whole worlds will come up again in Sect. 3.3.3.) Every time we come across some set of B-properties, we will also find its corresponding set of A-properties. Yet, if we assert that the supervenience between normative and non-normative properties and facts is not a matter of reduction, we must somehow establish that the connection between the normative and non-normative will hold across any two possible worlds despite the panoply of possible variations between worlds. This will create an intolerable dilemma for the non-reductive supervenience theorist, who must either accept reduction to explain the trans-world correlations or simply accept these correlations as brute and unexplained.

If there is trans-world correlation of A-properties with B-properties, then that correlation is not coincidental. If it were simply a coincidence that two properties were instantiated at a world (as it surely is for very large numbers of properties in any given world), then it is uncontroversial that there should be worlds at which that correlation does not hold. If the trans-world correlation at stake is not coincidental, but also not a matter of reduction, then either B-properties determine A-properties or some third set of C-properties determines both sets. The first option sounds enticing, and it expresses an intuition that drives many philosophers toward supervenience. But what form is this determination to

take, if it has to obtain across possible worlds? No form of analyticity (not even a robustly pre-Quinean one) fits the bill; if the supervenience is non-reductive, the ways in which normative facts supervene on non-normative ones are something that we discover empirically.

One possibility would that some non-normative properties somehow *cause* normative ones to be instantiated. Not all causal relations are strongly supervenient in this sense. For instance, some match-strikings cause some match-lightings, but there can be match-lightings in the absence of match-strikings, and match-strikings without match-lightings. The sort of possibility to consider here is one in which B-properties are instantiated only if A-properties cause them, yet B-properties are not identical to A-properties. There will be instantiations of metabolic properties only if there are instantiations of some redox properties, which can be said to cause the metabolic properties (perhaps alongside many others), but metabolic properties are not identical with redox properties.[8] Causation is a philosophically challenging subject, but this seems like a non-starter for normativity in any case. How might an instance of, say, inflicting unnecessary pain on a child *cause* the moral wrongness of that action? There is no provision for such causal pathways in our theories of physics (or elsewhere for that matter), and the wrongness here does not seem to be a *result* of the action, but rather a characterization of the action itself. Surely, we can cause *consequences* that we praise or blame, but in these cases, we appear to have non-normative features of a world leading to other non-normative features of that world, and we would still need to say how normative properties supervene on some or all of those non-normative ones. This sort of causality would give us the trans-world correlation needed for strong supervenience, but it does not seem to fit normativity at all.

Perhaps we could say that the non-normative facts and properties *compose* the normative ones. There is something intuitively appealing about this, in that it suggests normativity gelling into place when certain non-normative properties obtain, but it will not do for our present purposes. Why should particular constellations of non-normative facts and properties happen to compose normative ones? If we take it that more fundamental objects and properties can sometimes coalesce in ways that compose other objects and properties, then we may take it that there is an order and organization to the composite object that

is not captured at the subvening level. Such supervenience claims are a bit oddly placed here, though. Strong supervenience may hold for some sorts of composition relations, as when arrangements of elementary particles compose atoms, or arrangements of atoms compose molecules. But that is because being a particular sort of atom *just is* being a particular arrangement of elementary particles (i.e., property identity). If we are to say that non-normative properties compose normative ones here, then we must still explain how that composition gives us the distinctive character of the normative properties (i.e., their *to-be-doneness*). In the case of non-reductive supervenient properties like the metabolic ones mentioned above, there is at least a partial explanatory contribution to the supervenient properties from the subvenient base. Metabolic properties involve organisms' converting material into energy for their functioning, and the physical properties that compose metabolic ones do generate such energy.[9] (We discuss our own view of these sorts of explanatory contribution at length in Chaps. 8 and 9). But whatever a set of non-normative objects and properties might compose, simply creating another supervening layer will not explain the to-be-doneness without either adding "queerness" or running afoul of Moore's Open Question argument.

Consider an epistemic example. We would have to say that for some set Φ of non-normative properties, and any agents x and y, if x has Φ in w_j and y has Φ in w_k, then x has a justified belief in w_j and y has a justified belief in w_k. (Φ here might be something like a set of psychological properties.) If this is to be a non-reductive strong supervenience account, then Φ is not a *reduction* base for epistemic justification; it is just a set of properties that composes epistemic justification time and time again, yet also does not add "queer" properties to the world in the process. What we still lack is an explanatory pathway from the non-normative properties to the normative ones akin to the point about energy in the metabolic example, so insisting that they compose the normative consistently remains unexplained. (We cite psychological properties like reliability in making epistemic evaluations, but that does not show that reliability somehow *composes* justification in the present sense.) We could try other notions here—B-properties realize A-properties, for instance—but the problem will repeat, *mutatis*

mutandis, for those proposals as well. These are attempts to preserve the trans-world correlation strong supervenience needs without establishing just what would give rise to it.

What about our second option, that some third set of C-properties (where A≠B≠C≠A) determines both the A-properties and the B-properties? Consider an example: the property of *having greater mass than a helium atom* strongly supervenes on *having greater mass than a hydrogen atom*. For any possible worlds w_1 and w_2 and any entities x in w_1 and y in w_2, if x in w_1 and y in w_2 both have greater mass than a helium atom, then x in w_1 and y in w_2 both have greater mass than a hydrogen atom. Strong supervenience will hold, but not in a very interesting way; insight comes from knowing more about the properties of various atoms' microstructures. It is not that there is something intrinsic to the property of *having greater mass than a hydrogen atom* that somehow shapes, constitutes, composes, or realizes *having greater mass than a helium atom*. These correlations are determined by a further set of properties: the elements' respective masses and the conditions of their microstructures.

But there is no obvious candidate to play this larger role of causing or determining supervening normative properties that is somehow "behind" the scenes. If we say, for instance, that moral wrongness strongly supervenes on some subset of states of pain, and that we can be sure it does so because both sets of properties are somehow determined by some set of physical properties, then we preserve strong supervenience between pain and moral wrongness only at the expense of the explanatory import of pain. Once we cut out the middleman, we are once again left to ask how the trans-world correlations will hold without collapsing into reductionism. So this second option really has no legs, and with it goes any real hope of articulating non-reductive approaches to normativity in strong supervenience terms.

3.3.2 Weak Supervenience and Normativity

Next, we might consider weakening the supervenience conditions that must hold between normative and non-normative properties. Again fol-

lowing Kim ([1984], 158), we define a weak supervenience relation by saying:

> A weakly *supervenes* on B if and only if necessarily for any x and y, if x and y share all properties in B then x and y share all properties in A—that is, indiscernibility with respect to B entails indiscernibility with respect to A.

Note that we do not have the sort of trans-world comparison in this definition that we did in that for strong supervenience. Strong supervenience holds when something with a set of B-properties would be A-indiscernible from anything with those B-properties *in this or any other world*; weak supervenience holds when things *in a given world* that are B-indiscernible are A-indiscernible. Boyd ([1988], [2003a], [b]) argues in this vein for "homeostatic property clusters" of natural properties, coinciding in explanatorily salient ways, upon which normative properties could supervene.[10] Blackburn ([1984], [1993a], [b]) suggests that weak supervenience is a very hospitable position for metaethicists and moral theorists who want to avoid supernaturalism to take, and that it gives us an argument in favor of moral anti-realism. If moral properties are determined by natural properties (supervenience), but do not reduce to them (i.e., only weak supervenience), then there can be failure of the sort of trans-world correlation that strong supervenience suggests, though not a failure of intra-world determination. For the anti-realist, this sort of distribution can be explained by projectivism, while for the realist, it would remain mysterious (1984, 184–187). Such a view would still give moral judgments some grounding in natural facts, though some naturalists might not find that sufficient to place them in the pertinent sense.

Weak supervenience still creates a significant explanatory burden for anyone adopting it as a placement strategy. And another threat looms: weak supervenience may hold in cases where the correlation and A- and B-properties is incidental in a way that undercuts any claim to have placed the normative in the natural world. This might happen in cases where the supervening property does not vary, and thus its relation to the subvenient level is trivial. We may also face cases in which the base properties are so extensive and complex, and the manner in which it determines the supervenient properties so murky, that asserting weak supervenience does

not place or explain the normative in any substantial way. (In particular, we are concerned that blithely saying "the normative supervenes on the physical" exhibits this flaw.)

We suggest that the demand here can be articulated with a pair of conditions:

 (i) *The informative base condition* (IB): The B-properties inform our understanding of the systematic relations among the A-properties in some theoretically significant way without replacing them (hence, this is a weaker requirement than fully explaining the supervening level or deriving its laws).
(ii) *The Goldilocks condition* (GC): The B-properties must be a finite set that is not too small (which would invite reduction) nor too large (which would make the determination of A-properties trivial).

In cases where IB is met, but GC is not because the base property set is too small, we will get strong supervenience or reduction (e.g., being water supervenes on being H_2O). It seems at least logically possible that a single A-property could weakly supervene on a single *B*-property without A reducing to B, but we do not in fact find that in the cases that concern us here (e.g., being a wrong action does not turn out to supervene on causing pain, as many counter examples show us). In cases where IB is not met, we get incidental or mysterious correlation. For instance, being a perfect number less than 100 weakly supervenes on being a platypus, but only because the A-property here is necessary and unchanging, so platypushood makes no explanatory contribution to it. (This example gives us strong supervenience, too.) The real concern when it comes to normativity is that we may judge two cases differently, note that there are non-normative differences between them, and thereby assert that supervenience is preserved while never really spelling out how the non-normative does the determining. There are also dangers in having too expansive a set of base properties, but we will save discussion of those for a moment.

How might all this be applied? Suppose (for the moment) that the property of *being an enzyme* weakly supervenes on a set of chemical properties, but does not reduce to that set. In other possible worlds

where organisms evolved to have different sets of biochemical properties, the particular compounds that act as enzymes in the actual world might not serve as enzymes at all. The properties that make compounds enzymes are not an ad hoc collection, nor a vaguely bounded catch-all for a whole world. (This would lead to global supervenience, which we will discuss shortly.) The catalyzing roles that enzymes play will be partially accounted for by the chemical properties (their propensity to break various bonds, etc.), but the relevant sorts of catalysis that makes them enzymes will depend on their integration into further biological processes such as metabolism and immune response. The chemical properties that serve as a base here are well-bounded precisely because the determination role they play is informative, even though it is neither reductive nor a complete explanation of enzymology. Both IB and GC are thus met, so saying that the B-properties determine the A-properties is neither trivial nor evasive, and weak supervenience in such a case would not be mysterious at all.

Placement strategies for the normative formulated in terms of weak supervenience will be unacceptable in our view because they violate both IB and GC. They will fail to meet IB because we have no explanatory route from any non-normative supervenience base to prescriptive features of the normative. They will fail to meet GC because the non-normative properties that are relevant to the normative judgments we make will be too promiscuous.

Consider failures of IB here. The laws and mechanisms of, say, physics are frequently informative (if not complete) in explaining levels of properties that supervene on them, even when reduction fails. For instance, the thermodynamic properties (among many others) of the molecular mechanisms on which metabolic processes supervene are conducive to understanding the patterns we find at that supervening level. (Why do mammals eat so much? High metabolic rates to generate heat and regulate their body temperatures.) But we do not find this sort of systematically informative character in different layers of non-normative properties on which the normative might be said to supervene.[11] The laws of physics do not find a mirror in the normative, nor do the finer-grained details of normative discourse reveal themselves as smaller models whose details are determined by physical details. Just as we saw with reductionist pro-

posals, the "to-be-doneness" of the normative is alien to other levels of discourse, and descriptive accounts of various regularities are fundamental mismatches in explaining what it means to say that something ought to be done a certain way. The demands for explanatory links between weakly supervening properties and their bases have been loosened considerably as we move away from reductionism, but there are not additional resources to inform our understanding of how the non-normative could determine the normative facts.

The news for GC will not be any more encouraging. We have seen that accounts with very narrow supervenience bases end up being strong supervenience or reductionist accounts. A reasonable response to this is to note the holism of the non-normative properties that bear on normative distinctions. We may not come across single non-normative properties that determine normative ones, but clusters of non-normative properties may more clearly have such a determining role. So wrongful action does not weakly supervene on causing distress, but maybe it weakly supervenes on causing distress plus intentional action plus avoidable action plus… (some small, finite set of properties). Each term here will have to be non-normative (e.g., "intentional" will have to be cast in terms friendly to cognitive science). Nevertheless, even if we complicate the supervenience base here—and we readily agree that attention to more complex clusters of non-normative properties is a more enlightening move here—we only make it more apparent that there is no suitable supervenience base as required by GC. Any given non-normative property like causing distress can be shown to fall to counterexamples in which it plays no such determining role or gives rise to just the opposite sort of normative judgment. Causing distress to a child is generally a wrong action, but it is permissible (perhaps even morally praiseworthy) if one is telling a good ghost story. But it is again wrong if the child is acutely psychologically fragile, say, because of lingering effects of earlier trauma, and so on. And so on through rounds of expansion and other adjustments to the purported supervenience base. There is a temptation here to expand the supervenience bases until one has exhausted the variations. But there will not be a principled stopping point short of whole-world descriptions, and long before that point, we will have lost any sense in which the supposed determination of the

normative by the non-normative is well-bounded and conducive to a placement strategy for the normative.

So the prospects for weak supervenience placement strategies are just as bleak as they were for reductionism and strong supervenience. The way in which the base properties determine the supervening ones—even in this weaker sense—must be articulated in a convincing way, and we do not find resources for doing so. What gave both this and strong supervenience some teeth was our lingering sense that non-normative properties have something non-trivial to do with the judgments we make, and even if we are leaving this sort of supervenience behind, we should not lose sight of that appeal. Indeed, we will return to it in later chapters. But the problem of specifying a suitable narrow supervenience base suggests that at least one last possibility is to expand the base and weaken the supervenience even further instead, and this points the way to our last candidate in this chapter, global supervenience.

3.3.3   Global Supervenience and Normativity

One final possibility to consider would be a move to global supervenience.[12] As Kim (1984, 1987) and others have pointed out, strong supervenience would entail global supervenience, but not vice versa, and global supervenience has a modal import that weak supervenience does not. So in this subsection, let us consider ways in which global supervenience might hold, though strong supervenience did not. On Kim's formulation, we could say that A-properties *globally* supervene on B-properties when:

> Any two worlds indiscernible with respect to B-properties are indiscernible with respect to A-properties. (1987, 318)

By returning to trans-world correlations, global supervenience avoids one of the pitfalls we mentioned in Sect. 3.2. Two non-normatively indiscernible worlds could not differ in their normative properties and facts if global supervenience held. On the other hand, there might be more room for a non-reductive account since whole-world property sets are lining up

with one another, rather than individual properties or small finite subsets of them. Thus, we could more liberally say that all of the normative properties supervened on all of the non-normative ones without having to give more specific correlations among them. Global supervenience is thus a position as far removed from identifying supervening properties with base properties as one could get while still keeping the supervening properties within the confines set by the base. (For most thinking about supervenience, this amounts to keeping supervening properties within the *physical* world.) Haugeland ([1998](#)) defended a form of this in order to keep the mental within the realm of the physical, but do away with even token identities with physical objects and properties.[13]

The IB and GC conditions will still be concerns here, though. (Recall that they were conditions on naturalist placement strategies, not just weak supervenience accounts.) An account could meet the global supervenience condition while badly violating IB. Suppose we give an account, T, of normative properties that has them globally supervening on all of the non-normative properties. Assume that we have two worlds: w_1 with a set Γ of non-normative properties, and w_2 with Γ' which differs from Γ only in the timing of one beta decay in a region of space-time far removed from any sentient beings who might be bound by norms. Furthermore, assume that w_1 and w_2 have inverted sets of normative properties; for any true sentence N ascribing a normative property in w_1, not-N is true in w_2. If T permits this, global supervenience is preserved since w_1 and w_2 are B-discernible,[14] but IB is grossly violated. And global supervenience really amounts to an abandonment of GC altogether. If T permits the example described here, then there need be no systematic way in which variations in supervening and subvening properties vary with one another at all, much less a nice balance between chaos and property identity.

In some respects, a global supervenience relation may be almost trivially easy to satisfy. One can scarcely imagine possible worlds that differed from one another at all, but did not differ in some of their natural, non-normative properties, if one begins as a naturalist. But appeal to global supervenience is hand-waving for a naturalist if *the very fact that the relation holds* is presented as indicating explanatory work has been accomplished. So a naturalist pursuing this approach still must offer

some supplemental grounds to accept global supervenience as sufficient. We have seen the difficulty of articulating a non-reductionist alternative for the normative already. This appears to leave only one path for a global supervenience placement strategy. We would have to accept that the normative facts and non-normative ones are two aspects of the same world, each with their own order and logic that merit equal consideration and respect, but which simply will not unify with other explanatory projects.

This would seem to leave us to say that when a whole world's normative properties differ from another world's, their non-normative properties must too, but that we will never be able to say just how those non-normative differences determine the normative ones. We feel some sympathy with such a response, but as a *placement* strategy, it seems to succeed only by not really making any significant claims at all. It cannot offend the naturalist, because there are no supernatural entities posited, but it is as bloodlessly non-explanatory as it could be. Could this be resuscitated by emergentism, the view that some non-reducible properties emerge at certain levels of complexity of subvening elements? Most philosophers who have looked at emergentism have either taken it to be a form of strong supervenience or not a form of supervenience at all (e.g., Humphreys 1997). And even if we were to refit a notion of emergence to meet only the global supervenience condition, it would repeat a problem that has plagued all the placement strategies we have looked at so far. It would give us the *presence* of something—an entity or property—that stipulatively locates the normative without any further insight into how any such presence could do so without lapsing into queerness. This should not strike any naturalist, however moderate, as a suitable placement strategy.

Despite all this, global supervenience is not a condition that we want to deny wholesale. Where two worlds are indiscernible at the non-normative level, it would be mysterious for us to make very different normative claims about them. We think this is not because a significant ontological thread running through these issues has been uncovered, but rather because our consideration of other worlds is itself guided by some normative commitments. For instance, one very strong (if still defeasible in principle) norm of rationality is a respect for parity of reasons: roughly, when we consider a judgment or course of action, we should judge or

act in the same ways in cases where we have just the same reasons. Or conversely, if we judge or act differently in one case than another, there should be reasons for making that different choice.[15] To stipulate that two worlds were identical in all their non-normative properties would thus give us exactly the same empirical facts to consider in any judgment about normative matters (e.g., that pain was caused by a particular action), so a difference in judgments would require different empirical facts or else it would become arbitrary and irrational. But this is a normative commitment guiding us in how to make our judgments, not a discovery of some natural facts that constitute something normative. Normativity nudges us toward accordance with global supervenience, rather than the supervenience helping to explain normativity.

3.4   Eliminativism

Maybe the best placement strategy is no placement strategy at all. As we have noted, it seems highly improbable that you can give a suitably naturalist account of normative properties on which they are substantive properties placed in the natural world. (Mere aggregate properties are still available, but as we said, they will not suffice as means for placing the normative in the natural world.) There are at least a couple of different reasons for this. First, it seems unlikely that our normative concepts pick out a set of acts and objects that form a recognizable kind at the natural level. To see why these things belong together essentially requires the concept under which they fall. For example, the set of actions that are correctly categorized as "cruel" appears shapeless and gerrymandered at the natural level; only when one is competent in the use of the concept "cruel" can one see any shape to this set of objects. As McDowell argues:

> Supervenience requires only that one be able to find differences expressible in terms of the level supervened on whenever one wants to make different judgments in terms of the supervening level. It does not follow from the satisfaction of this requirement that the set of items to which a supervening term is correctly applied need constitute a kind recognizable as such at the level supervened upon… Understanding why just those things belong

together may essentially require understanding the supervening term. (1981, 145)

And indeed, if you consider the wide variety of things that can purportedly instantiate a particular moral property—say, *unfairness*—it is exceedingly implausible that all of these things have something in common at the natural level. An action can be unfair. So can omission (a *failure* to act); or a person; or an event, such as a race or a spelling bee; or an object such as a test; or an institution, like a company (or one of its constituent entities, like its human resources department or its grievance department); or the economic arrangements of a nation.

One might claim that the above instances—and, indeed, all instances of morally significant properties—trace their moral qualities back to an agent's (or agents') intentions, and so this is the property all have in common. However, the unfairness of a failure to act may have nothing to do with the person's intention (who may have failed to act out of ignorance, or distraction, or lack of consideration, or any of a number of reasons not related to any intention, good or bad). We can say the same for the unfairness of a test, or many other unfair acts or practices. An institution or economic arrangement that is unfair might be unfair without anyone intending this, or even despite the best intentions of its participants (maybe no one knows how to make it fair, or people acting with the best of intentions produce a system that is unfair, or no one has their hand on a sufficiently large lever of power to effect positive change).

It is easier to make the case that morally thin terms do not refer to substantive properties in the world. It is exceedingly implausible that all instantiations of a thin term like "unfair" (or "good" or "bad," etc.) have something in common at the natural level or form a recognizable kind at the natural level. But this argument is more troublesome when made regarding morally thick terms. Terms like "honest" seem to have a descriptive role, to pick out some natural property. Even Timmons, who challenges the dogma that all assertoric discourse must be descriptive and argues that morality is primarily a prescriptive enterprise, claims that "we have terms that are, in a sense, hybrid: they are partly descriptive and partly evaluative. I have in mind so-called thick ethical concepts like honest, courageous, and so forth" (1999, 132).

In the simplest case, an honest person could simply be one who always (or almost always) states what she believes. (Or, perhaps better: a person who never asserts what she does not believe.) A proposed analysis like the preceding will not do. The problem is not that the proposed analyses are too simple. The problem is that they attempt to analyze "honest" as a descriptive term, thereby ignoring the evaluative element of our use of this term. For two individuals can state deliberate untruths with identical frequency, and we may yet call only one of them dishonest. Thus, the extension of "honest" is not captured, well, extensionally.

Imagine a virtuous man, who hides Jews during the Holocaust. Presumably, this act of moral courage requires a fair amount of subterfuge, and he may well be required to lie frequently, even several times per day. His family may say among themselves, "Such behavior is difficult for him, as he is an honest man." Can he be an honest man, while telling many deliberate untruths per day? Surely he can, because we do not neutrally evaluate the number of his untruths; we evaluate them contextually. Another man, who tells a similar number of deliberate falsehoods each day (or even fewer), but in the pursuit of venal self-interest and sabotage of competitors at work, we would unhesitatingly describe as dishonest. But the moral context is different. And thus even morally thick concepts cannot simply be taken to describe in addition to evaluating: we do not ascribe or withhold a morally thick term like "honest" based on whether a person instantiates a particular descriptive property. The same descriptive property will sometimes merit the appellation "dishonest" and sometimes not. Rather, we ascribe or withhold a morally thick term like "honest" as part of an assessment of a person's actions and character in a particular context of evaluation, taking into account the point of the false statements, the moral worth of the intention behind these statements, and other morally salient features of the situation. Thus, we see that even morally thick properties have in the first instance an evaluative role and not a descriptive one—there is no descriptive property of "honesty" in the world that we are responding to when we issue our evaluations.

David Copp has recognized this point, to some extent. Thus, Copp attributes to expressivists like Hare the belief that "honest" (in addition to having evaluative content) "refers to some robust nonmoral property, such as, perhaps, the property of being disposed to assert only what one

believes" (2001, 40). Copp rejects this simple account, arguing instead that "if the term 'honest' is a moral term, then…it refers…to a robust moral property, such as, perhaps, the property of being disposed to assert only what one believes exactly to the extent that being so disposed is virtuous, or the property of being disposed to assert only what one believes exactly to the extent that being so disposed is called for by the ideal moral code of the relevant society" (2001, 40–41).

While this redefinition of "honest" shows sensitivity to the kind of case we presented above, it undermines any attempt to place the property of "honesty" in the world. For now, we have an interdependence between "honest" and another normative term, "virtuous," and this introduces a whole new web of interdependencies. Virtue does not require telling the truth if it would be *cruel* (but what is the difference between a cruel truth and a hard truth someone needs to hear?); and it certainly does not require telling the truth to the Nazis, as great *evil* will result from such truth-telling; but you should still tell the truth if you have done *wrong* (such as cheat on an exam), even if there are negative consequences for you, because the negative consequences do not *justify* your failure to tell the truth; and so on. Given the holistic interdependence between "honest" and all of these other essentially normative notions, which are themselves dependent on a raft of other normative notions, it seems implausible in the extreme that "honest" refers to a descriptive property that is going to be of any use to us in our theorizing (or in solving the placement problem). Again, the introduction of moral properties seems to create more problems than it solves. And if, as we argue in Chap. 7, we can get all that we need from a moral theory without the introduction of robust moral properties or facts, then we are better off without them.

One might argue, though, that we are not being sufficiently charitable to the naturalist and that the naturalist can give a naturalistic account of what unifies moral properties. For example, the naturalist might claim that moral properties are those properties that maximize utility (in terms of preference satisfaction), or in some other way satisfy human needs or desires, where this is given a naturalistic reading. But again, look at all of the things that can instantiate a moral property, such as we listed above. What these are all supposed to have in common is that they (say) promote human preferences in a particular way. But this is not a property

of these acts and items—it is more like a Cambridge property, not an intrinsic property of things at some non-normative level. So even if we can give a naturalistic explanation of why these things all have the same property (such as the property of "unfairness"), that is not the same thing as giving a naturalistic account of this property (i.e., that is not the same thing as specifying what they all have in common). A sophisticated position of this kind has been defended by Richard Boyd, who takes moral properties to be what he calls a homeostatic property cluster. Boyd writes of such property clusters,

> There are natural kinds, properties, etc. whose natural definitions involve a kind of property cluster together with an associated indeterminacy in extension. Both the property-cluster form of such definitions and the associated indeterminacy are dictated by the scientific task of employing categories which correspond to inductively and explanatorily relevant causal structures. In particular, the indeterminacy in extension of such natural definitions could not be remedied without rendering the definitions unnatural in the sense of being scientifically misleading. (1988, 196)

What unifies these various actions, character traits, institutions, and so on is their relation to human needs. Again, though, we wonder if an interesting or useful notion of a property emerges from this theory. Do these properties support counterfactual inferences? Do they support law-like generalizations? If not, then it seems like Boyd's attempted assimilation of morality to science fails. And we think the post-Gilbert Harman literature does not support such a robust characterization of substantive normative (particularly moral) properties.

The urge to place substantive normative properties in the natural world, as Boyd does here, is not without legitimate motivations. One of Boyd's chief concerns, for example, is whether observation can play the same role in moral reasoning that it can in science.[16] He worries that if there are no naturalistically characterizable, objective moral properties in the world, then the way the world is cannot epistemically guide and constrain our moral theorizing; and in particular, we cannot learn new moral facts from observing the way the world is. We would be no less apprehensive about such outcomes. But we believe this invites us to

invert the approaches to theorizing about normativity and how the non-normative can inform our judgments about it. Even if we do not posit substantive normative properties, our theorizing can still be answerable to the non-normative facts of the world, and these facts have an epistemic bearing on the moral claims we make. Thick moral terms, like those mentioned above, may invoke ancillary descriptive commitments (such as the death implied by "murder" or the distress implied by "harm") without collapsing into descriptive contents themselves. We will discuss this issue at length in Chaps. 5 and 6. The naturalist may feel that her very commitment to naturalism requires assimilating morality to science by positing naturalistic moral properties. But our approach will be that if we defend a theory of normative discourse that posits no moral facts or properties, then there can be nothing in our theory to offend against naturalism, and so the naturalist should be satisfied with the account given here. And finally, our contention is that the perceived need for an account with substantive normative properties was in the first instance driven by a representationalist account of language, according to which the primary function of declarative sentences of a language has content in virtue of corresponding to an independent reality. On such an account, the need for a substantive account of normative properties feels pressing, because how could normative sentences have content if there were no independent normative reality for them to represent? Given representationalism, the absence of normative properties or facts seems to lead one inexorably to eliminativism.

We contend that normative properties are not needed to do the work they are called on to do, and the very shapelessness and gerrymanderedness of the actions, virtues, institutions, and so on which instantiate any given normative property make the claimed existence of such properties dubious, anyhow. This very shapelessness fits better with the sort of account we are offering of how best to understand the relation between the normative and the natural. As we argued in Sect. 3.3.3, supervenience requirements should not be understood as a metaphysical thesis, but instead as a normative requirement concerning consistency. Such a requirement makes no commitment to the range of non-normative circumstances which would license (say) a particular moral judgment forming a recognizable property (much less a natural property!), and so the

shapelessness thesis fits well with our view of how supervenience require-
ments should be understood.

So we have come to an option that even many naturalists would seek
to avoid: eliminativism. If we are neither supernaturalists, nor reduction-
ists, nor non-reductive supervenience theorists about normativity, as the
rest of this chapter would imply, then we seem to be out of ways to place
it in the world. Given all the objections we have raised to standard reduc-
tionist and non-reductive supervenience accounts above, why not shift to
the position that normative facts and properties do not exist at all? We do
not believe that normative *stuff*—norms, or normative facts, or norma-
tive properties, or however we formulate this point—*exists* in the natural
world. Perhaps there are useful turns of phrase that suggest they are items
in the world, but we do not believe these should commit us in any full-
blooded sense that would entail ontological commitments.

But does this imply eliminating normative discourse altogether, as we
discussed in Chap. 2? Those who endorse eliminativism for normativity
generally do so because they have come to the conclusion that there is a
problem with normativity itself, and so it should be dismissed in all its
potential forms from our accounts of the world. Denying any ontologi-
cal commitments to additional normative entities is simply a matter of
bookkeeping after reaching this conclusion. However, there are indis-
pensible theoretical roles played by normative discourse and thinking in
normative terms, and that at least some forms of doing so are cogent and
well-supported. There are surely many terrible ways of thinking and talk-
ing in normative terms, but these are on a par with causal-explanatory
commitments that we reach in some form of error. We reject ontologi-
cal commitments here not because normative discourse is derelict, but
because it is sound, and no sort of *entity* available to us will underwrite
that theoretical role. To look for normative entities in the natural world to
explain normativity is to impose the wrong sorts of theoretical demands
on the account we give. To take the failure of such effort as an indictment
of normative discourse itself is to compound this mistake and distort our
self-understanding even more dramatically.

So our response to the eliminativists includes both a nod and a shake
of the head. We share their sense that there is no solution to the prob-
lem of where to *place* the normative in the natural world. But we will

argue that there is still a legitimate sense in which normative discourse may express truths, and interesting theoretical truths at that. Hawthorne (1994) distinguishes *ontological eliminativism* from *doctrinal eliminativism* about a subject. Ontological eliminativism entails that some range of entities does not exist; doctrinal eliminativism entails that some range of entities does not exist and that there are no true doctrinal claims involving those terms or entities. Thus, most of us will be doctrinal eliminativists about phlogiston and black bile, denying both that such things exist and that there are true claims about the actual world that invoke them. There may be true claims like historical claims ("Phlogiston was taken to be an oily fluid") or negative ones ("Oxygen is not phlogiston"), but none of any positive explanatory or other doctrinal significance. So we will happily concede ontological eliminativism about the normative, but not doctrinal eliminativism.

These efforts may strike some readers as revisionary rather than explicatory. Our sense is that they are not, and that the issues of the ontology of normative discourse are imposed on ordinary language by philosophers rather than originating there. We think that much of the consternation about normativity arises from philosophers being "bewitched by language," as Wittgenstein quipped. But we would argue for the plausibility of the account we would give even if it did require revision. The motivation for adopting doctrinal eliminativism about normative discourse and adopting an error theory toward it rests on an impulse to impose a crudely uniform way of interpreting their content: that true normative claims require distinct normative stuff that they represent. To strip out all normative matters would undercut the very possibility of all forms of discourse, including the natural sciences, for the sake of theoretical uniformity. That strikes us as the wrong tradeoff to make when there are viable alternatives. The chapters that follow constitute our attempt to articulate such an account.

Notes

1. This section develops several ideas introduced in Koons (2006).
2. Railton (1993) and Jackson (1998) exemplify this approach. Brink (1989) says that moral properties are "constituted" by natural ones,

which is somehow stronger than supervenience, so we would include him here. Actions themselves are troublesome here, as they do not reduce neatly to physiological event types or other non-intentional kinds. We will not raise objections to this for now, as we have other issues to raise about reductionism.

3. Just to name a few, Wilson (1998), Hauser (2006), and Casebeer (2003). Much of the work in this vein straddles ethics, psychology, and biology, but in ways these authors would contend are appropriate.

4. Boyd (1988) develops this notion.

5. This move in ethics has been extensively criticized by Horgan and Timmons in their series of "Moral Twin Earth" articles (1991, 1992, 2000). We do not intend to pursue their line of attack here, though; instead, we wish to explore whether the reductionist can explain the normative surplus which is not reduced while remaining true to her naturalistic commitments.

6. An objection along these lines was raised by Stephen Darwall when this material was presented at the 2005 meeting of the British Society for Ethical Theory.

7. Kim attributes this formulation to Brian McLaughlin.

8. Even this example is probably debatable. For instance, we will have to say that a metabolic process like digestion is *caused* by redox reactions that catabolize fats, carbohydrates, and so on rather than being type-identical with those. That seems to safely satisfy the non-reduction requirement, but whether the causal requirement is met is much murkier. We are really only making the claim that this is a dead letter for explaining normativity, so we will not belabor the point further.

9. We are assuming here that these roles in metabolism could be realized by multiple different molecular mechanisms (so there would not be property type-identities in play) and using the example because it formulates tidily. More ardent reductionists might think this is a case where reduction could succeed; if so, other examples would do.

10. Boyd (1988) also speaks at length about the "rational supervenience" of moral judgments on non-moral factual judgments, though this is not the sort of metaphysical thesis we have been considering in this section.

11. This is not to say that there are no informative connections between normative and non-normative discourse. There surely are, and we will have much more to say about this in Chaps. 8 and 9.

12. Global supervenience has been espoused by Lewis (1986).
13. Haugeland called it "weak supervenience" in 1982, but this term was later adopted by Kim and others for the position we described in Sect. 3.3.2. Haugeland later clarified that his position was a form of global supervenience (1998a, 3–4).
14. We assume that all worlds B-indiscernible with w_1 are still A-indiscernible.
15. Such a principle would presumably be restricted to judgments of some consequence. Whether to get coffee rather than tea before work one morning may be just a whim without undermining one's rationality.
16. See Boyd (1988, 206ff).

References

Blackburn, Simon. 1984. *Spreading the Word*. Oxford: Oxford University Press.

Blackburn, Simon. 1993a. Moral Realism. In *Essays in Quasi-Realism*, 111–129. Oxford: Oxford University Press.

Blackburn, Simon. 1993b. Supervenience Revisited. In *Essays in Quasi-Realism*, 130–148. Oxford: Oxford University Press.

Boyd, Richard. 1988. How to Be a Moral Realist. In *Moral Realism*, ed. Geoffrey Sayre-McCord, 181–228. Ithaca, NY: Cornell University Press.

Boyd, Richard. 2003a. Finite Beings, Finite Goods: The Semantics, Metaphysics and Ethics of Naturalist Consequentialism, Part 1. *Philosophy and Phenomenological Research* 66(3): 505–553.

Boyd, Richard. 2003b. Finite Beings, Finite Goods: The Semantics, Metaphysics and Ethics of Naturalist Consequentialism, Part 2. *Philosophy and Phenomenological Research* 67(1): 24–47.

Brink, David O. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.

Casebeer, William. 2003. *Natural Ethical Facts: Evolution, Connectionism, and Moral Cognition*. Cambridge, MA: The MIT Press.

Darwall, Stephen, Alan Gibbard, and Peter Railton. 1992. Toward Fin de Siecle Ethics: Some Trends. *Philosophical Review* 101(1): 115–189.

Fenske, Wayne. 1997. Non-Cognitivism: A New Defense. *The Journal of Value Inquiry* 31: 301–309.

Godfrey-Smith, Peter. 1989. Misinformation. *Canadian Journal of Philosophy* 19(4): 533–550.

Hampton, Jean E. 1998. *The Authority of Reason*. Cambridge: Cambridge University Press.

Haugeland, John. 1982. Heidegger on Being a Person. *Noûs* 16(1): 15–26.

Haugeland, John. 1998. Mind Embodied and Embedded. In *Having Thought: Essays in the Metaphysics of Mind*, 207–237. Cambridge, MA: Harvard University Press.

Hauser, Marc. 2006. *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York, NY: Harper Collins.

Hawthorne, John. 1994. On the Threat of Eliminativism. *Philosophical Studies* 74: 325–346.

Horgan, Terry, and Mark Timmons. 1991. New Wave Moral Realism Meets Moral Twin Earth. *Journal of Philosophical Research* 16: 447–465.

Horgan, Terry, and Mark Timmons. 1992. Troubles on Moral Twin Earth: Moral Queerness Revived. *Synthese* 92: 221–260.

Horgan, Terry, and Mark Timmons. 2000. Copping Out on Moral Twin Earth. *Synthese* 124(1–2): 139–152.

Humphreys, Paul. 1997. Emergence, Not Supervenience. *Philosophy of Science* 64: S337–S345.

Jackson, Frank. 1998. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Clarendon Press.

Kim, Jaegwon. 1984. Concepts of Supervenience. *Philosophy and Phenomenological Research* 45(2): 153–176.

Kim, Jaegwon. 1987. 'Strong' and 'Global' Supervenience Revisited. *Philosophy and Phenomenological Research* 48: 315–326.

Koons, Jeremy Randel. 2006. An Argument Against Reduction in Morality and Epistemology. *Philosophical Investigations* 29(3): 250–274.

Kripke, Saul. 1982. *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press.

Lewis, David. 1986. Against Structural Universals. *Australasian Journal of Philosophy* 64: 25–46.

Moore, G.E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.

Rachels, James. 2000. Naturalism. In *The Blackwell Guide to Ethical Theory*, ed. H. LaFollette, 74–91. Malden, MA: Blackwell Publishers.

Railton, Peter. 1993. What the Non-Cognitivist Helps Us to See the Naturalist Must Help Us to Explain. In *Reality, Representation, and Projection*, ed. J. Haldane and C. Wright, 279–300. New York, NY: Oxford University Press.

Ridge, Michael. 2007. Anti-Reductionism and Supervenience. *Journal of Moral Philosophy* 4(3): 330–348.

Rodriguez-Pereyra, Gonzalo. 2002. *Resemblance Nominalism: A Solution to the Problem of Universals.* Oxford: Clarendon Press.

Sellars, Wilfrid. 1953. A Semantical Solution to the Mind-Body Problem. *Methods* 5: 45–84.

Walton, Kendall L. 1994. Morals in Fiction and Fictional Morality. *Proceedings of the Aristotelian Society Supplementary Volume* 68: 27–57.

Wedgwood, Ralph. 1999. The Price of Non-Reductive Moral Realism. *Ethical Theory and Moral Practice* 2: 199–215.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations.* Trans. Rush Rhees, and G.E.M. Anscombe. Oxford: Basil Blackwell.

4

Truth and Pluralism

A major feature of our account can now return to the fore again. Part of our strategy in reconciling normativity with the naturalist themes described in our first two chapters will be to follow philosophers, such as Brandom and Price, in rejecting what we have called representationalism about meaning and content. On such a view, thoughts and descriptive parts of a language (paradigmatically, declarative sentences) are fundamentally representations of the extra-linguistic world, and truth and designation are the fundamental elements of analysis. If the objections we discussed in Chap. 3 to placing normativity in the natural world have bite, then representationalist accounts of normativity give rise to real problems: if there is nothing to place, then there is nothing to designate, no facts to state, and either all of what we say in normative discourse is false or it is so laden with empty, non-referring terms as to become vacuous.

One can reject representationalist assumptions for many reasons, and not everyone who rejects these assumptions will agree on what to take up in their place. Since our work is already quite expansive, we will not offer an extended elaboration of the grounds for our rejection. But one point that will inform the rest of this work is that to be an anti-representationalist is not to reject certain aspects of various forms of

scientific realism, or at least a realistic reading of the ontology of some theoretical discourses.[1] That is, if our best account of the physical commits us to, say, electrons, then that should be read as a commitment to the real presence of such entities. Our best theories may be wrong, and we may be mistaken in committing to the real presence of some entities and properties, but some commitments should be read in this way. We endorse some pluralist assumptions in this chapter that might rule out the most austere ontologies (e.g., only the stuff of physics), but our only actual goal here is to show that it will be possible to have truth-apt portions of normative discourse without additional ontological commitments. As we saw in Chap. 3, designating or pointing to some range of objects and properties never offers us any help in articulating what it is to make normative distinctions in the first place. Thus, we lay the groundwork here for a way of understanding normative discourse that does not depend on a special range of normative things that it represents.

To do this, we must set some boundaries for the alternative approach that we will pursue. First, we set out to introduce our own position on the truth of discourse that invokes normative vocabulary without appeal to supernatural entities or to reduction and supervenience strategies. Rather than adopting a specific position on truth within the existing literature (e.g., minimalism, correspondence), we indicate a range of possible deflationist approaches that are amenable to our purposes. We argue that these are better suited to all forms of naturalism, even those that are less friendly to normativity than ours. This paves the way for a more extensive elaboration of how normative discourse can include truth-apt sentences in Chaps. 5 and 6.

Second, we stake out a pluralist position on discourse and ontological commitment. As our first three chapters would suggest, we hold that there are numerous different types of discourse that address different theoretical needs, but will not reduce to one some fundamental discourse (such as, say, physics) even in the long run. Pluralism can take numerous forms of different strengths, however. Here, we consider (and reject) *strong* forms of discourse pluralism that insulate various forms of normative discourse from traditional attacks. While there are legitimate senses in which we may distinguish "regions" of discourse from one another, any such divisions hold only for certain theoretical purposes, and the

interweaving of different parts of the language is a deep and essential feature of each of them and the whole of our engagement with the world. This will preclude strong forms of discourse pluralism that promise non-overlapping magisteria for normative and non-normative discourse; we cannot shrug off ontological concerns about normativity as a result. A more plausible and informative pluralist approach would instead highlight the interweaving of different types of discourse and emphasize the variety and character of these intertheoretic relations. This paves the way for our more extensive elaboration of the interdependence of normative and non-normative discourse in Chaps. 8 and 9.

4.1 Truth

One assumption shared by many philosophers is that we must give a *substantive* account of truth (i.e., one on which truth is some sort of real property or relation, and indicating this relation is what makes an account of truth *informative*), rather than taking a quietist stance. Even philosophers who abandon representationalism about truth still often feel that something general can be said about what it is for a declarative sentence to be true, and feel that it is incumbent upon us to give an account of what this is. Given our rejection of reduction and supervenience accounts of normativity, along with our view that normative discourse will still be truth-apt, it will be necessary to offer a position on truth in general that does not require an entity, property, or set of conditions that elements of normative discourse designate. A theory of truth might be informative in satisfactory ways (i.e., not simply quietist), as we shall see, without being substantive in the present sense, and there are reasons for all naturalists to join us in some version of what we propose.

4.1.1 How Not to Not Worry about Truth (Why We Are Not Internalists or Relativists)

To some, attempts to get away from correspondence and other metaphysically substantive theories of truth smack of evasion—of seeking a

way not to talk about truth itself, but to talk about some anemic surrogate instead. We agree that some of these strategies will not suffice, and seeing why they do not will be instructive for our own account. Having abandoned correspondence and other representationalist theories, a common approach has historically been to give truth conditions in terms of members of a community satisfying certain epistemic or doxastic conditions. For example, Peirce (in)famously writes, "The opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by the truth, and the object represented in this opinion is the real. That is the way I would explain reality" (1878/1992, 139). Hilary Putnam defends what he calls

> the *internalist* perspective, because *what objects does the world consist of?* is a question that it only makes sense to ask *within* a theory or description of the world. 'Truth' in an internalist view is some sort of (idealized) rational acceptability—some sort of ideal coherence of our beliefs with each other and with our experiences as those experiences are themselves represented in our belief system—and not correspondence with mind-independent or discourse-independent 'states of affairs'. (1981, 49–50)

With representationalist theories of truth, objectivity was provided by the world: since truth was something like correspondence, our sentences answered to an objective, mind- and social practice-independent reality, and so worries about relativism were not particularly pressing. But if we define truth in terms of ideal rational acceptability, or agreement in the long run, or other epistemic or doxastic conditions, then relativistic worries loom large.

It is for this reason that advocates of such a view are usually at pains to emphasize that truth outstrips what a community might assent to at any particular time, perhaps at any time at all. For example, Peirce writes,

> [T]his view [of truth] is directly opposed to the abstract definition which we have given of reality, inasmuch as it makes the characters of the real depend on what is ultimately thought about them. But the answer to this is that, on the one hand, reality is independent, not necessarily of thought in general, but only of what you or I or any finite number of men may think about it; and that, on the other hand, though the object of the final

opinion depends on what that opinion is, yet what that opinion is does not depend on what you or I or any man thinks. Our perversity and that of others may indefinitely postpone the settlement of opinion; it might even conceivably cause an arbitrary proposition to be universally accepted as long as the human race should last. Yet even that would not change the nature of the belief, which alone could be the result of investigation carried sufficiently far; and if, after the extinction of our race, another should arise with faculties and disposition for investigation, that true opinion must be the one which they would ultimately come to. (1878/1992, 139)

Recall, also, that in characterizing his internalist conception of truth, Putnam refers not to what is accepted (even in the long run), but instead to "some sort of (idealized) rational acceptability." In this way, internalists about truth (to adopt Putnam's term) try to have their cake and eat it, too: they attempt to give a substantive account in terms of a community's practices (suitably idealized), and yet try to retain objectivity by assuring us that truth outstrips whatever the community's practices might be at a particular time. Certainly, if we care about truth, we should care about our community's epistemic practices. But there is a difference with being concerned with epistemic and evidential concerns surrounding truth-evaluation and identifying truth with some set of epistemic, evidential, or doxastic conditions. Indeed, there is a good reason to think that the internalist about truth is combining incompatible elements into a theory of truth, with incoherent results.

For example, consider Putnam's definition of truth above. Putnam is caught on the horns of a dilemma, caught between the demands of the internalism he embraces and the objectivity he seeks. For on the one hand, if we try to specify what counts as "rational" in rational acceptability, then we are specifying a standard of rationality according to our (internal) standards. We have reified truth and defined it according to a particular, historical, and contestable standard of rationality, and the question can (and should!) be asked, "Why should we accept this standard of rationality as final and ultimately binding and truth-determining?" By giving a concrete specification of the norms of rationality which determine truth, we have embraced the internalist aspect of the internalist theory of truth, but at the cost of the objectivity needed by any account of truth-apt discourse, and which any theory of truth must be able to preserve.[2] On the

other hand, if we do not specify in concrete terms what "rational" means, or specify wherein coherence consists, then we have not really given a substantive theory of truth at all—we have not specified the conditions under which a sentence is true. And so the theory is vacuous or at least discomfortingly incomplete.

The fundamental problem faced by internalists about truth lies in attempting to locate the external condition on truth (objectivity, out-stripping our actual practices) in something which is (given the tools at the internalist's disposal) inherently internal to our practices—what is rational, what people agree on, what people agree on when they are fully rational, a fully coherent set of beliefs, and so on. An account of what is rational must be internal to the practice for the internalist. Having rejected representationalist accounts of truth, the internalist cannot give such an account to vindicate the content of rationality. So what counts as full rationality cannot be any more external or objective than the notion of truth which is supposed to be built up out of this notion of rationality. If we define what is true in terms of what is rational, and then go on to ask whether a practice P is rational (i.e., whether it is true that a practice P is rational), then what answer shall we be able to give, except in terms of some other standard rationality internal to the practice (maybe even P itself)? So it seems like Putnam's internalism (unsurprisingly) does not have the tools to make truth sufficiently independent of our practices.

This problem of trying to build the external component of truth out of something inherently internal is tied to a deeper problem shared by inter-nalists about truth. Undoubtedly, all truth-apt claims we make are medi-ated by our language and our lived practices. For example, the sentence "Bob drives a Camaro" could not be true without the existence of a wide range of linguistic and other practices, all of which are historically con-tingent. But there is a difference between saying that all truth-apt claims are *mediated* by language and saying that their truth-aptness is *constituted* by elements within a language; and to slide easily from one to the other is simply a conflation. One sees this easy conflation, for example, in the work of Rorty, who counsels us

> to draw the consequences from Wittgenstein's insistence that vocabular-
> ies—all vocabularies, even those which contain the words which we take

most seriously, the ones most essential to our self-descriptions—are human creations, tools for the creation of such other human artifacts as poems, utopian societies, scientific theories, and future generations… From the point of view I have been recommending, any attempt to drive one's opponent up against a wall in this way fails when the wall against which he is driven comes to be seen as one more vocabulary, one more way of describing things. The wall then turns out to be a painted backdrop, one more work of man, one more bit of cultural stage-setting. (1989, 53)

In this passage, one can see the slide from "All truths are expressed using language" to "All there *are* are different languages." This leads us to Rorty's further view that there is no objectivity, only unforced agreement, and his definition of truth completes the conflation of "mediation by" and "constitution by" language and actual practice: "A liberal society is one which is content to call 'true' whatever the upshot of such [free and open] encounters turns out to be" (1989, 52, emphasis removed). A similar slide can be seen in Putnam's discussion of truth.

But the mere fact that truth-apt claims are conditioned by our practices does not entail that truth is simply a matter of whatever our practice says it is, that there is only language. To give a familiar example, baseball is uncontroversially a human invention, and the definition of a home run in baseball is purely a matter of convention. But even granting this point, the field of play has a specific definition, and whether a particular hit constitutes a home run will be a determinate matter, not merely one of communal agreement.[3] The umpire can be wrong, the fans can be wrong. Indeed, conceivably (if unlikely), everyone could be wrong about a specific play.

There is no easy inference from "X is mediated by our vocabulary" to "There is nothing but matters of vocabulary when it comes to determining the truth of X." Ultimately, then, the internalist, who wants to give up on representationalism but still have a substantive account of truth, winds up with candidate conceptions of truth that are trying to satisfy contradictory requirements: they are trying to settle facts while simultaneously remaining internal to the practice or language. At best, we wind up with the kind of dilemma we attributed to Putnam; at worst, we sim-

ply make the kind of conflation Rorty makes and give up on the external (objectivity) aspect altogether.

Ultimately, in rejecting representationalist theories of truth, the internalist about truth could not resist the temptation to identify truth with some terms in a language or some conditions of the social practice. In doing so, as we argued, the internalist wound up with a theory that was torn between its two competing elements, the internal and the external/ objective. But in rejecting such internalist theories, it seems that we are left without options, because we have also rejected the representationalist theories. We have rejected an account of truth in terms of robust word–world relations; we have rejected an internalist standard of truth; and so it seems that we have nothing left to appeal to in calling anything truth-apt. The problem is in thinking that objectivity requires a substantive theory of truth—that truth needs to be *identified* with anything substantive, whether this is word–world correspondence or some internal feature of our practice.

4.1.2 Deflationist Accounts of Truth and Normative Discourse

Rejecting intuitions about truth-evaluation will only take us so far, and one can hardly doubt that it is a very thin needle we aim to thread here. What we present here is intended as an elaboration of a view that has numerous precedents in the literature, rather than a wholesale defense. Most crucially, it will be an elaboration of the notion of truth-*evaluability* or truth-*aptness*. Certain forms of deflationism do not propose a substantive notion of truth, but do build in a number of constraints that make truth-aptness an objective matter, and still offer an informative account. In short, there is no single property, relation, or other substantive item to which we have access that lies at the heart of a notion of truth. On this point, we stand on the shoulders of giants rather than offering a novel account or substantial extension of an existing one.

Correspondence and various other substantive relations have crept into philosophy to elaborate truth-aptness, but naturalists should treat

them with greater scrutiny. As Price (1993/2011, 44–46), (2009/2011, 271–274) has noted, positing substantive properties of truth or representation actually amounts to a *worse* move for naturalists. The purported relata (linguistic, non-linguistic, abstract, concrete, etc.) to be related are wildly heterogeneous, and we cannot assign them a causal role in the natural sciences. (A naturalistic psychology might assign a causal role to speakers' representations, which both cause behavior and purport to be true, but their *truth itself* is not a causal factor.) This leaves us no path for placing them in the physical world, so substantive relations or properties of truth will be no friend to most naturalists. But even if we do not adopt a substantive theory of truth, we can preserve a sense of answering to the facts. Particular truth claims will depend on properties and relations in the world, of course. To assert or evaluate the truth of "Socrates died in 399 BCE" will depend on non-linguistic, non-conceptual features of the world at some point, as will countless other claims in countless ways. But invoking truth here, even explicitly, will not add a further relation or property to the mix that must be represented by our language and added to our ontology. As we suggested in the previous section, there is a substantial peril in identifying truth with some conditions internal to a set of social practices; deflationists of different stripes sense that there is a peril in *identifying* it with anything. (Thus, we would be equally wary of pluralist accounts of truth that purport to solve some traditional problems for theories of truth by permitting some plurality of properties to serve in truthmaking across different domains, so long as those properties are substantive and truth is identified with some of them for each domain.)[4]

The common core of deflationary accounts is that there is no additional metaphysical legwork in saying that a sentence is true than was involved in making the assertion itself. It would suffice for an informative account to instead articulate the formal, expressive, or pragmatic contributions that talk of truth makes. Our view, then, is consonant with philosophers from Quine to Horwich to Brandom to Price. If one could call all their approaches to truth "deflationism," while acknowledging the diversity among these authors, then we would be deflationists. What disquotationalists and minimalists from Quine to Horwich have convincingly argued is that there is no need to posit additional descriptive content to sentences invoking a truth predicate. No further

piece of the world is described, and no further ontological commitment made, when we call a particular sentence true. What the expressivism of Price and the prosentential/anaphoric accounts of Grover et al. (1975) and Brandom (1994) have convincingly argued is that truth talk will still have expressive and pragmatic roles that we cannot do without, even if no additional descriptive content is introduced. Lance (1997) has argued that these pragmatic functions are indispensable parts of any theory of truth, even those that embrace representationalist assumptions. While deflationists are united in the major assumptions they reject, there is no shortage of disagreement on other issues, leaving demanding technical challenges in articulating an account of truth. Horwich (2008) in particular is sharply critical of prosentential/anaphoric accounts in favor of his own brand of minimalism. We will not attempt to adjudicate these disputes in this work, as this task would require its own book-length treatment, and because any such resolution will not substantially affect the rest of the work we are doing.

So go our deflationary commitments when it comes to truth. Endorsing deflationism might be thought to leave us in just a precarious a position, however. In endorsing deflationism, we are denying that normative sentences are true in virtue of any substantive relation between these sentences and the world. The deflationist can say that "'Electrons carry negative is charge' is true" does not commit us to substantive properties beyond those of electrons and their charges. But if we have also denied that there are substantive normative properties, then what does that leave us to talk about in truth-apt ways at all? One wonders, then, how any normative sentence "gets a grip on the world" or on objectivity?

What can we say in a more positive vein at the moment? In both normative and descriptive discourses, there is both a framework established by sets of social practices and—when properly pursued—a dependence on and integration with the world in making the claims. For such sentences to be truth-apt, some of the conditions and criteria of application will be informed by involvement with the non-normative, non-social aspects of the world in which they are instantiated, even if that involvement is not by way of representing something. They may incorporate literally any way we encounter and engage with the world, and there will be no single, unifying type of relation that runs through all of

these sentences and evaluations. By committing ourselves to a social practice account, we are suggesting that there is no direct confrontation with the world and that the mediation of our encounters by those practices is neither optional nor disposable. In this way, we will sound much like Rorty and other neopragmatists. But what we will suggest in later chapters is that the practices we adopt respond to various practical necessities, and this permits a type of scrutiny of our own practices that outstrips whatever we take our commitments and entitlements to be at the moment; we can find that the whole community was wrong about some matter, even if the whole community agreed, and our being wrong "does not depend on what you or I or any man thinks" as Peirce put it.

4.1.3   A Preliminary Sketch of Truth-Apt Normative Content

So we have come to a point of asserting that normative discourse does not represent or commit us to substantive properties. And we have committed ourselves to a range of deflationary options in an account of truth that will ensure normative discourse remains truth-apt, so long as it remains "world-involving" in ways that we will elaborate in later chapters. But why should we think that normative discourse has to be truth-apt at all? If it does not have something in the world to represent, and its distinguishing feature is the sort of "to-be-doneness" we described in Chap. 3, then why struggle to offer an account of its truth-aptness? Perhaps it really is all merely expressive of our attitudes, nothing more than an announcement of our allegiance and a covert plea for solidarity. We do not feel this would do justice to what speakers, both in everyday discourse and in the most abstruse theoretical discussions, are saying and doing with their language. A crucial subset of the things we say in overtly normative terms really are declarative sentences that we assert. Error theorists are closer to the mark than emotivists in this sense. But our sense of urgency in preserving the truth-aptness of normative discourse also stems from a sense that those sorts of merely expressive approaches are in the grips of unduly sparse views of what speakers may assert and what the content of

overtly normative declarative sentences might be. To give readers a sense of what this might entail, we should say briefly here what we think is being asserted when we make such a move in normative discourse.

Consider the action-guiding character—the "to-be-doneness"—of normative discourse. To some degree, what we want to suggest is already accepted in accounts of prescriptive sentences and certain categories of speech acts. These are distinguished by their directing their audience to perform some actions or realize some goal. An imperative like "Drop and give me 20!" uttered by a drill sergeant or sports coach is taken as a direction to perform the action(s) specified; whether that speech act actually compels the audience depends heavily on features of the context at hand. For instance, someone may yell this at me when I am not a member of the military or the team in question, in which case I may reject the command for lack of authority. Some speech acts may also accomplish this in a given setting with sentences that do not include overtly normative vocabulary. "Your brother needs help," uttered by a parent to his child, may direct the child's actions to go help her brother even though its surface grammar suggests it is a report.

One payoff of accepting this category is a release from the usual demands of fact-stating discourse to describe or report something about the world. Imperatives, for instance, do not purport to represent, so there is no shame in their not doing so. The tradeoff is that such sentences will no longer be truth-apt, and accepting or rejecting them will be very different from judging a declarative sentence true or false. Ordinarily, in the case of imperatives, this is not a cost at all. A drill sergeant or coach who demands pushups is not even trying to report a fact to us, and queries about what the facts are would likely be met with bafflement. (And possibly lead to more pushups.) If there is no purport to report facts and their ancillary ontological commitments, then there is less in store to unnerve ardent naturalists, too. It will be tricky to say just how we get the authority necessary for imperatives, but the content of those utterances and their action-guiding character pose no additional concerns here.

But what of apparently declarative sentences asserted in normative discourse—those that feature "ought," "should," "must," moral claims, distinctions of correctness, appropriateness, reflection on interests, and methods? They seem to share this sort of action-guiding character with

imperatives, but they also appear to report something like other declarative sentences. We can make generalizations, singular ascriptions, and embed them in sentential operators, for instance:

1. All murders are wrong actions.
2. Abby's remark to Lucy was cruel.
3. If Abby's comment was cruel, then she should apologize.

All of these suggest that the purport to truth for these sentences is too deep to dismiss. Thus, declarative sentences in normative discourse have *assertion-like* conditions of application and *prescriptive-like* consequences of application. In making them, we are guided by concern for truth and our reasons, but "to-be-doneness" follows from accepting them. Such statements can be disputed by other agents in ways that subjective expressions of sentiment, allegiance, and other attitudes cannot. If I report a subjective preference for chocolate ice cream over vanilla, others cannot meet that with much more than shared expressions of favor or contrary expressions of disgust. But whether Abby was acting cruelly or not may be disputed in part by appeal to non-normative facts, even if resolution of such disputes is often difficult or elusive.

How are these two features compatible? We contend that what is asserted in declarative sentences such as (1)–(3) is that some target audience has some set of commitments or entitlements, and an immediate practical consequence of this being the case is that their comportment to the world and others should be directed in certain ways. In saying such a thing, we may additionally invoke our own authority to make such claims, either as experts who are more fit to make such judgments, or peers who can introduce something of normative import (e.g., make a promise). But we are not merely or even primarily invoking our attitudes and expertise any more than a chemist who is invoking their authority in asserting that water is H_2O. Declarative sentences in normative discourse are not "about the speakers themselves;" they declare that some target audience has some status that implies coordinating their actions in some way. The purport in making them is not bound to any agent, unlike the imperatives we mentioned earlier, which depend on the social standing of the agent making the assertion.[5]

We suggest that the common feature in the content of such overtly normative declarative sentences is the introduction of a privileged form of action or range of potential outcomes toward (or against) which an audience is guided. The audience here may be an individual in conversation, or a more abstract target of "all of us" and "how we do such-and-such." The resemblance to imperatives should be clear here. But the status of that privileged form of action or range of outcomes, as opposed to some other option from a much wider range of possibilities, will be a matter that agents can deliberate, discuss, and dispute, much as they would the grounds for any other declarative sentence to be asserted. Note that this is not to say that the preferred form or range of actions or outcomes is the best or the optimal in a global sense, though we may be committed to that in some cases. Many norms are simply matters of how things are done locally, and it is perfectly intelligible to ask whether another norm could replace the present one even if we acknowledge the grip of the present one. ("When in Rome…") And for normative declarative sentences, there will be empirical support and other reasons to settle firmly on that privileged range or outcome in the way there would be for sentences in a fact-stating discourse. Moreover, reflection on the action-orienting character of concepts central to normative discourse has given us a richly connected set of inferential roles (again, as we would find with fact-stating discourses), which permit us to reflect on them in systematic, theoretical ways, and embed sentences involving them in logical operators, including negations and conditionals. In these ways, the sentences expressing them bear all the hallmarks of other truth-apt sentences.

How do non-normative facts support assertions with action-guiding character in truth-apt ways? Critics might note that we can cite empirical evidence as reasons to use other non-declarative sentences that do not involve truth-aptness, too (e.g., the appearance of someone's name on a team roster in the pushups example above). Further elaboration will be needed, and it will come in later chapters. But for the moment, consider some examples and comparisons to illustrate the bearing of empirical facts on normative claims. Begin with a descriptive declarative sentence. Whether we should assert the sen-

tence, "Selenium has six naturally occurring isotopes," will be a matter of how our practices articulate a theory of chemistry and how the samples of selenium in the world actually turn out to be. Consider a normative example about scientific methods. Whether we should assert the sentence "Single observations do not confer epistemic entitlement to affirmative conclusions in experimental settings" will depend on how we articulate our scientific practices and how such inferences actually play out in guiding our reasoning. "How they play out" is a kind of dependence on the world. We adopt more stringent epistemic standards in building theories in general because systematic inquiry into the world suggests that its diversity makes one-case inductions hasty. Under other conditions, we might adopt different normative stances. If we somehow found ourselves in a world of much greater simplicity and homogeneity, such inferences might be much safer. For narrow ranges of phenomena, we might do this even in the actual world. A single piece of unexpected or very specific data may have enormous weight in a well-developed theoretical framework. Where we have well-developed frameworks for evaluating normative claims, much the same would be true.

Or consider the adoption of double-blind studies in medical research. As a matter of epistemic principle, their use became obligatory only by the mid-twentieth century. Prior to their adoption, researchers typically knew which patients received medications and which received placebos, while few checks were in place on the biases that might lie in even the most earnest observers' efforts. The first double-blind study in western medical literature came only in 1948 when Austin Bradford Hill used this method to test the effects of streptomycin on tuberculosis patients (Kaptchuk 1998). Thus, the empirical facts of failed efforts in the past, sloppy science in the present, and the efficiency and precision of experiments employing double-blind methods serve as partial grounds for accepting one general principle about experimental design rather than others.[6] A normative declarative sentence such as "Medical studies should incorporate double-blind methods" may be judged true in light of thoroughly non-normative empirical reasons, even if the claim is not *about* any of these empirical facts.

4.1.4 Normative Discourse and Quantification

If we adopt a deflationary account of truth, we will steer clear of one significant set of problems for an account of normative discourse that does not posit substantive normative properties in the natural world. But there is another problem lurking nearby, and many naturalists who share our deflationary commitments on truth will argue that commitment to normative properties in the natural world is not so easily jettisoned. It is, for example, a familiar tenet of Quinean naturalism that we are ontologically committed to the existence of a property or object if we quantify over that property or object in the best or canonical version of your theory of the world. We regularly use noun-like, (apparently) referring expressions without thereby incurring ontological commitments. I might say "We must act quickly, for Michael's sake!" without committing myself to the existence of sakes; or I might say "The average American mother has 2.3 children" without thinking that there is such an entity as the average American mother. Nor does the fact that we informally quantify over an entity entail ontological commitment: I might say, "All our fates are inextricably intertwined," without thereby being committed to fates in the long run. But if in our best theory of the world, we quantify over entities, *only then* are we ontologically committed to the existence of these entities. The Quinean naturalist will claim that we must, in fact, be committed to normative properties and objects, because we routinely quantify over them: for example, "All right actions maximize utility (or are commanded by God, etc.)." Our critics will say that unless we wish to embrace a systematic error theory, which we do not, we must acknowledge that this quantification commits us to the presence of normative properties and facts in the world.[7]

One possible reply would be to acknowledge the existence of these normative properties and try to give a reductive account of them, consonant with naturalistic theories in philosophy. We have rejected such approaches at some length in Chap. 3. Another approach would consider normative expressions with seemingly referring parts (e.g., "The foreman's cruelty inspired his workers' resentment") as idiomatic expressions which do not decompose into referring parts (i.e., "cruelty"). This approach does not seem promising. If normative sentences are non-decomposable,

idiomatic expressions, then how do we account for the validity of the inferences contained in the following passage?

> *Foreman Jones learned a hard lesson: if you are cruel to your workers, they will not be loyal to you when you need it. He was cruel to his workers. And when he needed it, they were not loyal to him.*

However, we do not think we need to concede the objection. There are accounts of quantification in the expressivist tradition within which we are working that do not rely on word–world relations to make sense of quantification. The objection relies on an "objectual" understanding of quantification. On objectual interpretations, "$(\forall x)\varphi x$" is true just in case every object in a domain is in the extension of φ. Successfully quantifying over goods, values, and so on, ontologically commits us to the existence of these objects or properties in the world. (And so, better to reduce or eliminate such items, Quineans would say.) But we see no reason to allow our ontology to be driven by our semantics in this way, rather than having our various theories of the world (scientific, moral, etc.) tell us what the world contains and having a semantic theory that can handle the richness of types of discourse and the variety of ontological commitments they embody (or do not embody). Thus, we endorse a substitutional interpretation of quantifiers that remains neutral on the ontological commitments of the terms over which it quantifies. Substitutional interpretations come in many forms, however, so the one we endorse requires some elaboration.

In its simplest form, the substitutional interpretation of the quantifier claims that "$(\forall x)\varphi x$ is true in an interpretation if φa is true on that interpretation for every term a in the language." Thus, universal quantifiers are treated as universal conjunctions: "[U]niversal…quantifiers are logical locutions that have the expressive function of making propositionally explicit conjunctive…substitutional commitments. Attributing commitment to a claim of the form (x)Px is attributing commitment to all claims of the form Pa" (Brandom 1994, 434).[8] One problem this simple substitutional theory faces is that such theories are adequate only when we have enough terms for every object that might be in the domain of quantification; but some things we want to refer to might be in prin-

ciple inaccessible. As Mark Lance writes, "specific neutrinos in distant galaxies, individual grains of sand in Death Valley, or thoughts of beings with whom we will never communicate. We can, however, quite easily quantify over each of these" (1996, p. 488). Lance's solution is to introduce the distinction between referential and arbitrary uses of singular terms. We use a singular term referentially when we mean it to refer to a particular entity, as when we say, "Bill has a headache." But we use a singular term arbitrarily when we do not intend for it to refer to any particular item. The key to having a satisfactory substitutional interpretation of quantification is to allow not just the substitution of any term in the language, but also of any *arbitrary* term. As Lance writes

> The substitutional rendering was indeed onto something in attempting to assimilate universal quantifications to infinitary conjunctions. The difficulty arose by neglecting the existence of arbitrary terms and, hence, in failing to recognize that these offer a sort of expressive resource unavailable to a language with only referential terms. The solution to the difficulties faced by substitutional interpretations of ordinary language nominal quantifiers lies in seeing that commitment to $(\forall x)\varphi x$ carries with it commitment not only to the referential substituends of this sentence, but to the arbitrary ones as well. (1996, 489)

So much for substitutional accounts of quantification. Our goal here is not to defend exhaustively an expressive account of each different type of semantic discourse, but mostly to show how such an account (already adequately defended by others) can be pressed into the service of a robust theory of normative discourse.

So in developing our response to the Quinean objection, we want to say that the fact that you quantify over a term or predicate does not imply that the term or predicate refers to an object in the world. We are committed to placing objects and properties in the natural world only if our best theories require their presence there. Paradigmatically, realistic interpretations of physical theories require the presence of real entities and properties as referents of their terms, counterparts of elements in their models, and so on. The same could be said of any theoretical discourse that purports to offer causal explanations (though there are admittedly many different views on how to interpret scientific theories, too). The

grammar of a theory may leave us with terms that appear to require such presence in parallel fashion, but we will not have to take up ontological commitments if the roles played by those terms can be articulated without the need for the presence of those entities or properties. Our claim is that the formal expressions of logic and semantics (such as quantifiers) can play the roles of articulating our commitments over domains with very different ontological implications, and that normative discourse plays expressive roles that can be articulated without substantive ontological commitments to entities and properties in the natural world.

Whether use of a noun-like expression ontologically commits us to objects in the world is not a question that can be settled all at once, for the entire language, by general semantic considerations. Rather, it is only settled on a case-by-case basis, by considerations that are shaped by the explanatory goals of the theory under consideration. Whether there are electrons in the world is a question not settled by semantics alone, but also by physics. Whether there are good or values in the world will involve an even broader set of explanatory concerns, well beyond semantic considerations alone. (To the extent that we engage in extended arguments about semantical considerations here, it is to try to demonstrate that when value theory shows that there are no such things as goods or values, to think that this means there are no normative truths is a reliance on a faulty representationalist semantic theory.) And recall that at every turn in Chap. 3, the *presence* of some property in the natural world offered us no help in explaining or understanding normativity.

This means that when we quantify over values, for example, we should not leap to the conclusion that such quantification commits us to the existence of values in the world. For example, if I say, "All goods are ultimately valuable because of their role in a deontological scheme," then this sentence need not commit us to the existence of goods. We rejected, above, the idea that such sentences should be treated purely as idiomatic; but what we would instead like to claim (following Price) is that sentences like these decompose into expressive functions, not into referential functions. In Chap. 6, we will argue for a larger taxonomy of expressive functions, and explain how to parse, expressively rather than referentially, sentences containing phrases like "All goods." But to give a preview of our argument, we will say that to talk about a good (e.g., health) is to endorse

a range of actions and outcomes (say, washing your hands before eating, having a balanced diet, getting exercise, etc.). This range of actions will itself be open-ended and open to revision. But at the same time, we do not want to say that this particular set of actions and outcomes is dictated by or oriented toward some specific property that is health, which constitutes a human good or interest. For as we will see in the next chapter, although what we end up endorsing as health will have empirical constraints, it will also depend on historically and culturally contingent aspects of social practice, and will also depend on other sorts of endorsements we make about the value of various kinds of engagements with the world. Health itself will be involved in a variety of normative entanglements and cannot serve as a fixed anchor for our theory of the good. In short, we will be able to say that goods (like health) are expressive because it is endorsements "all the way down"—we never hit a level where the simple facts about health dictate to us. There is no thing that is health, only a further set of endorsements, as we will elaborate in Chap. 5.

One possible objection, which we owe to Jim O'Shea and Niklas Möller in particular, merits attention here. They ascribe to us the view that quantification in *all* discourses should be interpreted substitutionally, not just in worrisome cases. We agree that a piecemeal solution would strike us as *ad hoc* and evasive. We also favor it out of a sense that the sorts of word–world relations objectual quantification presumes make bad fundamental semantic explainers in general. But if so, the objection goes, why worry about normative properties at all? Their appearance in normative discourse merely reflects a set of linguistic rules that play roles in the coordinating behavior with other agents (to put their role very, very broadly). Nothing metaphysically suspicious need follow from this, and the refusal to posit normative properties is an overcorrection. Again, there is a degree to which we would agree with this, as our rejection of doctrinal eliminativism might suggest. In each of the remaining chapters, we make an effort to preserve the ways in which we speak of "correctness," "cruelty," and other overtly normative vocabulary that acts much like terms for non-normative properties. Substitutional quantification itself does not entail any particular metaphysics, even if it helps with some.

However, to speak of an action being "right" or a method being "inappropriate" is to take up commitments in very different ways than we do

when we place something in the world via a theory such as physics, even if we are rejecting representationalist assumptions. Even if we take the pragmatic, inferentialist approach we have alluded to at many points, there are still intelligible questions about whether our theories require the same sorts of ontological commitments. Determining that will not be a simple matter of mapping vocabularies (as representationalism would imply), but rather a comprehensive critical look at all the explanatory strategies we employ. We seek very different things with different sorts of theoretical inquiry, and to throw all our theoretical commitments casually into the same basket would be a disservice to this variety that would flatten the landscape. In Chap. 3, we argued that to place the normative in the natural world would be to shoehorn items into the natural world that could not play suitable roles in our causal-explanatory strategies. At the same time, to pass non-normative items and explanations off in place of genuine normativity fails to capture the prescriptive dimension of the normative. Appealing to objects and properties, however we articulate those appeals, runs together very different sorts of commitments. Some distinction is needed to account for these different types of commitments, and as we elaborate in Sect. 4.2, we must *reconcile* those different types with one another in broad analyses of our theoretical projects. This requires some stinginess in doling out our ontological commitments in the name of pluralism, rather than greater license to add to our ontology across the board.

4.2   Moderate Discourse Pluralism

In some cases, "pluralism" implies what Huw Price has called "horizontal pluralism," or the view that there are multiple, equally legitimate theories or forms of discourse that all purport to serve the same linguistic task. For instance, metaethical moral relativists assert that there are multiple communities that have distinct approaches to morality, each with an equal claim to legitimacy, and that there is no community-independent perspective from which to make judgments on them. We have no desire to defend such a view here. By contrast, "[vertical pluralism] would be the view that philosophy should recognize an irreducible plurality of kinds of

discourse—the moral as well as the scientific, for example" (Price p. 36). We will endorse a form of vertical pluralism in the present sense, to be elaborated at some length in succeeding chapters. Having distinguished horizontal and vertical pluralism, we will focus on vertical forms and just call them "pluralism" for the time being. But there are stronger and weaker readings of what vertical pluralism entails. In particular, we are concerned with the basis for drawing distinctions between discursive frameworks and the degree of autonomy that this would grant each framework from the others. What we will reject here is a range of interpretations of pluralism about theoretical commitments that would preserve various forms of normative discourse by insulating one from another, without a reconciliation of the different commitments they entail for us.

The weakest forms of discourse pluralism might only suggest that there are various different types of discourse as a matter of empirical fact, though there is no reason to think that such differences are necessary or enduring. There are a number of different discursive frameworks, the weak pluralist will say, but these are contingent matters of the current articulations of our theories that should one day give way to a unified, uniform account if the world. Even a mad-dog reductionist could be a pluralist in this weak sense, since it commits us to so little. Strong forms of discourse pluralism would imply much deeper, persistent divisions and a concomitant immunity from one another's challenges. What we say in normative discourse, mathematics, physics, biology, and many other frameworks would each be an island unto itself, and ontological and theoretical commitments we make in doing, say, physics need not hamper what we say in mathematics or ethics. Each domain would be a "non-overlapping magisterium," carried out with indifference to what might be said or accomplished in others. We could not settle matters involving moral claims with the claims and methods characteristic of the discourse of the physical sciences, nor vice versa, but this is not a mark against either discourse.

A range of options might open between the strongest forms of discourse pluralism and the weakest ones, and we will adopt one in the coming pages, but first let us consider what might motivate stronger forms. Stronger forms of pluralism can be read as coupling two theses together—one we can draw from Carnap and one from Quine.[9] As we

might expect, Carnap's contribution here turns on his efforts to defend ranges of discourse that more austerely inclined empiricists and other naturalists would find suspicious. To defend the apparent appeal to abstract entities in mathematics, logic, and semantics, Carnap asserted that what we think of as a single language is in fact a number of different linguistic "frameworks," each with their own vocabularies, rules of inference, methods of taking up empirical input (or not), and tasks that drive their development. Carnap's claim was that what it would take to interpret and evaluate the truth of any given claim would be a matter of how to resolve certain kinds of questions *within* a framework using the expressions and methods *native* to the framework. Determining which sorts of objects are "real" or "exist" when it comes to the abstract entities of mathematics, logic, and semantics as though this could be settled outside of those frameworks was to misunderstand what sorts of questions we could pose. Instead, we should distinguish relevant, well-formed questions that could be posed *within* a framework from those about the viability of a framework as a whole:

> "[W]e must distinguish two kinds of questions of existence: first, questions of the existence of certain entities of the new kind within the framework; we call them internal questions; and second, questions concerning the existence or reality of the system of entities as a whole, called external questions. Internal questions and possible answers to them are formulated with the help of the new forms of expressions. The answers may be found either by purely logical methods or by empirical methods, depending upon whether the framework is a logical or a factual one." (1956, 22)

External questions in Carnap's sense are matters of whether to accept a particular range of expressions in a theoretical framework. Given the interdependence of expressions in a framework, this would often amount to an assessment of some significant portion of the framework (perhaps the whole) rather than any particular expression. It is difficult to imagine an external question about the adoption of, say, "oxygen" as a theoretical expression in the physical sciences that does not pose a challenge for all such element terms, and to the various explanatory models in which they play essential roles. This would quickly rise to the level of a challenge to

the viability of the whole theory, at least in its present form. When we imagine eliminating an item from our ontology (discovering it is *just not there*), we are usually imagining a kind of mistake that is internal to the theory and confirmed by use of the rest of the framework: for example, physicists might posit dark matter in their current astrophysical models, only to find later that its purported effects can be explained by other means.

To this, we can add the Quinean view that matters of ontological commitment are ultimately settled by the appearance of expressions in canonical versions of our best theories. Where our best theories have variables that must be read as denoting a range of entities, we are thereby committed to their existence:

> To be assumed to be an entity is, purely and simply, to be reckoned as the value of a variable… The variables of quantification, 'something', 'nothing', 'everything', range over our whole ontology, whatever it may be; and we are convicted of a particular ontological presupposition if, and only if, the alleged presuppositum has to be reckoned among the entities over which our variables range in order to render one of our affirmations true. (Quine 1953, 13)

Quine is not really a vertical pluralist in our present sense.[10] Science and other forms of discourse have no clear boundaries in his view and "knowledge, mind and meaning… are to be studied in the same empirical spirit that animates natural science. There is no place for a prior philosophy" (1969, 26). But if we join his suggestion that quantifying over something in the canonical form of a theory amounts to ontological commitment with Carnap's segregation of those theoretical frameworks from one another, we have a recipe for strong pluralism: many autonomous discourses, each with their own ontological commitments immune from the restrictions at work in other frameworks. As Price puts it:

> Once we recognize that there is more than one framework in use in ordinary language, and that there is no framework-independent stance for metaphysics, it follows immediately that the naive picture is misleading, unless specifically confined to one framework (in which case it can be

thought of harmlessly, as a metaphorical picture of what can be said in internal terms). (1997/2011, 137)

The "naive picture" in this case is the familiar naturalist ultimatum that moral discourse must either be brought into the sciences, demoted to some inferior status (merely expressive, idiomatic, etc.), or we shall plunge into some form of non-naturalism. A strong form of pluralism would imply that we face no such dilemma, simply because there is no one language game of settling ontological commitments that cuts across all the frameworks.

We share with proponents of these stronger pluralist views the sense that there is no standpoint external to our practices from which we can pass judgments on our ontological commitments. And we agree that indispensability in certain forms of theoretical discourse is all the reason we could ever ask for to make an ontological commitment. What does not strike us as plausible about these forms of pluralism is the immunity that is explicitly granted to each framework from the challenges of another. We address this in two stages: first by offering a problem for strong versions of discourse pluralism and second by offering a more moderate interpretation of its central assumption.

The problem for strong discourse pluralism is that the segregation of frameworks from one another that would permit Carnap's internal/external distinction is not the case for our various discourses, and would not be possible for any set like ours. To make that segregation possible and make the internal/external distinction viable, we must envision each framework as self-contained with its rules either explicit (or explicable), and its elements defined at least well enough to be brought under its rules. Making explicit such components of a framework may be an ongoing project, parts of which are less developed than others at any given time, and some parts of which may be assumed without defense indefinitely. Mathematics serves as a prime example of Carnap's frameworks, and one can understand why injecting concerns about empirical content into the most abstract of theories would strike him as misguided.

To preserve the immunity that Carnap's thesis purports to establish, it is essential to note that elements within a framework need an internal consistency and shared character that will make the features they con-

tribute available without introducing problems that the framework is not equipped to address. If a framework F includes some claim P, and the use of P is governed in part by its inclusion of some expression(s) that invoke inferential relations to some further set of claims Γ, then the use of members of Γ must also be governed by the rules and methods of F if the framework is to be coherent. Without that restriction, the formulation and resolution of internal questions become impossible.

In some cases, failure to meet this condition is simply an indication that the framework needs further development (e.g., elaborating the concept of an imaginary number in mathematics or the standings of a class of agents in a moral theory). That in itself poses no problem for strong discourse pluralism, though producing the needed developments may be very difficult. The more serious problem for pluralism is when the set Γ that P invokes involves expressions and commitments that are explicitly excluded or otherwise impossible to accommodate in F. For example, non-physical entities and forces cannot show up in the discourse of physics; if your proposal for an extension of our theory of quantum gravity involves positing helpful ghosts, this is a very bad thing for your proposal. You have smuggled in terms that have no home in that framework.

And for a strong pluralist seeking to preserve the internal/external distinction, strict containment will be crucial. Including elements from another framework in F will typically import the problems of the other framework into F, and F will not have resources to address them. If we place those helpful ghosts in our theory of quantum gravity, their power to affect changes will not be explicable in current laws and models and they will be a class of entities that the framework has no resources to describe. This will be a serious problem for strong pluralism, as the frameworks we have spoken of as segregated thus far actually overlap and integrate with one another in myriad ways. The framework of mathematics shapes almost everything said in physics; the laws of physics constrain possible explanations in the social sciences; normative claims inform discourse about methods in the sciences. Examples will come from every corner of our language.

There are a number of possible responses here. Perhaps the simplest is to argue that the supplemental materials in some Γ (e.g., the math needed to do physics) are simply reformulated within other frameworks. So there will be some fragment of the framework of physics—call it *math-*

in-physics—that provides those resources. But this would simply recapitulate the problem at hand within a framework, preserving segregation at the cost of coherence. Abstract objects will still be outliers in physical theories, methodological norms in ostensibly non-normative theories, and so on. The strong discourse pluralist might argue that some frameworks can be ported around while working in other frameworks as a set of background resources—a sort of discursive "toolkit" that we can take anywhere with us, but not one that has a deep impact on the framework that is borrowing from it. Thus, mathematics might be in the toolkit as we do physics, or economics, or offer a theory of justice. Three problems arise for this reply. First, it does not seem to be a good fit across the board for all the overlaps. Some features of physics or chemistry will be very deeply woven into explanatory models in biology, for instance, rather than simply popping up in incidental details. Second, the sense in which the toolkit resources are simply "in the background" is less promising on closer examination than it might seem. When a moral theorist claims that a particular distribution of resources is unjust, or that particular policies and practices are unjust because their environmental consequences harm distant or future persons, the quantitative measures and physical details are central to what is being claimed. Other numbers would be more just, as would other physical conditions, and these are not incidental details. Third, this proposal does not tell us how the apparent disparities between frameworks in the toolkit are to be resolved. Mathematics, physics, and logic would seem to be candidates for the highly central role a toolkit would play, but they already commit us to wildly different sorts of objects, properties, and methods. (And all of them have their normative side.) What it would be to bring the resources of, say, mathematics to another discourse like biology is not a trivial matter of dropping terms in on the fly. Rather, how to treat non-mathematical phenomena in quantitative terms is crucial to the articulation of those non-mathematical discourses, woven in from their very inception despite the considerable differences between them. We see these as substantial, enduring problems, and contend that there is no viable version of strong pluralism.

Fundamentally, strong discourse pluralism fails to do justice to the way in which commitments in one framework have consequences for other frameworks. Commitments must be reconciled across various frameworks.

There is a positive version of this thesis: work done in one framework (e.g., mathematics or chemistry) can make genuinely fruitful contributions to another framework (e.g., biology). Many of our examples in this section have made use of this observation. But there is also a more negative version of this thesis: rational and explanatory demands on our theories will not allow us to   make commitments in one framework that cannot be reconciled with the commitments in another framework. To do so would vitiate the explanatory import of both of the frameworks that we did not reconcile with one another, as insulating them from one another makes the commitments we make in either one optional and disposible. Where we permit ourselves to abide such incompatibilties, the force and significance of each of set of commitments is undercut. Our moral theory cannot commit us to facts about human psychology that the best scientific accounts do not permit. A biologist or neurologist cannot posit a mechanism that would violate laws within the framework of physics. And we cannot blithely posit normative properties and facts in the physical world if we are going to make causal explanation the final court of appeals in determining our ontological commitments there. So the pluralism we adopt must recognize that even if different frameworks use different tools and different vocabularies in service of different interests, the books must be balanced at the end of the day, and commitments reconciled across all frameworks.

A more viable alternative would be *moderate* discourse pluralism, the view that there are theoretically informative divisions between different types of discourse, but that the divisions do not preclude other types of integration. For a moderate discourse pluralist, divisions between different theories, etc., will be genuine, but worth noting only with certain types of metatheoretical concerns in mind; the rule will be for unity across different types of discourse. Any division serves some metatheoretical purpose for us, and there is no metatheoretical purpose that we must always pursue, or which trumps all others. As a general maxim, we can say that *we have grounds to adopt pluralism when reduction and replacement fail, but some form of discourse succeeds.* Reduction should be a familiar notion to readers; "replacement" here is the sense in which a theoretical successor may replace a predecessor without any pretense to preserving or reidentifying the predecessor. (Modern biochemistry *replaces* theories of the four humors in this way.) Many non-reductive approaches to biological, psychological, social, and many other phenomena

could thus be interpreted as types of moderate pluralism. We think that there is great significance to the latter half of the maxim—that a form of discourse *succeeds*—that has been sorely neglected. Theories in the natural sciences purport to accomplish certain things, to have goals at which they are directed and which they (hopefully) reach. The same could be said for normative ethics, mathematics, and other forms of discourse that are not empirical science. In this way, there is an interest or set of interests that we pursue when we conduct any given one of them, and any significant demarcation among them will reflect a difference in the interests pursued in each of those forms of inquiry.[11] And crucially, the goals of each discourse are not isomorphic or reducible to one another in any thick, informative sense. What we are *trying* and *succeeding* in doing when we argue against the justice of a given policy, or prove a conjecture, is fundamentally different from the goals we pursue with a causal-explanatory account.

We have reason to adopt a pluralist approach when we find different theoretical projects succeeding in the pursuit of their goals without apparent conflict, and therefore without apparent competition to serve the same interests. Biological theories do not conflict with physical ones in this sense, as no biological term will be incompatible with the laws of physics, even if the laws of physics themselves make no use of the biological terms. Biological theories are, however, directed at very different explanatory goals, giving rise to different vocabularies, methods, pedigree problems, models, and so on. "Success" in this context should not necessarily be thought of as a final, settled state of a theoretical project (we doubt there ever is such a thing), but rather the sort of fruitful, ongoing development of our best accounts in which problems are frequently provisionally resolved, even if doing so poses new and interesting ones as a result.

This gets to the heart of our aversion to positing substantive properties to account for normativity. Normative discourse, we would argue, is not properly understood as an attempt to *describe* something in any familiar sense. Thus, it can be undertaken in parallel with a plurality of other theoretical projects with which it does not compete. The sort of closed frameworks Carnap described would not be a necessary condition for a discourse, and the inclusion of more distant resources (math in moral theory, evidential norms in scientific discourse, etc.) would pose no threat at all. In some cases, the refinement of a theory might greatly

narrow permissible contributions from elsewhere. For example, what it is permissible to introduce in a chemical description or explanation is very rigorously prescribed and very restricted: no helpful ghosts, but also no just institutions, no feedforward neural networks, nor invocations of countless other legitimate concepts and expressions from other discourses. These restrictions are not formal necessities to establish the very possibility of a discourse, but rather increasingly sophisticated explications of strategies conducive to the best long-term pursuit of the interests served by the discourse. Formulating an explicit version of those interests and projects—what makes for a physical theory, or a theory of justice, or a psychological theory, etc.—is itself an ongoing project and one of the tasks of any discourse that is going to be an enduring project for us.

Some analogies would be in order here to illustrate the differences at hand. Strong pluralism would suggest that each type of discourse is an island unto itself. Things happen on each one, we might find ourselves on any one at a given time, and what happens on each island stays on the island. Moderate pluralism would suggest that discourses are more like neighborhoods or regions, each with their own character and distinct industries, but plenty of shared streets and lots of business done between them. For all their different flavors, they still interweave to give us a city. So rather than discursive or theoretical *frameworks*, we will refer to these as theoretical *discourse regions*. For some purposes, it will be fruitful to think about only one region on certain theoretical questions, but in the long run, *there is only one terrain*.

This more moderate form of discourse pluralism has many of the benefits of stronger forms. It welcomes the insight that a wide range of expressive functions might be played by the vocabularies and other resources in different discursive regions. As we will elaborate at some length in the coming chapters, it will allow us to preserve the truth-aptness of wide ranges of discourse that some naturalists have argued we must discount. But it will come at a certain cost. If our moderate pluralism is right, then worries about the ontological commitments implied by many of our discourse regions have some teeth. If normative claims are not segregated in the way that strong pluralism suggested, then we must ultimately reconcile them with their fellows in other regions. There is no insulating normative discourse from the demands of integration and consistency with physics and other sciences on a moderate discourse pluralist view. Instead, we

must articulate how normative discourse serves a distinct set of interests, and how it expresses ranges of practical commitments and entitlements. We must articulate how it expresses the guidance of our actions, in other words, but without injecting the presence of entities or substantive properties with causal roles to play. But we must do so in ways that lay bare how normative discourse is integrated with other discourse regions across the wider terrain, rather than by isolating it from our other forms of inquiry. Articulating the different expressive functions required to do this without losing the truth-aptness of normative claims will be a substantial challenge. We believe this will be worth the effort not simply because it preserves normative discourse, but because it suggests a new, fertile way of conceiving of naturalism. This would give us an account in which normative discourse was not quarantined, but rather conjoined synoptically to give us a fuller, richer story of what it is for beings like us to engage the world as we do.

Notes

1. A "realistic reading" in this sense is also compatible with certain kinds of anti-realism. Bas Van Fraassen (1976) argued *for* realistic interpretations of physical theories (speaking of their unobservable entities was purport to refer to such things, not solely a permissible move within a scientific practice) but *against* truth as a goal of scientific theories, favoring empirical adequacy instead.
2. Such theories present other challenges; for example, Michael Dummett has pointed out that Putnam has to give up on bivalence.
3. Typically, a home run is hit when "A fair ball passes over a fence or into the stands at a distance from home base of 250 feet or more. Such hit entitles the batter to a home run when he shall have touched all bases legally." (MLB Official Baseball Rules 2014, 6.09(d)). Actual application is a little more tricky, because "home run" is a distinction that is ruled by an umpire immediately in the typical case, but is also a distinction made by an official scorer in others. Thus, a typical home run is a ruling made by an umpire, but an "inside the park" home run is so recorded by the official scorer (as opposed to calling the play an error or a hit plus an error).
4. One alternative here would be approaches that were *both* pluralist *and* deflationary, but used different deflationary strategies for different

domains—disquotationalism for some, anaphoric for others, and so on. We are not aware of anyone who does precisely this, but we would think of this as a further deflationary approach and not object to it in principle.

5. Cf. (Kukla and Lance 2009, Chaps. 1–2) on "agent-neutral" and "agent-relative" normative statuses, and Price (2003) on the "third norm of truth."

6. The efficiency and precision here may also be entirely negative in any given instance, for instance, more rapidly identifying dead-end hypotheses.

7. Quantification may also tend to focus our attention on singular terms, only to neglect how predicates could commit us to normative properties. More attention to predicates follows in Chap. 6.

8. Note that Brandom is using "(x)" as a universal quantifier here, rather than "$(\forall x)$."

9. Carnap is clearly the prime example of such a strong pluralist, and we find him endorsing Quine's view on quantification, after the quote below (Carnap 1956, 42–45) though Quine will not be a vertical pluralist. Huw Price sometimes sounds like one (1993/2011), (1997/2011), though less so at other times.

10. Quine embraced ontological relativity, but in Price's terms, this would be horizontal pluralism. That is, there might be multiple theories vying to serve the same explanatory task and no external perspective from which to choose between them.

11. Much the same could be said for arts and practices that were not theoretical in anything like the senses that scientific accounts, epistemology, or normative ethics aspire to be. There are interests that we pursue in dancing, playing jazz, vegetable gardening, or the untold number of different practices in which communities engage. But our project—securing a place for truth-apt normative discourse constrained by the world without reverting to supernaturalism—is already very broad for a single manuscript, so we concern ourselves only with those that have theoretical purport for the present time.

References

Brandom, Robert. 1994. *Making It Explicit*. Cambridge, MA: Harvard University Press.

Carnap, Rudolf. 1956. *Meaning and Necessity: A Study in Semantics and Modal Logic*. Chicago: University of Chicago Press.

Grover, Dorothy, Joseph Camp, and Nuel Belnap. 1975. A Prosentential Theory of Truth. *Philosophical Studies* 27: 73–125.

Horwich, Paul. 2008. Varieties of Deflationism. *Philosophical Topics* 36(2): 29–43.

Kaptchuk, Ted. 1998. Intentional Ignorance: A History of Blind Assessment and Placebo Controls in Medicine. *Bulletin of the History of Medicine* 72(3): 389–433.

Kukla, R., and M. Lance. 2009. *'Yo!' and 'Lo!': The Pragmatic Topography of the Space of Reasons*. Cambridge, MA: Harvard University Press.

Lance, Mark. 1997. The Significance of Anaphoric Theories of Truth and Reference. *Philosophical Issues* 8: 181–198.

MLB. 2014. *MLB Official Baseball Rules.* Major League Baseball.

Price, Huw. 1993/2011. Metaphysical Pluralism. In *Naturalism Without Mirrors*, 34–53. Oxford: Oxford University Press.

Price, Huw. 1997/2011. Naturalism and the Fate of the M-Worlds. In *Naturalism Without Mirrors*, 132–147. Oxford: Oxford University Press.

Price, Huw. 2009/2011. The Semantic Foundations of Metaphysics. In *Naturalism Without Mirrors*, 253–279. Oxford: Oxford University Press.

Quine, W.V.O. 1953. Two Dogmas of Empiricism. In *From a Logical Point of View*, 20–46. Cambridge, MA: Harvard University Press.

Quine, W.V.O. 1969. Ontological Relativity. In *Ontological Relativity and Other Essays*, 26–68. New York: Columbia University Press.

Van Fraassen, Bas. 1976. The Scientific Image. *Journal of Philosophy* 73(18): 623–632.

5

Interests, Embodiment, and Constraint by the World

Thus far, we have made three central claims. First, we have argued that some form of naturalism, or at least allegiance to versions of some tenets associated with it, is a preferred methodological starting point for philosophy. Second, we have argued that efforts to solve the placement problem for normative discourse via reduction or non-reductive supervenience approaches cannot deliver convincing accounts. Third, we have argued that this sort of moderate naturalism should not lead us to abandon an understanding of our own normative discourse that is both truth-apt and constrained by the world in an important sense. Few philosophers have tried to reconcile all three of these themes. Those who say we cannot place the normative in the natural world tend to eviscerate normative discourse, replacing it with some form of projectivism or norm-expressivism (e.g., Blackburn, Gibbard), while others opt to eliminate it altogether or treat it fictionally (e.g., Mackie, Joyce). Those moved to keep a world with robust normativity often find themselves moving explicitly away from naturalism (e.g., McDowell, Wiggins), or toward supernaturalism (e.g., Adams, Craig). To many, there will be at least a whiff of an inconsistent triad in these three claims.

Obviously, we do not think that this is the case. To demonstrate how we might have these three cakes and eat them, too, we develop the idea of an interest in this chapter in anticipation of its role in an expanded notion of content for normative discourse in the next chapter. This is a term of philosophical art to some degree, in that we use it in a way that is broadly consonant with ordinary usage and informed by philosophical tradition, but also one that we will refine and elaborate in various ways. Although philosophical speculation should begin with our practices as we actually find them, this should not commit us to an "ordinary language" approach to philosophical questions in which our only role is to codify and clarify what is already implicit in ordinary practice. Philosophy, like science, is sometimes in a position to challenge ordinary practice on theoretical grounds. Moreover, ordinary usage of the notion of an interest does not offer us the depth or clarity on certain matters that are needed to address our inquiry. So we take it that our proposal is not radically revisionary, largely because we think this notion is still too loose to establish a canonical version that would require such revision. We are instead taking a fairly broad, loose ordinary notion and suggesting a refined philosophical variant to be drawn from that field of possibilities.

We choose this approach because we believe that interests (suitably elaborated) are both indispensable in the articulation of normative discourse and rooted in our embodied engagement with the world. In this way, they preserve the sort of normative discourse—robust and truth-apt—that we defended in Chap. 2 without requiring additional ontological commitments. But they are also matters that are impressed upon us and shaped by our engagement with the world in ways that compel us to answer to the world, rather than simply answering to one another. And this distinction will be crucial in defending the objectivity of some normative discourse. To the degree that a notion of interests has been introduced to recent philosophical discussion, they have often been treated as wholly within the sphere of human experience and discourse, "spinning in the void" to borrow a phrase from McDowell. On such a view, the most we can hope for in reflecting on our interests is near-universal assent, or ideal consensus in a Peircean vein. We believe this is a mistake, both in its failure to appreciate what it is to have interests and in its unduly narrow

conception of how the world might inform our judgments even when we do not seek to represent its objects and properties. That we have some interests rather than others is a matter impressed upon us by the world, and how to articulate and best pursue those interests are questions on which our answers can be constrained by the world (even if our interests are not entities in the natural world).

5.1   Interests and Engagement with the World

We will begin by remarking on a widely discussed fact about normative discourse, namely, that it is difficult to find uncontroversial examples of normative facts being ineliminably invoked in causal explanations. In Harman's venerable example, suppose we round the corner and see some hoodlums setting a cat on fire. The best explanation for why we form the judgment, "That is cruel," does not invoke cruelty or posit moral properties. Certain psychological facts enter into the equation, but, Harman argues, cruelty need not.[1]

Harman uses the causal impotence of normative facts to argue for moral nihilism (or at least a watered-down version of moral nihilism). If one has a prior commitment to the claim that the only legitimate discourse is causal-explanatory discourse (and that other sorts of discourse whose purpose is practical, rather than descriptive or explanatory, cannot be fully legitimate or objective), then one will be an error theorist about normative discourse to the extent that one discovers that normative discourse is not in the first instance a descriptive or explanatory discourse. Such a view of the function of legitimate discourse-types stems from a representationalist view of language, according to which language has content in virtue of representing the world (via, for example, referring to objects or factual states of affairs).

But again, we can make the case that normative discourse serves an expressive, rather than a causal-explanatory, role, and hence can be legitimate even if not causally efficacious. Sellars notes that normative facts only enter the causal order as the object of intentional states. Thus, he writes, "obligation enters into the causal order only as an element in the

intentional object of a mental act," that is, "via facts of the form *Jones thinks (feels) that he ought to pay his debt*" (1953, 222). However, as we have repeatedly urged, ontological eliminativism (that some range of entities does not exist) does not entail doctrinal eliminativism (that some range of entities does not exist *and* that there are no true doctrinal claims involving those terms). Sellars writes,

> [O]nce the tautology 'The world is described by descriptive concepts' is freed from the idea that the business of all non-logical concepts is to describe, the way is clear to an *ungrudging* recognition that many expressions which empiricists have relegated to second-class citizenship in discourse, are not *inferior*, just *different*. (1957, 282/§79)

We have staked out a position according to which normative discourse is not representational discourse, but rather serves a practical and expressive role. We are thus not committed to normative properties (at least, as we noted earlier, to substantive properties). In staking out such a view, it is incumbent upon us to respond to a number of worries about the possibility of truth and objectivity of normative discourse, if it is not ultimately representational. But these worries are not an odd assortment of unrelated complaints; they are rooted in a common set of concerns. One concern (discussed previously) is that if normative discourse is not causal-explanatory, then it may appear theoretically isolated, not connected to the main body of our theory, and hence (by our *own* lights) theoretically suspect. A second, related worry is whether by endorsing a social practice account of normative discourse, and denying that normative discourse is representational, we have made it impossible to give an account of how the non-normative world constrains our normative discourse.

As noted above, most of this book will be taken up showing how a social practice, expressivist account of normative discourse can address these worries, and related worries about the objectivity, truth-aptness, and other important features of such discourse. We will add the final details in Chaps. 8 and 9, where we demonstrate robust contributions from various non-normative discourses to different varieties of normative discourse. By showing how non-normative discourse contributes to nor-

mative discourse, we can show precisely how the former constrains the latter, even if normative discourse is not, in the final analysis, a species of causal-explanatory discourse.

But let us emphasize why we think some of these concerns, at least, are misguided. In particular, the concern that our normative practices could spin free, unconstrained by empirical inputs, is motivated by two assumptions that we reject as betraying a fundamental misunderstanding of the nature of the normative. The first assumption is that the mental and the physical are separate ontological realms in principle, and that it is unproblematic to imagine engaging in a normative practice that was disengaged from any physical activity altogether. The second assumption (which is more Kantian that Cartesian) is that rules can be understood formally, prior to their instantiation in practices. These two views—dualism and intellectualism—encourage the view that the normative can spin freely. Intellectualism encourages us to think that our grasp of the normative can consist merely in an understanding of rules, separate from practice—that one could possess "knowledge that" without any "know how" (and, indeed, that the former is prior to the latter in the order of explanation). Dualism further encourages us to think of intellectual activity as something in principle separate from our existence as embodied creatures. Together, these assumptions allow us to paint a picture of a mind, manipulating a system of rules that are cut off from any embodied practice (and possibly from any *body* at all!), not engaged with empirical reality, and unconstrained by how the world actually is.

We reject both of these assumptions. The view that our normative practices could be carried on in a way that is fundamentally disengaged from the world ceases to make philosophical sense. First, norms can only be understood as in the first instance implicit in a variety of practices. Explicit utterances of norms can only be understood as making explicit what was already implicit in practice. To reverse the order of explanation is to make it impossible to understand how we could grasp and follow norms, as authors from Wittgenstein to Brandom have pointed out. From the scientist in her laboratory to the man reading the news in his home, we are all bound by epistemic standards, by norms of evidence which govern the formation and sustaining of beliefs. In our interac-

tions with each other, we are bound by norms of moral behavior, usually unacknowledged, but implicit in the behavior we accord each other. And every time we read, speak, or even think, we are implicitly following the norms of our language, semantic, and syntactical norms—norms which most of us cannot even fully state.

One can explicitly state (and argue for) a norm that is not already implicit in practice. Thus, one living in a slave state can argue for the abolition of slavery; a scientist can argue for double-blind studies as the gold standard for medical research (before the wide adoption of this practice); the Académie française can argue for the adoption of the word "courriel" (instead of the foreign borrowing "e-mail"); and so on. But such moves can only be understood against (and as parasitic upon) a background of norms which are implicit in moral, epistemic, semantic, and other practices.

Second, the Cartesian picture encourages us to think of the mind as something that could carry out various activities—thinking, rule-following, norm-abiding—in splendid isolation, without engaging anything at all: not a body, much less a world beyond a body. But this strict separation cannot be maintained. To be sure, there is a difference between the mental and the physical—to talk of mental phenomena is to invoke the normative, whereas to invoke physical phenomena is to invoke the causal. This is not to say that the mental is, at bottom, non-physical; it is only to say that we cannot describe it only using the language of the physical. But we are embodied, part of the natural world, and our lives and practices are carried out in a way that is fundamentally enmeshed in the world.[2]

The rejection of these two assumptions (intellectualism and dualism) renders it difficult, if not impossible, to understand our normative practices as entirely isolated from the empirical, freely spinning without any input from the world. Thus, it follows that norms essentially arise in the context of practices, and these are practices that we necessarily engage in as embodied creatures. There is no question of a separate, inner (mental) realm of normative activity. Our practices, which are in the world, are the primary locus of the normative, and indeed we must explain the normative in terms of these practices, and in terms of our practical embodied engagement with the world.

It is worth citing, at this point, Herbert Simon's widely quoted example of an ant tracing an erratic path across a beach: The "complexity [of the ant's path] is really a complexity in the surface of the beach, not a complexity in the ant" (1969, 64). Simon, in his discussion, draws the more general conclusion: "An ant [A man] viewed as a behaving system, is quite simple. The apparent complexity of its [his] behavior over time is largely a reflection of the complexity of the environment in which it [he] finds itself [himself]" (1969, 64–65, emphasis removed from the original). The lesson we draw from this is that we cannot understand human behavior except as embodied and in practical engagement with the world. One cannot imagine a system of norms as a set of rules, arising not out of any practice, but as a pure exercise of the intellect, because this fundamentally misconstrues the relation between norms and practices. We can only understand norms in terms of practices, and we can only understand human practices in terms of their engagement with the world (just as we cannot understand the ant's path without reference to the beach).

Rebecca Kukla and Mark Lance (2014) are particularly eloquent on this point. As they note, many pragmatists emphasize that our practices are embodied. But noting that our practices are embodied in no way captures the ways in which our normative practices fundamentally arise out of our engagement with the world. As they write (when discussing the practice of food preparation, and the various norms and activities which constitute it),

> [I]magine trying to move your fingers in the way one does when shelling peas, without peapods being present. The skill is not one of moving hands through airspace while contingently sometimes doing so in the presence of a pea-pod, but rather a skill of responding to the mass, volume, inertia, and structural resilience of the pea-pod. (2014, 25)

Thus, *merely* pointing out that our practices are embodied falls far short of capturing the inherently world-involving nature of our practices.

In its most radical form (a form we need not endorse for the purposes of this book), rejecting a sharp dualism of world and agent takes the form of the extended mind hypothesis, which holds that

human cognitive processing literally extends into the environment surrounding the organism, and human cognitive states literally comprise-as wholes do their proper parts-elements in that environment; in consequence, while the skin and scalp may encase the human organism, they do not delimit the thinking subject. (Rupert 2004, 389)

Anthony Chemero (2009) traces the origins of this hypothesis to James and Dewey. In James's essays on radical empiricism, he was keen (among other goals) to expel neo-Kantian dualism. This is not the traditional dualism of non-physical mind and physical body. Rather, this is a dualism of experiencer and experienced, consciousness and content, according to which "experience is indefeasibly dualistic in structure," (James 1904a, p. 478) displaying an object plus subject dualism. This is a theme echoed in "A World of Pure Experience," where James writes,

The first great pitfall from which such a radical standing by experience will save us is an artificial conception of the relations between knower and known. Throughout the history of philosophy the subject and its object have been treated as absolutely discontinuous entities; and thereupon the presence of the latter to the former, or the 'apprehension' by the former of the latter, has assumed a paradoxical character which all sorts of theories had to be invented to overcome. (1904b, 538)

Dewey decries a Cartesian dualism of mind and matter, which portrays the mind as "a separate and isolated mental world in and of itself, self-sufficient and self-enclosed" (1925/1958, 15). Such a picture of the mind "has on its hands the problem of how it is possible to know at all; how an outer world can affect an inner mind; how the acts of mind can reach out and lay hold of objects defined in antithesis to them" (Dewey 1929/1958, 15). For Dewey, of course, the paradigm instance of intelligent action is skillful performance, in the world; intelligent performance is worldly and world-involving:

In well-formed, smooth running functions of any sort—skating, conversing, hearing music, enjoying a landscape—there is no consciousness of separation of the method of the person and of the subject matter. When we reflect upon experience instead of just having it, we inevitably distinguish

between our own attitude and the objects towards which we sustain that attitude…reflection upon experience gives rise to a distinction of what we experience (the experienced) and the experiencing—the *how*…This distinction is so natural and so important for certain purposes, that we are only too apt to regard it as a separation in existence and not as a distinction in thought. Then we make a division between a self and the environment or world." (1916/2008, 173)

Later writers, influenced by this pragmatist tradition, have emphasized the extent to which intelligence and cognition are world-involving. Arguing for the extended mind hypothesis, Clark and Chalmers ask us to

consider the use of pen and paper to perform long multiplication (McClelland et al. 1986; Clark 1989), the use of physical re-arrangements of letter tiles to prompt word recall in Scrabble (Kirsh 1995), the use of instruments such as the nautical slide rule (Hutchins 1995), and the general paraphernalia of language, books, diagrams, and culture. In all these cases the individual brain performs some operations, while others are delegated to manipulations of external media. (1998, 8)

Wagman and Chemero cite voluminous experimental data which seem to suggest, as they write, that "a hand-held tool is experienced as an extension of the body" (2014, 115) and approvingly paraphrase Merleau-Ponty's claim that "a blind man who is adept at using a cane to navigate does not perceive the cane, but the world at the end of the cane" (2014, 115). In addition to James, Dewey, and Merleau-Ponty, Chemero also mines Heidegger for insights into the ways in which external objects are fundamentally incorporated into intelligent action. As Dotov, Nie, and Chemero write,

Most human activity, Heidegger argued, is absorbed, skillful engagement with entities in the world. When we are coping skillfully with the world, we experience entities around us as ready-to-hand. To use Heidegger's example, a hammer is encountered ready-to-hand, as a piece of equipment, when it is being simply used to drive in nails. Our engagement with entities ready-to-hand does not involve explicit awareness of their properties; instead, we "see through" them to the task we are engaged in. When we are

smoothly driving in nails with a hammer, our focus is on the thing we are building not the size or shape or color of the hammer. (2010)

Only in a "breakdown" situation—where things do not function as they ought—do we focus on the hammer as an unready-to-hand object. These authors argue that ready-to-handness is a phenomenon that can be verified experimentally. Wagman and Chemero interpret ready-to-handness in terms of 1/f scaling:

> 1/f scaling occurs when the components of a system are so tightly integrated with one another that they cannot be understood independently (Bak et al. 1988; van Orden, Holden and Turvey 2003). Systems with this type of tight integration are often said to exhibit *interaction-dominant dynamics*. This technical term can be read quite literally: a system exhibits interaction-dominant dynamics when the interactions among the components dominate or override the dynamics that the components would exhibit separately. (2014, 118)

Not only do many cognitive systems display 1/f scaling (meaning the elements are not modular, and their functioning can only be understood as an organic whole), but some components of such 1/f systems extend "beyond the body periphery" (2010, 118), meaning that the cognitive system is comprised in part by items outside the human organism. Dotov, Nie, and Chemero claim to have demonstrated this by having test subjects play a simple computer game using a computer mouse, a tool with which most contemporary subjects are familiar. During normal play, the human-mouse system displayed 1/f scaling, suggesting that the mouse is actually a constituent of the cognitive system. (By extension, Dotov, Nie, and Chemero imply that generally speaking, a ready-to-hand tool is a constituent of a human cognitive system when it is being skillfully used.) But the computer program was designed to periodically and temporarily disrupt the function of the mouse, which lead to the kind of "breakdown" situation resulting in the mouse's being regarded as unready-to-hand. And as predicted, the human-mouse system did not display 1/f scaling during these periods.

Of course, the extended mind hypothesis is extremely controversial. The more general insight we wish to emphasize—and this is a thought that is present, in various degrees, from James and Dewey, to Heidegger, to Lance and Kukla—is that we cannot make sense of intelligent action, much less normative practices, in isolation from engagement with the world. Intelligent action is, in the first instance, action engaged with the world. To treat "pure" thought, disengaged from the world, as though this was somehow fundamental in the order of explanation, or the paradigm of intelligence, is, as Dewey puts it, like assuming "that a hibernating bear living off its own stored substance defines the normal procedure, ignoring moreover the question where the bear got its stored material" (1925/1958, 278). We take it that this insight—which stands in opposition to the twin errors of intellectualism and dualism outlined above—is a defining feature of pragmatism, and one of its anchoring points. Thus, we align ourselves with the pragmatist movement identified by Lance and Kukla as having its

> roots in the classic American Pragmatists such as Dewey, James and Peirce, and often also in the early work of Heidegger and his French successor such as Pierre Bourdieu and Maurice Merleau-Ponty. This group has productively focused on embodied practice as the ineliminable site of human meaning. (2009, 3)

5.2 Normativity, the Natural, Entanglement, and Layer Cakes

Thus, even though we do not hold that there are substantive normative properties, normativity can only be understood in terms of practical engagement with the world, and with other agents. The various ways in which the natural and the normative are interrelated will be developed in some detail in the final three chapters of the book. But we can start spelling out the relation in a schematic way in this chapter, with an eye toward (a) allaying some of the concerns our account might have given rise to, and (b) laying the groundwork for the later chapters.

One way to understand our view of the intimacy of the connection between the normative and the natural is by contrasting it with a competing view of this relation. This competing view holds that we can separate the world into a factual substrate, and a normative superstructure. These philosophers adopt what we will call (borrowing from Marc Lange 2000) a "layer cake" view of the relation between the normative and the natural, according to which the former sits on top of the latter like the top layer of a cake. On this view, the natural explains the normative (via the normative being composed by elements of the natural, or supervening on them). Although some normative perception is surely psychologically non-inferential, it is possible in principle to imagine cognition of the world in purely non-normative terms. That is, we can imagine cognizing a purely non-normative world, upon which normative properties are layered as emergent or epiphenomenal properties. In its strongest form, this picture takes the form of a Humean or neo-Humean projectivist theory, according to which we encounter a world which is wholly devoid of normative significance, and then project a set of values onto this world.

This layer cake picture of our engagement with the world is flawed. In the first instance (and this criticism is a direct descendant of McDowell's criticism of Blackburn's projectivism—see McDowell (1981)), it is a mistake to think that we can "disentangle" (McDowell's word) important normative concepts into a factual component and a normative component. On the picture described in the previous paragraph, we encounter the world as a factual landscape, and then upon certain groupings of objects and actions, we project values. But this picture presupposes that we can group together the various objects and actions into the extension of the normative term (e.g., "cruel") prior to the projection of the term, and then project this value or norm onto this previously cognized extension. But, as McDowell argues, it is simply not plausible that all (or even most) normative terms have an extension that can be comprehended independently of the relevant normative concept. That is to say, *the normative concept must already be in play for us to be able to grasp the relevant extension; we cannot grasp this extension prior to the normative concept, and the concerns and interests embodied in it*. Thus, valuing *in principle* cannot be a two-step process, which proceeds from cognition of a value-free world to the addition of a value-laden "layer" on top. The cognition of

various natural features of the world and the cognition of these features as normatively salient (or their incorporation into various normative judgments) are, as it were, equiprimordial.

The layer cake view also fails to take into account the way in which the normative is "baked into" our practices. Non-normative items are not devoid of normative import to us prior to our taking up interests, because there is no position "prior" to having interests, even in an ideal constructivist sense. That is to say, there is no useful sense in which we encounter a world prior to having interests (and hence taking up normative terms and concepts), or conceptualize a world in an interest-free way. As noted in Chap. 2, we endorse the broadly *pragmatist* view that "we do not encounter reality in the stance of a cognitive subject, but rather we practically cope with the world in such a way that it is given to us as a field of practical significance" (Honneth 2008, 30). Gibson famously writes that we perceive the world in terms of affordances, "action possibilities," a view of cognition that is fundamentally out of step with the idea that we cognize a world of pure fact, with no relation to our own position in the world as agents. The way we see the world is in the first place conditioned by our concerns. A person does not see the crazed ax-murderer dispassionately and then (based on her desire not to be slain) run away. She *sees* the situation as dangerous; her very conception of the situation is bound up with her concerns, and this conception explains why she acts as she does. We encounter a path *as* a path because of our interest in getting from point A to point B. So, we are sympathetic with John Haugeland (a Heideggerian, even if he would not describe himself as a pragmatist), when he laments the excesses of Cartesianism:

> Among Descartes's most lasting and consequential achievements has been his constitution of the mental as an independent ontological domain. By taking the mind as a substance, with cognitions as its modes, he accorded them a status as self-standing and determinate on their own, without essential regard to other entities. Only with this metaphysical conception in place, could the idea of solipsism—the idea of an intact ego existing with nothing else in the universe—so much as make sense. And behind that engine have trailed the sorry boxcars of hyperbolic doubt, the mind-body

problem, the problem of the external world, the problem of other minds, and so on. (1998b, 207)

What Cartesianism and the layer cake view have in common is that they view our relationship with the factual world as one in which we are separate spectators, who observe (and to be sure, interact), but are in a fundamental sense *at a distance*. Only on this view of our relation to the world does the idea of projecting values onto the world make sense as a story of the way in which valuing occurs. But this picture mistakenly presupposes that we could have a set of interests *separate* from our engagement with the world, and then project this interests onto the world in an act of valuation. But our interests only exist as embodied in the natural world, and not apart from it. For example, our interest in being healthy essentially depends on our concrete embodiment, and the way this interest manifests itself (and the way in which we promote this interest) fundamentally depends on the way the world is. There is no such thing as holding this interest prior to being embodied and embedded in the world, and then projecting it onto the world conceived without prior interests. This is not to deny that we can entertain interests we do not have in a hypothetical fashion, or look upon our world while setting aside some that we do. But the layer cake picture suggested something stronger than that: that we can and do start without embodied interests in a fundamental way, adding them only later and only in an involuntary fashion.

It might be tempting to suggest a kind of hierarchy or priority of interests that would emerge from Cartesian subjects, concerned only with the factual before turning to interests and normative language and concepts. If we were pure cognizers in the Cartesian mode, we might think of an interest in acquiring knowledge as the first, most general interest to be projected from that underlying rational nature. But this seems highly implausible, even as an idealized reconstruction of our taking up the interest. A person chosen at random might have no reason to pursue knowledge about metallurgy, or the composition of Neptune's atmosphere, or William Shakespeare's education, or the sociology of college fraternities. Her interest in acquiring knowledge is specific to actual problems and concerns, not knowledge in the abstract. A person might have a specific

interest in acquiring knowledge about the life of Admiral Nelson, how to grow a tastier tomato, this season's *X-Factor* contestants, how to lose weight, and so on. Talk of "big-picture" interests (like knowledge) does serve a particular purpose, as we shall see in Sect. 5.3, but this is primarily an expressive purpose, which allows us to relate different sorts of activities together, to help make rational sense of different communities who have different practices (but which we can see as serving interest we recognize), or to organize our pursuit of more specific interests in a more systematic way. The larger point is that in general, people do not pursue these larger interests—Smith does not read about the New Horizons space probe's exploration of Mars with the thought that this will serve her general interest in acquiring knowledge. Rather, Smith does so because it serves her more specific interest. And of course, you cannot understand the emergence of her specific interest in the make-up of Pluto, or planetary exploration, and so on *independently of how the solar system, and hence the larger world, actually is*. Thus, people's activities are organized around the pursuit of specific interests, and these specific interests can only be understood by reference to the way the world is; they are not things that can be understood independently of how the world is, and then projected onto it. The layer cake picture of normativity, where there is a factual level on bottom, and a normative level above, cannot properly account for the entanglement of the normative and the factual.

We do not wish to deny that it makes sense to distinguish between the non-normative and the normative, to talk about causal-explanatory discourse regions and normative discourse regions. Given this, though, it is not plausible to think that we cognize the contours of some factual property (say, cruelty), and then project value onto that. It is not even plausible to think that the property of cruelty forms a kind recognizable in isolation from the values and interests embodied by the normative conception of cruelty; this kind of sharp separation between the factual and the normative is not tenable.

But does not this view—our rejection of the layer cake view of the relation between the normative and the natural—raise a difficulty? We want to deny substantive normative facts and properties; but at the same time, we are arguing that the primary locus of the normative is in our practical engagement in the world, that the identity of some items (particularly

artifacts) is inherently bound up in a network of norms, and that where our normative practices engage the world, there is no way to disentangle the normative and non-normative elements. Do not these concessions commit us to a more complicated metaphysics? Do not they commit us to admitting that there are norms out there in the world, and that an expressivist view of the normative is not adequate to the phenomena it hopes to explain?

We will flesh out our social practice account of constraints on normative judgments by the natural world in much greater detail in Chap. 7, but for now, let us show how some details of this account let us address the worry outlined above. Language is an inherently interpersonal affair; only by taking the social (as opposed to the purely personal) view can we understand the inherently normative aspect of language. To take Wittgenstein's well-known example, following a sign-post cannot be explained as an instance of my following a rule, because the rule would require an interpretation, which then leads to Wittgenstein's famous regress. Nor (and here is the part that concerns us) can it merely be a matter of doing as we are trained to do, although this is certainly part of the story. For in saying, "In following the sign-post, we are doing as we are trained to do," we are only giving the causal side of the story—and just telling the causal side of the story does not express fully the idea of following a sign-post (1953, 80/§198). What is missing? This: "a person goes by a sign-post only in so far as there exists a regular use of sign-posts, a custom" (1953, 80/§198). What we should draw from the example is that there is no spooky normative property inhering in the sign-post; nor need there be any normative property at all in the sign-post. What there is a social practice—which consists not merely of a group of people trained to respond to the sign-post in a particular way, but which also contains a meta-practice in which such object-level attitudes and responses can be made explicit, and made fodder in the game of giving and asking for reasons. The same is true for any other item which we engage normatively. Even though artifacts like hammers cannot be identified as such independently from a set of norms, that does not mean that the norms inhere in the object. Rather, this means that such artifacts have their identity as part of their place in a web of social practices, and their identity—say, as a hammer—fundamentally depends on their place in this network. Again, a hammer is not just a

thing with a particular shape, or size—it is something that is to be used for pounding nails, and it has this identity in virtue of the surrounding practices. There is nothing spooky in this; nothing that commits us to a complicated metaphysics, or to normative facts or properties inhering in objects.

But has this view of things driven us back to the layer cake view of the normative? By giving the above, social practice account, have we admitted that we engage with a norm-free world, and our practices impose norms on this world? No; for it does not make sense to talk of engaging with a world prior to our practices, at least not a conceptual and rational engagement that might serve as the base of a layer cake. (So non-human cognizers and even non-cognizers engage *in a much looser sense* with the world around them without practices, but not in a fashion that provides a basis for the conceptual and normative resources that we deploy.) Separating our social practices from the physical world is like trying to separate sound from the air (or other physical medium which transmits it); it fundamentally misunderstands the essential involvement of the latter in the former. Our only engagement with the world is via our practices in the first place, and hence we encounter a world that has the various practical significances it does in virtue of the ways in which elements of this world are incorporated into our practice. But the point that we can only engage the world (or at least certain elements of it) as normatively imbued is a conceptual or epistemological point, not a metaphysical one. It does not entail that normativity must now inhere in the non-social world. For example, we argued above (borrowing from McDowell, and echoing an earlier argument Sellars made against phenomenalism) that the factual and normative elements of normative concepts like "cruel" cannot be disentangled. That is, it does not make sense to think we can independently cognize the factual extension of "cruel" (i.e., independently of the normative concept *cruel*), and then impose or project normativity onto this factual extension. The extension cannot be understood apart from the concept, and the interests and values it embodies. Thus, we must learn the concept and the extension at the same time. But to say all of this is to make a point about the concept *cruel*, not to make a metaphysical point about a substantive property of cruelty in the world. Thus, we need concede neither the layer cake view of norma-

tivity, nor the view that normativity inheres in objects independently of our practices and interests.

(An important note: We are speaking loosely of "extensions" here. As we have noted repeatedly, we do not propose to commit ourselves to anything more than mere aggregate properties, and we do not take the extensions to have any explanatory force for our account. The very shapelessness of the "extension" of a concept like cruel makes it implausible that it forms anything like a substantive property at all, in any ontologically useful sense. "Extension," for us, means no more than "the class of actual and possible items and events that we would consider in licensing application of the term 'cruel.'" It is doubtful that the membership of this class would even be definitely defined for a large number of cases.)

Kukla and Lance tie this question ("How do objects in the world come to have normative significance for us?") to the earlier question we discussed ("How are our normative practices not merely embodied, but essentially world-involving?") To attribute natural objects with some kind of inherent normativity would be embrace anti-naturalism; but how else to make sense of this normative import? As noted above, we engage in normative practices as embodied creatures, but that is only half the story:

> [E]very normative practice is something we do with our bodies. But what our bodies do is manipulate and negotiate the material world. Hence our normative practices world-involving; material things and features play integral roles in the pragmatic structure of these practices and therefore take on normative significance from within the practice. (2014, 24–25)

Thus (to use their example), the fact that it is raining may mean that it is time to harvest the grapes, or to end the baseball game. But the rain has no normative import apart from our (socially constituted practices) of wine-making, sports, and so on:

> [R]ain has concrete normative significance from inside these practices. The rain need not 'tell us' anything or 'hold us' to anything. We are the ones who institute, maintain, and practice the norms of vinification, baseball, fashion, and so forth. But we cannot do this except as embodied beings who engage with rain and its absence; within such engagements, rain has specific normative meanings and consequences. (2014, 26)

Thus, we can hold (without resorting to some kind of objectionable anti-naturalism or non-naturalism) that objects and events can have normative import—but they do so only within the context of our (essentially world-involving) practices. Of course, by arguing that the normative is instituted by our social practices, we leave ourselves open to the old accusation of relativism, which seems to dog social practice accounts of normativity. We will address this objection at some length in Chap. 7.

A fundamental nexus in this natural-normative connection will be our interests. Interests entangle the natural and the normative in a number of ways: our interests (e.g., our interest in health) are fundamentally shaped by the natural (i.e., by our natural environment, by the ways in which we are physically embodied, etc.); our interests are connected to purposive action; and so on. Thus, one step in connecting the normative to the natural will be in giving an account of interests on which they are not psychological states, or reducible to psychological states (or to any other set of natural facts). As with other features of normative discourse, talk about interests serves an expressive role (about which we will say more in Chap. 6). But let us begin by saying a bit about the role interests play in explaining and justifying action.

5.3 Varieties of Interests

Central to our approach here will be appeals to the interests of agents who can engage in normative discourse and self-correction. Broadly speaking, we are creatures to whom things can matter, and thus we have interests in how they turn out. To ascribe interests to ourselves (or to others) is to give expression to the fundamentally goal-directed character of our activity, which is deeply woven into the world in which we are embedded, as we described in the previous section. But much like our view that we should not reify normative properties, we will make a similar case that we should not reify interests. While they are essential to every form of discourse and activity in which we engage, and play an essential role in the objectivity of normative judgments, we should resist the temptation to place them in the world as causal explainers.

5.3.1 Interests as Pragmatic, Rather Than Psychological

In line with our general expressivist strategy here, we deny that interests are items that will be added to our ontology. Although we may, as a matter of shorthand, speak of interests, we think it is more enlightening to look at the expressive role of interest-talk. We will argue that to ascribe an interest is to take a normative stance—to privilege some goal or end, and to prioritize actions or conditions conducive to this end. Thus, having an interest in confirming the Standard Model entails privileging this end, and also undertaking a commitment to perform various sorts of actions and behaviors in service of this end. While interests are fundamentally world-involving (one cannot have an interest in confirming the Standard Model, or harvesting the crops before the rain comes, independently of how the world is, and independently of our concrete, embodied engagement with the world), interests are reducible neither to states of the world, nor to psychological states.

To the degree that an account makes goal-directed human action central theoretical elements (as we do), it will need a notion of interests. Our interests are fundamentally conditioned by how the world is, and by our conditions of embodiment. Thus, having an interest is never solely a matter of personal or communal assent; the world constrains our having of interests. These interests, in turn, constrain the shape of our normative practices. Much of the pragmatist tradition has emphasized the importance of interests to our engagement with and conceptualization of the world, but this has often been at the expense of attention to how the world with which we engage shapes the interests that we adopt. Inattention to this latter theme allows normative discourse (and on some views, *all* discourse) to lapse into an idealism where our claims are left "spinning in the void."[3] But the conditions in which we find ourselves embodied are not so mutable, nor so optional. For beings like us, the conditions under which we are embodied will place us in various positions of practical urgency, and the interests we undertake to address by our actions and practices are the complement of those conditions.

To ascribe an interest (to ourselves, or to someone else) is a way of characterizing our practical engagement with the world—of adopting a practical attitude of endorsement toward something (baseball, harvesting the crops on time, ending human trafficking, etc.). To speak of an interest is a way of endorsing a certain behavior or outcome, but crucially, can also be a way of coordinating and projecting behavior into the future. For example, we may consider knowledge to be a good, but what specific bits of knowledge are deemed worthy of pursuit depends on our interests more generally: is it *worth* knowing how many words Borges wrote during his lifetime, or how many angels can dance on the head of a pin? Within a specific discipline, the interests driving that discipline will determine what questions are pursued and which ones are not: an electrical engineer will not (in his or her professional capacity) pursue questions concerning, say, the best way to maintain pH balance in a saltwater aquarium, or who will dominate next year's Paris fashion scene, although knowledge concerning these questions may be a legitimate interest that others have.

Ascriptions of interests, what it is to pursue them, what succeeds, and whether we can abandon them are all matters partially determined by the conditions in which our practices are embodied. The sort of constraint we have in mind here is not causal or descriptive, and it does not require the presence of an interest-as-entity to causally interact with us. (It is not a force acting upon us like gravity or electromagnetism, nor even intimidation from other agents.) But the constraint involved is substantial enough that worrying about our interests will have a kind of practical necessity for us and the only resolution of it will compel us to attend to the conditions in which we are embodied. Speaking of experience, Kant warned us that thoughts without intuitions are blind, while intuitions without concepts are empty. There is a parallel to be drawn here with interests and the conditions of our embodiment. Interests considered in the absence of the conditions in which we are embodied would leave us with the sense that they are arbitrary and somehow lacking objectivity. But to attend only the physical conditions and causal details of the world—stripping out any sense of us as *directing* our actions toward interests—is to lose sight of what it is for us to inquire about our world in the first place. Appeals to interests as we conduct them will therefore not be matters on which resolution is solely a matter of social assent and collective choices

about what to value, as someone like Rorty might suggest. Nor are they matters of causal determination by forces external to our social practices. These two prevailing views on the nature of normativity are two sides of the same coin, but we should embrace neither one nor take them to exhaust the possibilities.

That we have an interest in health, for instance, and that it takes the form it does, depends on the fact that we are embodied in the precise ways that we are. What precisely constitutes health is a question that is addressed from within our practices—one that cannot be isolated from the conditions in which we are embodied, but one that also cannot simply be read off of the features of the world. To take an example to be discussed in Chap. 6, homosexuality was once considered in the West to be a mental *illness*, but has not been so, at least in the USA, since the mid-1970s. Thus, while facts of biology are going to place significant constraints on how our interest in health is structured, the facts of biology do not simply determine this interest in the absence of other normative considerations. How to structure our practices to serve an interest such as health can only be determined in relation with other practices (such as economics) and the interests served by these other practices, along with all of the involved worldly conditions (from how easily various hormones used in medical treatment can be synthesized, to how costly it is to extract the materials used in hospital construction from the earth). What it means to be healthy, and thus to pursue health, reflects our further commitments to what sorts of levels of activities people can and should engage in—many types of practical engagement entail leaving one's home and interacting with others, for instance, and thus impairments to our mobility, communication, impulse control, or cognitive capacities to plan courses of action would all be detriments to our health. That some count as diseases and disorders and others do not reflects standard conditions for embodied creatures like us. An inability to fly is not a health problem, but an inability to walk might be. But for beings with other very different practical concerns, a very different articulation of their interest in health may emerge. There is a matrix of entanglement and interdependence between interests, the practices by which we pursue them, and the conditions in which these practices are embodied. How our practices should be structured to serve our inter-

ests (and how we should conceive of our interests in the first place) is an ongoing discussion, shot through with fallible commitments, and bearing holistic dependence not just on various worldly conditions, but also on other practices and sets of interests.

However, the having an interest (or our ascription of it to someone else) is not always a matter of ascribing some thought or conception of that interest to an agent or the members of a group. It is a *normative* status, not a psychological one. As a normative status, it *may* be acknowledged by someone, and acted upon, but this acknowledgment does not *constitute* having the status. One might have an interest and never reflect upon it, or never be motivated by it, nor even realize that one has it. To ascribe to Jones an interest in being treated with dignity is not to *describe* Jones (Jones may be suffering from dementia and incapable of ascribing interests to herself); rather, it is to take a normative stance about what matters to her. Making such a case would require us to appeal to her aversion to suffering, even where she might lack the conceptual resources to organize her own behavior, but her deficits would not be disqualifiers by themselves. The same goes for infants and other adults with severe cognitive disabilities, and maybe about non-human subjects that can suffer but cannot represent and reflect on their representations of the world and themselves. One can incorrectly *deny* that one has an interest, though doing so will often require some epistemic perversity.[4] The seriously ill patient who refuses to seek medical help out of sheer willfulness, or those who deny climate change out of contempt for any scientific challenge to their practices and preferences, are all denying that they have certain interests, but their subjective psychological states do not settle the matter.

This picture of interests allows us to draw together several threads from this and earlier chapters. Let us frame this discussion from an intriguing (but potentially misleading) quote from Peirce: "The word 'ought' has no meaning except relatively to an end. That ought to be done which is conducive to a certain end."[5] We should examine what is correct about this quote, and also what is potentially misleading about it. As we argued in Chap. 1, we can only be normatively engaged with things insofar as they engage our interests. As we noted (in borrowing Richard Kraut's example), if you were to tell us that smoking had this

property of *badness*, but had no ill effects on human health (did not cause cancer, or emphysema, or decreased lung capacity, etc.), then you would have offered no reason to refrain from smoking. Smoking's badness, insofar as it failed to connect with any human concern, could not constitute a *reason* not to smoke. For something to be reason-giving (positive or negative), it has to promote (or hinder) our interests. This is the sense in which Peirce is correct in saying that all "oughts" are relative to an end: reasons must always relate to an interest of ours. The essential role of interests in normativity was thus one thread of our discussion. The second thread is that insofar as we have interests *qua* embodied creatures, insofar as interests are essentially world-involving (but not reducible to states of the world, and not items added to our ontology), we further cement the connection between our account of normativity and our modest naturalism. If I say, "Smoking is bad because it ruins your health," I am saying that for a creature who is embodied as you are, who has a particular practical orientation (toward the good of health), you have a reason not to smoke. None of this is to reify interests or to claim that they are mental states, or biological states, or anything else. But insofar as I hold you to be committed to preserving your health (whether you acknowledge this commitment or not), I attribute to you a reason not to smoke.

However, (and this is the third thread to tie together), Peirce's quote is misleading in that it implies that all normativity is means-end—that all rationality can be reduced to a series of hypothetical imperatives. Recall our example from Chap. 2: If Bill says he really *wants* to lose weight (and you might think he really *ought* to lose weight), and then eats an entire gallon of ice cream in a single sitting, what can the Humean say? He has pursued a means (eating a gallon of ice cream) to an end (namely, immediate pleasure and gratification). He has only acted irrationally if you think he acted on the *wrong* end. But that is something you can only say if you have an account of constitutive rationality. Without this, an account of means-end rationality ceases to be an account of rationality, and merely becomes a descriptive statement that a person who in fact has a particular end will in fact pursue a means to that end. Thus, an account of normativity that tries to reduce all normativity to means-end reasoning is not an account of *normativity* at all.

This is where interests come in. Interests, on this account, must serve a number of roles and satisfy a number of constraints: (a) they give us reasons for action, in the Peircean sense above, which they can only do if (b) these interests are themselves normatively significant items, a lesson we have learned from Korsgaard and Hampton. Concerning (b), we noted a few pages back the sense in which there is a matrix of entanglement and interdependence between the practices and conditions in which take up reasons for action and our interests. Thus, while our interest in being healthy provides us with reason to engage in particular behaviors, our evolving standard of what behaviors are morally permissible may also alter our conception of this very interest (health) as the historical example of homosexuality and mental health demonstrates.[6] This shows that there is a complex, holistic interdependence between interests, reasons for action (which are dependent on a whole network of interests and practices), and the revisability of each element of our practice. To say that there is such interdependence between interests and practices is not to deny that the practices and conditions in which take up reasons for action arise from our interests, any more than embracing the theory-ladenness of observation is to deny that our body of empirical theory ultimately rests on observation. It is, however, to deny a foundational, unrevisable, and infallible role to whatever set of interests we might endorse at any given time. Thus, interests are ultimately thoroughly normative in nature (which satisfies Korsgaard's concern on a theory of action), while at the same time giving rise to the practices and conditions in which take up reasons for action and our interests (giving meaning to Peirce's dictum).

An interest that is shared by members of a particular epistemological community can not only coordinate action but also allow for the projection of coordinated action into the future. Consider, for example, physicists' interest in providing experimental support for the Standard Model, thereby enhancing our understanding of the physical universe. Again, we take it that having an interest in knowing that the Standard Model is true (or approximately true) does not entail that there is a property or object *knowledge* in the world; rather it is to endorse or encourage some courses of action with respect to the Standard Model. This interest helps coordinate and project the actions, particularly when it is made explicit. So, for example, a significant step in confirming the Standard

Model would be confirmation of the existence of the Higgs Boson. But conducting the experiments to confirm the existence of the Higgs Boson required the construction of the Large Hadron Collider (LHC), which took 10 years and the collaboration of 10,000 scientists to complete. A further four years of experiments have yielded strong evidence (in the form of observations "consistent" with a Higgs boson) confirming the Standard Model (CERN Press Office 2012). This example exhibits two particular features of interests that concern us here: first, this interest in confirming the Standard Model by confirming the existence of the Higgs Boson allows for a (large) group of people who share this interest to coordinate their behavior in a way that allows for the mutual satisfaction of this interest. Thus, we see that one pragmatic purpose of an interest lies in endorsing a set of behaviors, thereby allowing for a collective or group endorsement of a set of behaviors, which facilitates cooperative or group action. Second, the interest also involves *projecting* this set of actions into the future. It involves the projection that we will continue to have a similar practical engagement at some time in the future, and that the projects and questions whose investigation we endorse now will be ones we continue to endorse at some later point.

As we look to these features of coordination and projection, we see more deeply into the normative character of interests. When we examine an interest like "Confirming the Standard Model", we see that having this interest commits one to an *open-ended* set of endorsements and projection of endorsements, not simply patterns of behavior exhibited by agents to date, nor to particular psychological dispositions they may have at a moment. Thus, a physicist working in 1980 might have been committed by this interest to endorsing certain courses of action (e.g., work to complete the Superconducting Super Collider in Texas) which did not pan out (an instance of the uncertainty in our projection of endorsements), but is now committed by this interest to a different (but no doubt partially overlapping) set of endorsements. Nor can we stop there. Endorsement of a course of action like working to complete the LHC contextually implies the endorsement of further courses of action, such as holding a conference call with so-and-so, checking the liquid helium levels in magnets 53–82, ordering donuts for today's meeting, and so on. Many of these will involve further endorsements of further courses of

action. The idea, then, that an interest simply is a certain *object* or even fixed disposition of an agent is implausible in the extreme, as it would (like the property of being unfair, discussed in Chap. 3) be the most shapeless, gerrymandered, and disjunctive property ever conceived, and certainly not fit to do the work a reductionist has in mind for objects to do. Thus, any attempt to reduce the notion of an interest will find that it is just shared, goal-directed activity all the way down, with no non-normative bedrock in which to anchor an account.[7]

Where agents engage in goal-directed action, we may ascribe an interest in some outcome or the activity of pursuing it to them. Those agents may explicitly consider the interests in question, or they may not; whether or not they do will not fully determine the appropriateness of such ascriptions. Thus, there is a difference between an action *expressing* an interest and a person *acknowledging* an interest. Meeting one's friends for drinks might serve one's interest in enjoyable socialization, even if one does not specifically formulate this goal. On the other hand, one can act with a specific interest in mind, as, for example, when one goes running specifically with the goal of getting in shape. Indeed, there is an even broader distinction to be drawn between *having* an interest and having a personal or cultural *conception* of an interest; one can have an interest even if one is unable to formulate a conception of this interest (or even act in a way that embodies a commitment to this interest). We could imagine entire communities that were unexpectedly inarticulate about their interests, despite linguistic and conceptual abilities much like ours. Unexpected as that might be, they would still have them.

How can interests be ascribed in this fashion in the absence of explicit acknowledgement, perhaps even in the face of an agent's or community's denial that they have such interests? We suggest that this is possible because the possession of an interest is never solely a matter of personal or communal assent, or membership in a community that endorses some set of social practices, but rather a meshing of such practical resources with the conditions in which they are embodied in the world. Interests are not solely a matter of voluntary choice, but involve a complex interplay between one's embodiment conditions, the surrounding social practice, one's life plans, and so on. Choice plays an important role—Jones may have an interest in attending this weekend's Renaissance Faire, or in veri-

fying Smith's observation of a habitable exoplanet, or in improving her jump shot, Brown might not have any of these interests, due to having formulated a different life-plan. But to say that some, even many, of our interests are subject to choice is not to say that most or all of them are—various moral and epistemic interests are not optional, and the same goes for various prudential and other sorts of interests.

If we talk about interests at this level of specificity, though, it becomes difficult to know which interests to ascribe to a person. (And doubly so if acknowledgment is not a necessary condition for having an interest.) Jones might have an interest in learning more about the US Civil Rights movement, but she might equally well have had a satisfactory and valuable experience learning about China's Great Leap Forward. Should we say she also has an interest in the latter topic? But then do not a person's interests become indefinitely many, uncountable? To address this question, we must articulate how it is possible to theorize systematically about interests and make every ascription subject to scrutiny.

5.3.2   General Interests and Normative Theorizing

Often, accounts of practical reasoning start with big-picture lists of interests that are ostensibly basic to or implicit in all other practical concerns. Compare, for example, the list of basic goods in natural law theory, virtue ethics, or Martha Nussbaum's (2000) capabilities. Nussbaum's capabilities include things like life, bodily health, affiliation, and control over one's environment, among others. Concern for, say, access to a particular medication would be justified in terms of its mitigating some symptom, which, in turn, would be justified by facilitating the exercise or acquisition of a capability (e.g., sumatriptan mitigates the symptoms of migraine headaches, which, in turn, allows us to function in various ways.) Other authors have called these *basic* interests, but for reasons that will become clear, we will instead call these our *general* interests. By contrast, we can also distinguish a notion of *local* interests here, as well. For the moment, we can say that general interests will be those that exhibit the greatest generality, and are not subsidiary to further sets of interests. So at first pass, they will at least include items often identified as intrinsic goods like

health or knowledge. Local interests will exhibit much greater specificity and locality to the projects in which particular agents are engaged; they will be the more familiar foci of our concrete, day-to-day practices. Thus, one might have a local interest in harvesting the crops before the rains come. This local interest will make rational (and will also motivate, to the extent that this interest is lived by the person) a number of subsidiary actions, such as making sure the combine is serviced and fueled, arranging contracts for the purchase of the crops, perhaps hiring extra help in anticipation of the additional labor required, setting the alarm early each day in anticipation of a long day of work, and so on. While we will argue that talk about general interests does have a pragmatic role to play, we deny that our local interests are sustained or justified solely by "general interests" in these ways. Rather, people may just as legitimately act to promote interests that are more concrete, specific, and directly world-engaging. General interests like health may play a role in our account, but such an interest need play no role in justifying or explaining the farmer's actions, and need not exist as a psychological state or other ascription to him in every case.

Local interests serve a dual role in practical reasoning. First, they serve to justify a practice, or set of actions. Harvesting the crops before the rain comes is ostensibly a legitimate interest, and we can endorse an interest like this one, but we can also criticize an interest as trivial, immoral, a poor use of resources, or on many other conceivable grounds. An interest in getting the crops harvested before the rains can justify a wide range of actions; an interest in collecting the whole range of Beanie Babies is possibly too trivial to have justificatory force; an interest in reviving Jim Crow laws is terrible. Second, an interest can be taken up by a person, can become lived or explicit in reflection, such that the person organizes her actions in pursuit of this interest. Thus, a person who takes on board an interest like harvesting the crops before the rains come may, as a consequence, perform all of the actions listed in the previous paragraph. The normative and motivational features align in other ways, too. We can criticize someone for *failing* to take on board an interest we ascribe to them: we can argue that Reynolds has a deep interest in getting her crops in before the rain comes, but is failing to act on this interest, and so is being foolhardy or irrational.

As our examples here would suggest, we take it that we may attend to interests at different levels of generality for different purposes, and that different formulations of those interests may compel different agents in different ways, or not at all. Health seems like as widely held an interest as we could attend to, for instance, while proving a particular theorem might be the province of very few or even a single mathematician. Note that the impacts of those pursuits can vary as well: each of us has an interest in his or her *own* health, but our lonely mathematician will prove a result *for everyone* if she is successful. The theoretical temptation in prioritizing general interests is to offer some set of interests that had the greatest such generality and urgency and assign them a privileged explanatory or justificatory status in our reasoning about norms. We could then think of other interests as deriving their importance from those general interests somehow. We then pursue the local ones not for their own sake, but for the contribution they make to the pursuit of the general ones. We think there is a theoretical value to certain ways of taking up this approach, but we would not endorse versions of it that make certain interests transcendental necessities or features of a natural order to be discovered, as orthodox versions of virtue ethics and natural law theory do for some normative matters.[8]

We think a more plausible understanding of what we called general interests is that they play an expressive role that emerges upon certain types of reflection on local interests, but which does not assign them a prior or fundamental status from which others derive their importance for us. The sense in which they are fundamental or prior cannot be an historical one; we are not descendants of communities that cared about general interests like health and knowledge, but not local interests like more effective farming and anticipating threats in their surroundings. We should actually find just the reverse. Any actual instantiation is going to be one in which the conditions in which agents are embodied have already circumscribed possibilities and pushed some things to the fore. In actual, everyday coping—even for agents and communities distantly removed from explicit articulation—a general interest like health will have already manifested itself in interests in food and shelter, and even those will be manifested in more specific interests given the conditions in which they cope. That is, their interests in *getting food* will show up

as interests in getting whatever happens to be an option locally (perhaps by fishing for those near water, hunting for those on the plains, etc.) and interests in shelter will reflect local conditions in things like materials and climate (ventilation in the tropics, insulation in the arctic, etc.).

For the moment, the point to note is that whether we think of those interests as explicit in a philosophical account or implicit in everyday coping, there are no reasons to assign general interests a conceptual or historical priority over local interests. They are not general in the sense that they are independent of local interests, nor are they determinants of the sort of urgency that local interests have for us. But that is not to say they are worthless distinctions, nor that we should try to eliminate them altogether. We can suggest a set of moves for the introduction of general interests and invoking them as reasons that parallels the story Sellars gave about sense data and thoughts. Suppose Jones, the mythical genius who first proposes sense data to explain the behavior of his fellows, has a brilliant mythical sister whom we will call Janet. Growing up with Jones, Janet finds herself immersed in explanations of other speakers' behavior in reporting and knowing the world in terms of internal episodes that arise in light of "the impingement of physical objects and processes on various parts of the body" (Sellars, 1956/1997, p. 109/§60). She wonders about commonalities in the behavior of her fellow agents—commonalities that encompass the epistemological, observational, and linguistic but extend to include *all* varieties of behavior and moral, prudential and other forms of giving and asking for reasons.

Janet notes that she and others do, try to do, and sometimes report feeling satisfied by their efforts to build a lean-to, gather ripe tomatoes, explain why their tomatoes are now ripe, avoid death in the jaws of sabertoothed tigers, talk others into dancing with them, and vast arrays of other activities. When she asks why people do these things, they sometimes say very immediate things—"I wanted to build a lean-to" or "It was going to eat me"—and sometimes they suggest a bit more—"We have to eat, we like tomatoes, and this is when they taste best." Crucially, she might also note that variations on these practices and actions are taken up by the same people in different circumstances. (They bundle up in the cold, but wear less in the summer; they avoid eating some things for their extremely pungent taste, but eat plenty of ginger that seems to help them

digest.) In a moment of ingenious intrasubjective comparison, she might propose that there is a common way to articulate to how these behaviors and practices are directed toward their goals that supersedes the ways in which they are expressed and accounts for the variations according to circumstance. Her friends in the tropics build lean-tos, harvest mangos, and avoid standing water. Her friends in the arctic build igloos, spear fish, and cover all the skin they can when the wind picks up. But Janet notes that what she has introduced is distinct from Jones's introduction of inner episodes as *posits*. What draws her observations together is the very heterogeneity of the causes and conditions involved, and the sense that their open-ended character would be poorly explained by any causal account. What all of these conversations do is *justify* a way of going forward, and the unity among these many diverse cases is simply not a further kind of prediction. What we do in unifying these cases under a more theoretical notion of an interest is to recognize and affirm the ways in which others might (and do) direct their activity, not simply anticipate them. To conceive of her interests and theirs is to seek ways of *sharing* or *disputing* the ways they go forward, not just *observing* and *anticipating* them. A different category is needed for such a manner of engaging them, and thus she allows us to move from simply describing their motivations to evaluating them by the addition of this category. She can assert that for all their local differences, the motivations and satisfactions behind all of these practices belong to the same theoretical category—an *interest* such as "health."

Like any account that appeals to more general features, this has the virtue of making both her arctic friends and her tropical friends' practices seem less random, and it has a degree of counterfactual robustness to it. If we were to transplant the arctic folks in the tropics and the tropical folks in the arctic, they would probably seek locally appropriate sorts of shelter. But the value of this distinction does not lie in this predictive dimension. It is rather that they will have *reason* to (and we can expect them to) take up the same or relevantly similar practices because those practices *serve some interest* that was common to both communities all along. Janet notes that she and her friends have had a sort of interest all along (getting the lean-to to stand up, having tomatoes to eat, etc.) but that her theoretical move here is to suggest another layer of the account, one on which what appeared to be diverse elements of our social and practical spheres have

a commonality that is indispensable for certain purposes. Interests will thus be of service in *justifying* all of these practices and adjustments, not just describing the regularities in them. And thus, we have a role for the notion of a general interest, though we have come to them *from* our local interests, rather than treating the general interests as axioms or generators of more specific interests.

Thus, a general interest plays a role in justifying *myself* (or *ourselves*) to someone and directing the courses of action we take going forward. That is, it serves the discursive activity of excusing or entitling oneself to some disputed commitment in a manner that will likely appeal to a general interest to which others are already committed. If you spot me reading a book and ask me why I am doing so, I will answer with some local interest; if someone is persistent enough, we will come to appeals to general interests. So, reading this book improves my understanding of an academic debate, which allows me to participate in those scholarly practices of exchange, all of which serve into our general interests in knowledge and control/autonomy. At that point, our spade is turned, as Wittgenstein would say. We may tell a story about general interests that integrates them with one another in some way, but we will not appeal to a further layer of rational motivators for local interests. Affiliation may give us more opportunities for pleasure, and knowledge may give us better prospects for control and autonomy, but there is no further, more general layer of interests that these serve. And if some local interest really does serve a general interest or interests that we seek better than our alternatives, then this serves as a reason for an action or practice.

General interests have a certain pragmatic thinness though, even greater than "good," "right," and similar terms in ethics. If you ask me why I shovel snow from the driveway of the little old lady at the end of my street, and I say, "It was right to do so," I sound a bit lofty but not entirely out of step. If you ask me why I keep looking at my watch nervously, and I say, "It serves my general interest in knowledge about the world," I am either being obtuse or serving up a punch line. Despite their airs of pragmatic centrality, general interests are really most at home in a kind of philosophical discourse that does not regularly take more colloquial forms. In fact, people with very sophisticated, well-articulated

sets of local interests may be hopeless in thinking about general ones, just as many morally good people do not make very good moral theorists. This is not a failure of their efforts in those specialties any more than not offering a theory of moral motivation marks a failure of most moral agents. Talk about general interests is itself a kind of discursive practice that serves a set of interests that are not in play in most other situations.

Given all of this, we are in a position to see the expressive role of general interests more clearly. Local interests have normative force on their own; we need not appeal to general interests in justifying the actions and practices that are taken up in pursuit of local interests. However, general interests can be appealed to in order to unite disparate practices, and show how the local interests pursued under these practices are licensed as species of the same genus. This is more striking when we engage in cross-community discourse. There, by having discourse about general interests, there is a way to ask whether very different sets of social practices that do not map neatly onto one another might still be permissible. For instance, if we come across communities that practice avuncularism in their filiation, like Taino or Chamorro communities (at least in the past), then their practices appear to license males abandoning their biological children; appeal to general interests might allow us to say that despite these differences, health and affiliation are still pursued for vulnerable members by different familial arrangements. So the direct purpose to which general interests are being put is the evaluation of sets of local interests (whether to adopt this or that research project, how to get along with communities that differ enormously from our own, etc.) and thus resolving disputes that center on tradeoffs between concrete, embodied concerns. So, discourse about general interests can inform our decisions about how to get along with the avuncular Chamorros so that further concrete interests we have in, say, interacting with other communities in commerce and other overlapping practices can proceed. Or we might weigh the priority that different scientific projects adopted by our community have when we distribute resources, so that we may balance the concrete benefits our own community gets from emphasizing some of its projects over others. Rarely, if ever, are general interests in themselves concerns for anyone outside of philosophy, and without a bearing on local interests, they could hardly be worth any pursuit at all.

With that general proviso, there is still a type of discourse about general interests that does seem to focus on general interests themselves. A single local interest may serve more than one general interest (e.g., pursuing an interest in a team sport will tend to bring affiliation and likely improve our health). There is a place for discussion about how general interests like these will tend to rise and fall together, and how such a list should take shape in the first place. How does pleasure go along with knowledge? Or affiliation with health? As we have seen, the thinness of general interests precludes a purely abstract discourse even in this case. We can talk about how pleasure, knowledge, health, affiliation, and the rest all comingle and complement one another, but any claim we are entitled to make about those connections will have to be a nuanced one based in more specific (if still highly general) sets of local interests. The pursuit of knowledge really does facilitate pleasure in many cases. At the same time, other practices involving the hunt for knowledge will be net losses for most who undertake them, or communities as a whole. The thinness of any claim like "knowledge brings us pleasure" or "health and affiliation help us achieve autonomy" will only be remedied by thinking about the more concrete interests that weave together below them and attend to the details of how and why they weave together in just those ways. If we assert that some set of practices or the pursuit of some local interest serves a general one, that commits us to saying that in the long run, we will find most similar local interests—the ones we lump under knowledge, for instance—well-served by some practice or action. When we make an appeal to general interests in cross-community discourse (say, by appealing to the benefits of education for women in the developing world and its salutary effects on other interests, like the health of their families), then we are asserting such systematic, long-term service to many interrelated local interests on their parts that parallel something in our own.

Having said something about the role that general interests are to play and some things about the force they have as reasons for consideration of other interests, a penultimate conclusion suggests itself here: we do not really act or sustain practices in the service of general interests. We certainly act and sustain practices in the service of interests, but not the general ones. By the time flesh-and-blood agents are actually acting,

individually or collectively, and directing themselves at goals, those goals are what we have been calling local interests. The difficulty we had in saying just how one could draw the character of a local interest from a general one was not accidental, it was indicative of a more fundamental difference between the two categories. Whether or not a local interest is made explicit or some conception of it becomes part of an intention, that local interest can be a goal and achieving it can be satisfying in a way that is not similarly available for knowledge, health, or autonomy in general. This does not make local interests real and general ones unreal: appeal to each of sort plays an expressive role, and we want both. But commitments to the two categories are strikingly different, and any sense of commitment to the general interests will have to be cashed out in terms of practical commitments to the local ones.

We will close this chapter with two cautionary notes. We have emphasized repeatedly the world-involving nature of interests. One might think that this at least throws us back on a supervenience account of interests, and their relation to the non-normative. As we noted at the end of Chap. 3, there is a sense in which it would be strange to deny global supervenience, at least. But we argued there that supervenience, rather than being a metaphysical thesis, simply follows from prior views on rationality: parity of reasons demands that we should judge or act the same in cases where our reasons are the same. But in worlds where all the non-normative facts are identical, we would have the same empirical facts to consider in formulating our normative judgments, and so a difference in judgments would require different empirical facts or else it would become arbitrary and irrational. But this is a normative commitment guiding us in how to make our judgments, not a discovery of some natural facts that determine the judgment for us. Normativity nudges us toward accordance with global supervenience, rather than the supervenience helping to explain normativity. But the key point we wish to emphasize is that these empirical facts are inert in the absence of interests. That the rains are coming does not bear its importance on its sleeve, but if you have an interest in wine production (and know that they rains will make the wine watery, if the grapes are not harvested now), or if you are an

avid sailor, or a baseball player, and so on, then those facts will matter to you. In the presence of a set of interrelated practices and interests, this fact has significance, and can give us reason to perform actions (harvest the grapes, cancel the game, etc.). These empirical facts have the grip on us that they do only in the light of interests, and thus any sense that the non-normative facts could *determine* normative matters is simply misplaced. In a fundamental sense, our objection to traditional naturalist accounts of the natural-normative relation (such as supervenience accounts) is that they miss the point of evaluative discourse in the first place. To evaluate is not in the first place to describe, and so locating the factual element of an evaluative statement is heading down the wrong track in the first place. But worse, this factual element will never generate for you that which is the central element of the evaluative statement in the first place: namely, its *evaluative* or *action-guiding* character. In short, while we do not deny global supervenience, we embrace it for different reasons than do most naturalists, and we do not ask it to do the work that most naturalists do.

However (and this is our second cautionary note), if we distance ourselves from traditional naturalist accounts, we are also not offering a transcendental account of our practical interests. As we said in Chaps. 1 and 2, we take the project of elaborating normativity to be one undertaken fallibly and without the incorrigible sources of philosophical intuition that have undergirded many projects in the past. Instead, it must begin with the practices and perspectives we already inhabit, and from such a position, any interest we may actually pursue is fair game for challenge and revision if the terrain should change. That stance does not preclude some interests, broadly construed, having a kind of stickiness or permanence that makes them of enduring theoretical interest. An interest like health has a pedigree as long as human social practices, and we can scarcely imagine what a community that did not pursue it in some fashion would be like. Other interests will be far more optional or disposable. No one *has* to worry about collecting 8-track tapes, or getting wider lapels on their zoot suit, or enhancing their social media profile, and given the transience of most social practices, there will almost certainly come a point where no one even remembers such interests, much

less pursues them. We should expect various levels of stickiness to our interests between these, and there will be theoretical payoffs to elaborating why these differences emerge.

Notes

1. In Chap. 8, we will explain the causal impotence of the normative in terms of the lack of what we will call the *vertical contribution* of normative discourse into factual discourse. There, we define vertical contribution as a type of integration, whereby we account for some features of a higher-level discourse in terms of a lower-level one, as when (for example) molecular properties help us explain features of classical genetics. Vertical contribution will not require a full reduction or elimination of the higher-level discourse. We deny that normative discourse is vertically integrated into a lower-level nonnormative discourse in these manners, or that normative things are somehow composed of non-normative things. However, there will be other interesting contributions from non-normative discourse regions to normative ones, which we describe in Chap. 9.
2. From Heidegger onward, others have argued that understanding humans as agents, bound by norms, essentially involves understanding us as engaged with the world. Lance and Kukla argue in their recent book that "our social placement as agents within the discursive community and our responsiveness to the normative claims of the empirical world are interdependent" (2009, 211).
3. The most obvious figure endorsing such moves would be Rorty, but the list could go on.
4. Denial that we have some interest is not prima facie incorrect, but we presume there will be a strong default to accepted formulations of interests.
5. This is found in his *Collected Papers* 5.594, 1903.
6. We discuss this example above, but also at greater length in Chap. 6.
7. Compare Brandom's (1994, 45) discussion of a sanction as being internal to a system of norms.
8. Our sense that this needs to be made explicit is informed by conversations with Matthew Burstein and Mark Murphy.

References

Brandom, Robert. 1994. *Making It Explicit*. Cambridge, MA: Harvard University Press.

CERN Press Office. 2012. CERN experiments observe particle consistent with long-sought Higgs boson. *CERN*. July 4. http://press.web.cern.ch/press-releases/2012/07/cern-experiments-observe-particle-consistent-long-sought-higgs-boson. Accessed 23 July 2013.

Chemero, Anthony. 2009. *Radical Embodied Cognitive Science*. Cambridge, MA: The MIT Press.

Dewey, John. 1925/1958. *Experience and Nature.* New York: Dover Publications.

Honneth, A. 2008. *Reification: A New Look at an Old Idea*. Oxford: Oxford University Press.

James, William. 1904. A World of Pure Experience I. *The Journal of Philosophy, Psychology, and Scientific Methods* 1(20): 533–543.

Kukla, R., and M. Lance. 2014. Intersubjectivity and Receptive Experience. *The Southern Journal of Philosophy* 52(1): 22–42.

McDowell, John. 1981. Non-Cognitivism and Rule-Following. In *Wittgenstein: To Follow a Rule,* eds. S. Holtzman, and Christopher M. Leich, 141–162. New York, NY: Routledge.

Nussbaum, Martha. 2000. *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge University Press.

Rupert, Robert D. 2004. Challenges to the Hypothesis of Extended Cognition. *The Journal of Philosophy* 101(8): 389–428.

Sellars, Wilfrid. 1956/1997. *Empiricism and the Philosophy of Mind.* Cambridge, MA: Harvard University Press.

Wagman, Jeffrey, and Anthony Chemero. 2014. The End of the Debate Over Extended Cognition. In *Neuroscience, Neurophilosophy, and Pragmatism: Brains at Work with the World*, ed. T. Solymosi and J.R. Shook. Basingstoke: Palgrave Macmillan.

6

Action-Guiding Content

In this chapter, we will offer an expressivist account of declarative sentences that appear in normative discourse. "Normative discourse," in the present sense may stretch to include large numbers of linguistic forms, and we will not attempt to catalog and explain them all. Instead, we will offer expressivist characterizations of those employed in overt normative *claims* and suggest some ways in which they may be related to others. The common core of all the different parts of this account will be a concern for our practical interests and projections of how the adoption of various commitments and entitlements would serve those interests going forward. This pragmatist move will allow us to preserve a very robust account of normative discourse without the need to posit a new range of objects or phenomena in the world that our discourse would then represent. There will be no placement problem for normative discourse because there will be nothing to place. Our account will involve a proposal for a considerable extension of the resources for analyzing a language. In addition to many elements familiar from other accounts, we articulate different types of action-guiding character that pieces of normative discourse can have, which we will call *action-guiding modes* here. In some ways, these will remind speakers of accounts of speech acts, but

the crucial difference is that we are suggesting these modes are built into the very content of these sentences and expressions, not added afterward by our performance with them. We are suggesting that it is not normative discourse that is somehow impoverished, but rather the set of theoretical tools by which we have attempted to understand it.

This chapter takes up a project that has been suggested by the work of Terry Horgan and Mark Timmons. Readers of their work will note that their objections to reductionism in moral semantics (1991, 1992, 2000a) led them to reject that program in ways that our work in Chap. 3 echoes. Over the last 25 years, they have argued extensively that there were no plausible factualist accounts to be had in metaethics, but that a move toward non-factualism and expressivism in moral judgments did not entail non-cognitivism. Moral judgments could express genuine assertions of beliefs once we recognize that there are types of content beyond the descriptive and the merely attitudinal that serve action-guiding roles (2000b, 2006a, b). We can now see much of our work as preparing the way for the sort of cognitive but non-factual account Horgan and Timmons suggested.

We hesitate to call such an account "expressivist." The historical associations between that term and various forms of non-cognitivism are very strong, even if they are not necessary ones. And for that matter, non-normative assertions express attitudes and commitments as well, so the very terminology that suggests normative discourse is a separate sort of thing—a runt, really—seems to stack the deck from the start. However, our expressivist account will differ from many predecessors that bear that label in that at least some overtly normative sentences will be either true or false, and their truth or falsity will not be a matter of cataloging what the predominant or canonical attitudes of our community happen to be. So we call our account an expressivist one with some hesitation and caution to the reader. Our view will be fundamentally different from those of emotivists like Ayer (1952, Chap. 6) and Stevenson (1937, 1944), or even more sophisticated non-cognitivist expressivists like Gibbard (1990, 2003). Such accounts take normative discourse to be expressive of endorsement of our norms, and a proper accounting of attitudes we happen to have (including how some of those attitudes might compel us to revise others) *exhausts* this expressive role. Normative discourse bottoms

out in our actual attitudes on such accounts and to ask whether those attitudes are true or correct is to commit a kind of category mistake. We contend that this is not the case. Normative discourse (including its declaratives) is externally constrained in ways that make it as truth-apt as familiar forms of non-normative discourse, even though it does not represent anything in the ways that metaethicists have traditionally expected descriptive language to. As readers will expect from earlier chapters, this position grows in part from our view that those representationalist accounts are misguided from the start, and abandoning those expectations for normative discourse introduces no special problems for this part of our language.

Sect. 6.1 outlines the most general features of normative discourse and how making such claims can be externally constrained without reverting to some form of representationalism or lapsing into non-cognitivism. Sections 6.2 and 6.3 describe how these general features cash out more specifically in language about goods and language invoking rightness or correctness.

6.1   Action-Guiding Content

6.1.1   Normative Declarative Sentences and the Embedding Problem

Given that we take our position to be a cognitivist account of normative discourse, we now need to say how such an account permits some assertions and some genuinely true or false sentences. (We emphasize "some" here simply because normative discourse may also include questions, commands, and many other items, as well.) Critics can rightly point out that everything we have said thus far is compatible with non-cognitivist forms of expressivism. If we took the content of normative discourse to be expressions of subjective attitudes, or allegiance to norms, or prescriptives, they would all still be interpretable in the interest-pursuing terms we have laid out. We do think that there are expressions of subjective attitudes and norm-allegiance, but we do not think that all forms of

normative discourse belong to those categories, as some have an action-guiding character that is not a report of our attitudes and allegiances. Many will accept that there are some parts of the language that guide action, but they may hesitate to count them among the truth-apt sentences. Imperatives are helpful in thinking about this point. It is uncontroversial that there is a kind of action-guiding character to something like a command, expressed with an imperative; but imperative sentences do not even purport to truth. We can even say that imperatives have *some* sorts of external constraints. For instance, "Drop and give me 20!" makes sense as an imperative only for suitably limbed creatures in a world with gravity.

A familiar lesson to be drawn from Geach's (1960) and Searle's (1962) versions of the "embedding problem" is that in their simplest forms, overtly normative sentences occupy the same grammatical forms that descriptive sentences that we assert do, for example:

(1) Murder is wrong.

If we analyze this in terms of expressions of subjective attitude, allegiance, or as prescriptives, we avoid Mackie's "queerness" problem. But Geach and Searle noted that if we treated normative sentences like (1) solely as expressions of our aversive attitudes, then we cannot account for the role of those sentences in conditionals like "If murder is wrong, then Jeffrey Dahmer committed numerous wrong actions." Like a good conditional, the antecedent here (along with some background assumptions) implies the consequent. But bare expressions of attitudes do not have logical and semantic consequences of this sort; "Yaaaaaaay Orioles!" or "Boooooooooo Yankees!" are not even in the game of being truth-apt. They might arguably suggest reports of our attitudes (for example, "I hate the Yankees") that can be true or false once indexed to a speaker, but that link is fairly trivial and will not shore up the roles that normative sentences appear to play.

This point is sometimes generalized to suggest that non-descriptive discourse cannot fit into logical operators at all, but this is at least somewhat premature. Consider a few examples with imperatives. (What we take to be embedded imperatives for the moment are italicized for emphasis.)

(2) If you can't take the heat, then *get out of the kitchen*!
(3) *Put up* or *shut up*!
(4) *Come over here* and *give me a hug*!

In each of these cases, we appear to successfully embed one or more commands in a place where truth-apt sentences normally go. If readers think that the embedded components in (4) are invitations rather than commands, our earlier example "*Drop* and *give me 20!*" will do. This may be illusory, however, as negation does not work so well. It does not seem that

(5) Don't drink the water.

is equivalent to

(5′) It's not the case that *drink the water*!

In fact (5′) does not seem to make sense at all. Strange things also happen when we put imperatives in the antecedent position, or if we take the contrapositive of the conditional that seemed passable above:

(2′) If *don't get out of the kitchen*, then it's not the case that you can't take the heat.
(2″) If it's not the case that *get out of the kitchen*, then it's not the case that you can't take the heat!

Sentences (2′) and (2″), like sentence (5′) make no sense at all, even if we let the negation into the imperative in the antecedent. Biconditionals fare no better:

(6) *Get out of the kitchen* if and only if you can't take the heat.
(6′) You can't take the heat if and only if *get out of the kitchen*!

Perhaps (6) could be intelligible if we read it generously as implying a command issued to someone whose welfare was in danger but whose presence was urgently needed. ("We really need your chopping skills

on this salad, but we don't want you dying of heat stroke..") But this is probably taking charity a bit too far if we are thinking about the role of imperatives more generally, given that (6′) should be equivalent but not even the charitable reading sounds right there.

Why do (2)–(4) work and things go off the rails thereafter? Some parallels can be drawn here with Austin's (1970) "biscuit conditionals," for example, "There are biscuits in the cupboard, if you want some." Surely, they are there or not regardless of my desire for them, so the usual notion of implication does not appear to be at work. As DeRose and Grandy (1999) put it, most analyses of this phenomenon have started with the intuition that "with the biscuit conditionals, it's the conversational relevance of the consequent that seems to be contingent on the antecedent" (p. 406).[1] An analogous point can be made about other examples. The conversational relevance of conjunction and disjunction might be in elaborating joint requirements and inclusive or exclusive possibilities, respectively. So (3) and (4) are simply commands that are complex—satisfied only by multiple conditions, or by a number of conditions within a range. In this case, the structural appearance of conjunction or disjunction may be superficial, perhaps even borrowed from declaratives in order to make those more complex conditions explicit.[2] That role might endure even when truth-aptness is not at stake, as when we make demands or offer options to others. We might thus say that the satisfaction conditions for a complex command like these would invoke a conjunction or disjunction of genuine assertions; if I tell you to do three things (and all of them), and we cannot subsequently assert that all three are the case, then you are not done yet. (*Mutatis mutandis* for exclusive and inclusive disjunctions.) But none of this turns *the commands themselves* into truth-apt sentences.

Conditional commands mimic the sort of list-like structure of conjunctions and disjunctions, telling the audience that the result of checking the antecedent may give rise to a commitment to the consequent. The superficial conditional form in (2) looks even less like a genuine conditional when we notice that it cannot be manipulated in customary ways, so the sort of "limited borrowing for expressive purposes" interpretation appears appropriate here, too. One possibility that should be considered here though is that in some contexts, (2) might

be interpreted not as involving a genuine command, but rather a truncated normative sentence: "If you can't take the heat, then *you should* get out of the kitchen," where "you" in those two instances can be read as an indefinite pronoun akin to "anyone."[3] Truncated sentences of this form seem to occur elsewhere in everyday speech. If I am sitting with a friend, hearing her stories of awful things that happen at work, I might say, "Quit your job. That place is killing you." But this should not be read as an actual command; it is prudential advice more perspicuously stated as, "*You should* quit your job." We could imagine moral and epistemic advice of analogous sorts, and thus we find ourselves in the realm of hypothetical imperatives.

We have only considered some imperatives here, but this glimpse of the wider topography is enough to illustrate the difference to be drawn with declarative sentences that are overtly normative. While imperatives and some other grammatical categories may serve this role of guiding the actions we take and the goals at which we might direct them, normative declarative sentences do this and also function like other descriptive one in taking their place among pieces of logical vocabulary and purporting to say something truth-apt. There is no comparable shortcoming or feature to be "explained away" about the normative sentences embedded in these examples:

(1) Murder is wrong.

(7) It is not the case that you must give 20 % of your income to The Church of Scientology.

(8) Either you should commit to raising your children well, or you should not have children at all.

(9) If there is a possibility of bias, then you should make relevant parts of the study double-blind.

(10) If murder is wrong, then Jeffrey Dahmer committed numerous wrong actions.

(11) Murder is wrong if and only if killing in cold blood is wrong.

And as Geach and Searle noted, it will not do to construe these as merely covert expressions of individual attitudes, or even reports of

communal assent. Note the differences in the inferential roles in these examples:

(12) X is right.
(13) Our community affirms that X is right.
(14) S is committed to "X is right."

Whether or not (12) is true, it is certainly not equivalent to, nor does it imply (13) or (14). Nor does (13) or (14) imply (12). (The relativist might argue that (13) does imply (12), but if so, this is a further theoretical commitment, not an account of the sentences' meanings.) To call something *wrong*, or *incorrect*, or say that *we should do otherwise* is to ascribe to others and oneself a commitment to avoid it, but also to at least purport to say something that cannot simply be identified with one's attitudes or the stance that one's community happens to take. Some of these will strike us as odd when manipulated as we did earlier, for example:

(9′) If it's not the case that you should make the relevant parts of the study double-blind, then it's not the case that there is a possibility of bias.

But the oddness of this contrapositive is that it suggests an implication based on a condition that we cannot generally confirm on its own (i.e., we cannot decide what methods are appropriate prior to knowing what sorts of biases are in play). So we would not generally accept such a conditional at face value, but that hurdle is epistemic, not semantic or logical.

So, our work will suggest that the prevailing forms of non-cognitivist expressivism mischaracterize many elements of normative discourse. This invites challenges from error theorists who will agree that some parts of normative discourse behave like declarative sentences, but that these parts are mistaken when we assert them. So what do we assert with sentences such as (1) and (7)–(11)? What sort of propositions could be in play here if, as we have argued, there are no entities or substantive properties in play? We propose that claims in normative discourse assert *matters of our commitments and entitlements themselves*. They are assertions *that*

ranges of action and other conditions are required, privileged, permitted, discouraged, and prohibited and all the other familiar modes of direction that speakers already engage in with considerable facility and fluency. We have been developing an account of discourse that emphasizes usage in its broad strokes and inferential roles in its details, on which the content of expressions (both sentential and, in the long run, subsentential) are understood fundamentally in terms of the conditions and consequences of their application as they are integrated into our social practices. Descriptive, non-normative claims typically prompt us to comport ourselves in various ways to parts of the environment in which we are embedded. Normative claims assert different positions on *how* to manage our commitments and entitlements, but normative declarative sentences will not "bottom out" in some non-normative substrate such as a substantive natural property. This should not surprise us. Entities and properties, no matter how arcane, are not intrinsically action-guiding in the way that normative claims straightforwardly are, and the temptation to reify the elements of normative discourse (or fret that no entities and properties are available) is to misconstrue the task at hand.

We emphasized in Chap. 3 that while we agreed that we should not posit special entities and substantive properties to account for normativity, we should not abandon the position that normative claims are truth-apt. What we are proposing is that the assertible content of normative declaratives is of a different sort—one in which the search for or addition of additional items is of no help to us, but which still have objective purport. So why not embrace fictionalism here? Because the general features of their integration into larger networks of inferentially related contents and the ways in which we answer to evidence are fundamentally the same for normative claims and non-normative ones. The need to hedge our bets, as the fictionalist would suggest, depends on accepting both this point about the uniformity of truth-aptness and a monolithic account of content that uniformly implies substantive ontological commitment for each of its expressions. As we saw in Chap. 3, such attempts to account for the content of normative claims always run up against substantial problems, so when faced to choose between our options, we are suggesting that holding fast to the monolithic picture of content is the mistake. To make this point compelling, we must articulate how such content

could be assertible in a manner that does not lead us to substantive truth conditions (both because this would lead us back to substantive normative properties and violate the deflationism we have endorsed), but which does involve an evaluation of those assertions in light of something other than social assent to their assertibility. What we should therefore offer are:

(A) ways of thinking about how non-normative discourse and the world can serve as reasons to inform judgment or revise usage in systematic ways;

(B) an articulation of the different modes in which action can be guided by parts of normative discourse, both on a claim-by-claim basis and in a general systematic way, to give us general features of normative discourse like good-talk and right/correct-talk.

The next two subsections take up each of these tasks in turn.

6.1.2   Evaluating Normative Claims

So we take it that normative discourse includes sentences with which we make genuine assertions, and not just expressions of attitudes, allegiances, prescriptives, and so on. They will thus be items that make purport toward truth, though we must now elaborate how they could do so without reverting to descriptive or representational accounts of their content. The deflationary accounts of truth we endorsed in Chap. 3 may permit us to sidestep certain questions about the metaphysics of truth, but that will be for naught if the assertion of the action-guiding content in normative sentences requires us to posit entities and substantive properties again. But something about the non-normative world needs to make a difference to normative discourse even if we are to have this more robust expressivist account. Threading this needle obliges us to say how judgments over action-guidance may be *informed* by descriptive content and conditions in the world without being *representations* of those items. Non-normative conditions should be at least parts of the sets of reasons that justify our making normative judgments and assertions, even though the normative judgments and assertions are not *about* the non-normative

conditions in the same way that a description is *about* the conditions or object described.

This will be possible for us by including the consideration of interests as described in Chap. 5, which will serve as a lens through which non-normative conditions can have a relevance that informs our evaluation of normative judgments and assertions. In our most reflective, sophisticated moments, we may make our interests explicit and take them as reasons to endorse or revise our courses of action and practices. The self-correction that lies at heart of rationality presupposes a self-recognition of ourselves as engaged in the pursuit of those interests. An action or outcome matters to an agent in characteristically rational ways only once they can take up stances toward their own conduct and practices concerning how they address the practical purposes they serve (i.e., the interests we have in taking those actions and adopting those practices). The central features distinguishing the content of thought and discourse with normative content will thus be a recognition of some interest and attention to an action, outcome, or practice's service of that interest.

First, let us say a bit about how what we assert with declaratives can be understood in terms of practical potential. The pragmatist tradition has emphasized unpacking reports of what appear to be the facts of a static universe instead in terms of potential differences. To describe an object as solid, for instance, is to say that it would act in certain ways if we engaged it in certain ways. Emphasizing difference in this manner may also lead us to the familiar objection that there are some descriptions that involve parts of our world with which we can never bodily interact (e.g., Peirce's (1878/1992) diamond buried at the bottom of the sea, or a physical event outside of our light cone). So we should take care not to make this point about practical differences in verificationist terms, even if many of our examples suggest features of everyday perceptual experience. For example, we could look at our commitments and entitlements as speakers and the differences it would make to assert that something was solid, rather than asserting it was liquid, gaseous, or any other option. (That we can use it to resist certain forces, that we anticipate it behaving in certain ways, that we are entitled to act and use something because we are entitled to expect such behavior, etc.). And these differences might serve different interests in different ways. Something solid is no use to us when

we need to lubricate something, but it may be just the thing for building permanent structures. But to *describe* something as solid and attend to the differences that *would* make is not yet to adopt any of those as directing us to act in any particular way. We can consider what differences it would make without acting upon any of them in a given instance; the description informs us of practical differences without prompting action on them. To grasp a piece of descriptive discourse does not immediately commit us to acting in pursuit of any of the interests that the descriptive content might serve.

When we assert a normative declarative sentence, we do in fact adopt some interests (more or less explicitly) and privilege some forms of action guidance for their service of those interests. A normative declarative like "This is how you should hold the violin" (perhaps coupled with a demonstration) both commits the speaker to the defense of that manner of action-guidance (*how* to hold the instrument) and attributes a commitment to the audience act in some privileged fashion (*actually holding* it in certain ways). We can offer similar formulations of this for privileged goals that do not involve such limited ranges of action, as we will see later in this chapter. To make such an assertion is to be subject to familiar responsibilities to defend one's position in the light of reasons. The reasons will require us to use the sort of cataloging of practical potentialities offered by descriptive declarative sentences and assess how they will serve the interests covered by actions taken in accordance with the normative claim. ("You should hold it this way because doing so allows you to move your fingers more easily.") Our adoption of a normative declarative may be conditional: "If there is a fire, we should help our disabled colleague get out of the building" does not commit us to act immediately to help them leave the building. But to assert any such sentence is still to commit to perform or permit some privileged range of actions or the pursuit of some privileged goals. For any normative declarative sentence that we might assert, we will face an open question roughly of the form, "Will doing things *this* way accomplish what it needs to, all things considered?"

The most straightforward sort of evaluation here is comparative and instrumental. Whether one should hold an instrument in a particular fashion may be evaluated in light of the results of doing so (a player's ability to reach all parts of the fingerboard, the tone a player generates, etc.).

Other ways of holding the violin may be equally effective, or even superior in facilitating the playing of music. But we state our case in terms of "service" to our interests to stave off the temptation to read these matters exclusively in such comparative, instrumental terms. A normative claim is appropriate just in case the manner in which it guides our actions serves those interests in some privileged fashion, and there may be different bases for privileging them. The most obvious way in which we might privilege the way in which something serves an interest is by being the *optimal* manner in which we could pursue that interest. Thus, if X is the most efficient manner in which we can satisfy that interest in some measurable way in comparison with alternatives that we can imagine, then we often take that as a license to assert that one should do X (or be X, or use X, etc.). For instance, we have such a license to assert that certain scientific questions ought to be investigated with double-blind studies. To leave them out would be to pursue our interests in sub-optimal fashions, given our epistemic interests, in that false conclusions and dead-ends are much more likely.

However, optimality is not the only form that such assessments of service to our interests may take. Some normative commitments may be constitutive of a status or role we have an interest in pursuing, but which does not admit degrees or leaves no room for alternatives. The status of *being an epistemic agent at all* requires a concern for the truth of our beliefs and utterances; those who exhibit a systematic indifference toward such matters lack a feature that is partially constitutive of their status as agents and leaves them without a place in such practices at all. (It is important not to think of examples of bad epistemic agents here, e.g., those who are dogmatic or sloppy about their beliefs in localized ways. Rather, think of those who embrace a global dogmatism that makes them perversely averse to all questions of truth.) This is not a judgment about optimality, as there are no viable alternatives over which we are privileging it. We might also endorse a particular normative claim simply because something is the default action, practice, and so on and serves a role in a larger array of practices, even though we recognize that other alternatives might serve us just as well and our commitments are contingent or even arbitrary in the larger scheme of things. For instance, wearing darkly colored clothing to a funeral is the norm in Western cultures,

and adherence to those norms contributes to the respect and support of the family and friends affected by someone's death. One *ought* to wear black or gray to a funeral or memorial service out of respect, even if we realize that a community's practices could just as easily compel us to wear bright colors, attend naked, and so on.[4]

As we suggested in Chap. 5, our view is that there are no fundamental or foundational interests, and no interests completely divorced from all others. The modes of evaluating how actions, practices, and so on serve our interests are thus plural and varied, and in evaluating how all of these are to be pursued, we must adopt a type of reflective equilibrium approach by which we strive to articulate the best ways to balance all of those interests. Normative discourse is the part of our language in which we articulate and adjudicate proposals for positions to occupy that equilibrium. For instance, we have interests in understanding the healthy functioning of the human body, but in formulating our approach to medicine and research, we balance those interests with concerns for the welfare and dignity of potential subjects. Reaching reflective equilibrium here involves a projection of how our interests will manifest themselves for us in the future, how our practices and actions will bear on their satisfaction, and how our successes and failures in meeting those interests will affect the understanding we currently employ. Even if we agree that an interest like health is as persistent as they come, what being healthy entails is a matter that we will adjust according to changes in technology, agriculture, and a wide range of social practices that agents may opt into or out of in the future. (We will discuss this example in more detail in Sect. 6.2.) This makes the process of seeking reflective equilibrium expansive, complex, and perennial. It is woven into everything we do, every way we evaluate ourselves and our futures, and it is the day-to-day life of anything that can rationally engage the world. Philosophers will not have exclusive purchase on it, nor should we regret that they do not.

In deliberating which sentences to assert, we will look to the actual successes and failures, and to the actual constraints placed upon us by the conditions in which we undertake our practices. Those conditions will be reasons for us to make normative claims, but the normative claims themselves will not describe the conditions that license the assertions. The sorts of reflective equilibrium that we reach in the pursuit of our

interests described in Chap. 5 will give us rational grounds to make these endorsements, subject to dispute and challenge on various empirical grounds. Sentences and beliefs with action-guiding content do not purport to represent non-normative objects and conditions in giving their action-guidance, but we will be bound to attend to those conditions in evaluating and deliberating over the normative claims we affirm.[5] And in this way, normative claims both will not be the mere expressions of our attitudes or allegiances, and they will not be merely the ways that communities voluntarily elect to talk with each other. This will give us a robust notion of objectivity without representation or correspondence, as we shall argue at greater length in Chap. 7.

Before we move on, we should first make explicit a response to a potential objection here. Imagine an interlocutor objecting:

> *Hasn't this become a supervenience account, just like the ones you rejected in Chap. 2? You seem to be committed to supervenience in one of two ways: either (i) you are committed to saying that the truth of normative claims is constituted by the actual deliberative processes of extant communities, or (ii) you are committed to saying that the truth of normative claims supervenes on sets of interest-satisfying conditions external to our practices and communities.*

We think the first of these is simply a misreading of our position. At no point have we *identified* the truth of normative claims with the deliberative processes of extant communities. To do so would be to give non-minimal truth conditions for normative claims—"'N' is true iff norm-deliberative practices D entail N"—a move that we have denied our account undertakes with its deflationary commitments. To look more directly at the view we have offered, option (1) would be a commitment to some view that reverted to treating normative discourse as a kind of description of the output of a set of deliberative social practices. As we have emphasized, this is simply a misconstrual of the content of normative discourse. To make the charge that we are committed to (1) is to mistake the epistemic, reason-giving role of deliberative discursive practices for a constitutive one, and this is simply not our view.

The second challenge here is more subtle, but our response will involve some of the same features. "Interest-satisfying conditions" here should be

read as designating whatever non-normative conditions matter to norm-deliberative practices, either as successes in our pursuits or facts about the absence of features that run counter to them. Thus, diminishing the causal influence of cognitive biases by certain methods (making us more likely to make true claims), or the prevention of various bodily harms by concrete efforts to reduce those effects would be relevant interest-satisfying conditions.[6] The objection in (2) is that we are entitled to assert a normative declarative sentence if in fact the interest-satisfying conditions are met, and not if they are not, and therefore, what we say with the normative sentence must somehow supervene on them. And this challenge is more substantial than the first, because there is a sense in which it comes closer to our own view; we *do* want to say that non-normative properties like inhibiting the effects of bias or bodily harms *are* relevant to normative deliberation. On our view, normative claims will not be descriptions, but it makes a difference to those normative claims whether in fact the things we aim to do get done. But this does not entail a commitment to supervenience as an *explainer* of the content of normative declarative sentences. At most, this requires the very weak sort of global supervenience described at the end of Sect. 3.3. There, we agreed that a view would be suspect if changes in the truth of normative claims were possible when there were *no* changes in the global sets of non-normative facts (and thus swung completely freely of the non-normative world). However, those non-normative facts still do not determine the truth of the normative sentences, as the interests in play in reflective equilibrium will not be among those non-normative facts. Indeed, the interests will not be substantive properties at all, and attempting to place the goal-directedness of our actions and practices, as properties in the natural world drains them of their character, as we also discussed in Chap. 3. So the non-normative world thus matters to our judgments, but not in a manner that compromises our account.

6.1.3 A Botany of Action-Guiding Modes

So if there are genuine normative claims and the sentences involved in them are truth-apt, but we are denying that there are normative proper-

ties for normative vocabulary to designate, how should we approach an account of the meaning of those sentences and the overtly normative vocabulary that occurs in them? We would argue that such discourse is best understood in terms of the inferential roles of its sentences and subsentential expressions. Inferentialist semantics have many critics, of course, but normative discourse is a region of natural languages that is particularly well-served by attention to the use being made of these expressions and poorly served by efforts to describe their content in terms of representation. To drive this point home, we will give a very general botany of the sorts of practical import normative claims express. In Chap. 5, we discussed the role that interests play in licensing normative claims, and have since made an initial case that these claims could be evaluated in a manner that made them truth-apt even if there was no purport to describe objects and properties in familiar ways. We might think of that as an extended examination of what might serve as an "input" condition that leads us to make a normative claim. (Normative terms can occur in non-declarative language as well, and those need an account, but we are focusing on truth-apt language in this work.) To flesh out our account of action-guiding content, we need to say more about how it could be said to "guide" action in a non-causal sense to give us the relevant sorts of "outputs"—commitments to act, avoid, or value something. What this will give us is an articulation of what normative discourse does that is itself normative and practical "all the way down," telling us what is incumbent on speakers in the management of commitments and entitlements. We will thus supplement existing accounts in the inferentialist and expressivist literature with three elements here. Normative declaratives will have (a) a target audience, (b) a focus, and (c) an action-guiding mode or modes. The last of these will elaborate the sense of action-guidance in the most direct and distinctive manner, but the different modes included there will presuppose a target audience to guide and a focus on which to guide them.

A normative declarative will have a target audience, not simply in the sense that an act or inscription is performed in the presence of other speakers, but that its assertion expresses the attribution of commitments and entitlements invoked by its content to the target audience. This is most familiar in first- and second-person attributions: "You should apologize"

with the relevant "you" being settled by the details of a speech act in a conversational context, and similarly for uses of "I." A third-person attribution ("Flynn should apologize to Liz") will do much the same, even if those third parties are not present in the conversation. But a declarative may also attribute commitments to larger classes of agents, even if no particular agent is its target audience in a given utterance or inscription ("*Members* must pay their dues by 31 December"). Even a declarative sentence made in the most dialectical fashion, not pointed at any partner in conversation or restricted to any class of agents like, "Murder is wrong," attributes commitments to a target audience in this sense: it attributes them to all possible agents. (Those attributions can be challenged and defeated, but the declaratives here still express that attribution.) The target audience can thus have the widest possible scope in some cases, and a much more narrow one in others. But this should not obscure the point that the declaratives here are being asserted, and the attribution of those statuses is not a report of the speaker's attitudes. Declaratives in normative discourse will express a sort of to-be-doneness that is "agent-neutral" to borrow a term from Kukla and Lance (2009); that is, they guide with an authority that does not belong to the agents in virtue of who they are, as a command from a superior to a subordinate might (though the target audience may be a single person, addressed by another, as in the "You should apologize" example above). Making such claims will by default commit the speaker to their content, just as making an assertion of a non-normative declarative would, but this is a matter of squaring our own practical commitments with what is being asserted, not a means of reducing that content to the attitudes.

Normative declaratives will also typically invoke some set of non-normative commitments on which its action-guiding commitments and entitlements will then be focused. ("Non-normative" here meaning that the vocabulary with which we express them is not overtly normative, e.g., physical vocabulary.) We will call this set of commitments the *focus* of a normative declarative. Thus, a normative declarative such as

(1) Murder is wrong.

invokes for speakers the sorts of commitments we take up in asserting that something a murder (that someone dies, intent, etc.), and expresses

a familiar sort of aversive action-guidance that we will describe in further detail below. Were we not disinclined to use representationalist semantic vocabulary here, we might just say that the focus is what the declarative is *about*. This might be taken to imply the need for some designatum that it represents, though. A more complete articulation of the focus of a normative declarative would involve an elaboration of the sorts of reidentification commitments by which we typically assert that some action was (for instance) murder, what other claims might evidentially license the assertion of (1), and what further commitments and entitlement would follow from it (e.g., that agents would be entitled to step in to prevent it). Note that a focus is not a syntactical category like being a subject term in a sentence. "Murder" occupies that position in (1) and focusing the action-guidance require the descriptive commitments associated with that expression. But "You should use a No. 10 scalpel for abdominal incisions" focuses on surgical procedures and the tools used in them, not whoever the "you" addressed here might be. As this example also suggests, the focus may be complex, with numerous discernible sets of commitments related to one another. "Terminally ill patients should have access to the most aggressive forms of pain management" may be broadly said to have medical practice as its focus, but this is more informatively articulated as a complex arrangement of commitments involving patients, their prognoses, and dosage levels for medications. As we shall elaborate in Sects. 6.2 and 6.3, the focus of a normative declarative will regularly be a range of actions and conditions with many different degrees of overlap and similarity to them. Thus, complexity will be the norm, no pun intended. We say a focus "typically" involves some non-normative commitments because it may be possible to generate some normative declaratives that express tautologies without having a focus that can be specified descriptively. "Goods are good" and "You should do what you should do" are intelligible, apparently normative, and trivially true, but do not narrow their action-guidance at all with the help of non-normative vocabulary. We take it that these should still be treated as declaratives, however, and that it is best to think of their foci as the broadest possible construal of some normative category (*all* the goods, *all* the obligations, etc.). This makes them less immediately informative (as tautologies generally are),

but will also allow them to serve as general statements governing the use of thin normative concepts.

These first two features of the content of a normative declarative could characterize descriptive declaratives. That much should be expected, as we find them woven into the larger linguistic landscape in just the ways we do other declaratives. What distinguishes them is what we are calling their *action-guiding mode*. A normative declarative will guide responses—actions, commitments, and entitlements—for its audience with respect to its focus. Each mode will have a type of valence as it, were, guiding its audience on the output side of its inferential role by licensing a distinct array of exit-moves involving action types or outcomes on which we focus. (Paradigmatically, "X is right" would have licenses to do X as consequences, as well as material inferences to other normative declaratives such as "We value X" and "X is to be encouraged.") That guidance may be favorable or disfavorable toward the focus, demanding, permissive, discouraging, and more in various combinations and degrees. It might also have some types of neutrality we will elaborate. We might say that normative declaratives express a practical indicating or directing what is to be done with their focus—the content they share with descriptive declaratives. (Action-guiding modes will belong to whole sentences, though the role of overtly normative vocabulary will often be strongly linked to a particular mode.) Our suggestion here will at least remind readers of the pragmatic dimensions added in many speech acts (e.g., making use of a declarative sentence like "'The door is open" as a [somewhat coy] command in a conversational context and thereby adding an additional pragmatic dimension to it in the act that is not necessarily present in the meaning of the sentence itself). But we are suggesting that at least for normative declaratives, the action-guidance is an essential dimension of the meaning of the sentences themselves, not a feature added by some acts of utterance. While some of our vocabulary will be reminiscent of conventional speech act categories, we are not talking about speech acts here. Fine-grained details distinguish the different items of normative vocabulary from one another; we cannot hope to map this terrain entirely. But for our purposes in this work, there is no need. We suggest that there are a number of broadly shared general types of guidance at work in the action-guiding modes of normative declaratives. (It is thus

compatible with what we offer here that there be other modes for other types of discourse, even for non-declarative parts of normative discourse. We make no claim to completeness in that sense here.)

The different action-guiding modes at work here can be characterized in terms of what commitments and entitlement to act we attribute and undertake when someone asserts a given declarative. These may be attributions to ourselves when we are making the assertions. What we attribute and undertake will often be symmetrical, that is, we both attribute a commitment to someone making an assertion and undertake it ourselves. But we may also attribute a commitment or entitlement without undertaking it (as when we note someone's view, but hold off on adopting it ourselves), and can undertake commitments that we do not attribute to others. Each of them may be thought of as a kind of normative character expressed by a sentence, or in some ways, by individual overtly normative expressions. They are the different ways in which a normative declarative sentence can be said to guide our actions, as we have been suggesting. We take these to be fundamental dimensions of the meanings of normative sentences and subsentential expressions, and so an important caveat is in order here. If these modes are fundamental to the meaning of every element of normative discourse, then our account cannot make use of logical operators fit for whole sentences, particularly negation, in articulating them. So we take it that there are both positive modes of action-guidance and negative ones, as we shall elaborate, and not simply fewer modes and their negations. We concentrate on five modes in particular here: *direction*, *proscription*, *endorsement*, *repudiation*, and *permission*. Some of these terms are perhaps reminiscent of other accounts of pragmatics and deontic logic, though we will make distinctive use of them here. We should also note that for many normative claims, the modes will overlap (e.g., some claims express both endorsement and direction) while others will be incompatible with one another. We will return to this point shortly.

A piece of normative discourse may be said to express *direction* when it attributes a commitment to perform an action or one of a class of actions, or else pursue some privileged goal. Pieces of normative discourse that express direction express that some action is to be done, or that some type is generally so. This would be a stronger sort of practical commitment

than endorsement (which we shall elaborate shortly), in that one may endorse some action or outcome without attributing a commitment to actually take it or pursue it. In these ways, expressions of direction will be reminiscent of commands in indicating what is to be done, but they will not emanate from the authority of one who uses the declarative, as commands do. Thus, a second-person attribution expresses direction most overtly and familiarly:

(15) You ought to go apologize to Maria.

To say this to you is to direct you to go perform an apology, which seems very much like a command. But the direction expressed here can come in sentences that are not in the second-person form most common to imperatives, nor even implicitly involving a speaker delivering an imperative to an audience. The second-person features of (15) are not essential to the mode that it expresses. For example, we could make a parallel assertion from a third-person perspective, or as an assertion of a generalization:

(15′) Anne ought to go apologize to Maria.
(15″) Anyone who insults someone without reason should apologize to the person who is insulted.

Each of these claims expresses direction to an agent or agents in general to act in certain ways, though not as a command. Someone uttering (15′) need not say this to Anne or Maria, and whoever utters (15″) to an audience may do so with no knowledge of whether insults have been lobbed or not. But all of them have the import of attributing a commitment to perform some course of action to an agent or agents. They *direct* someone (or everyone) to do something, or do something in some particular way, but they present this as a matter not constituted by the attitudes of the speakers or their own personal authority.

A declarative that expresses *proscription* attributes a commitment to reject or avoid some action or outcome. When we express proscription, we do not simply negate direction. Sentences that express proscription introduce a rejective stance without expressing further affirmative

commitments. Other commitments and entitlements will follow from sentences that express proscription, but they are not simply negative functions performed on those other sentences. Proscription-expressing claims will be most familiar from ethics:

(1) Murder is wrong.
(16) Insulting Maria was cruel.
(17) Breaking promises is forbidden.

Nor does expressing proscription entail the endorsement of an action's complement. To say that murder is wrong, expressing proscription as you do, does not entail the endorsement of non-murderous actions as a whole. Proscribing murder does not thereby license or endorse beating someone so severely that they are disabled for the remainder of their lives, for instance, even though that is a non-murderous action. In many cases, some of what lies outside the range of proscribed actions and outcomes will actually be far worse and even more strongly proscribed. ("Like you said, I didn't break my promise to Joe. However, I did murder him.") Associations between proscriptions and other endorsements will require attention to context and detail, even though some generalizations will hold with very few exceptions. We think that much of this line of thought is at least implicit in the literature on practical reason, but we emphasize it to make clear that proscription is a primordial type of practical import in its own right, not merely a derivative of some other more fundamental one.

A declarative may also express *endorsement* of an action, outcome, or present state. Expressing endorsement of an action or action type in the present sense implies not only allowing it to happen but also affirming its performance and the status of agent(s) who perform it. Where I express endorsement for X, I not only affirm the actions, commitments, and entitlements of those who do X but also undertake those commitments myself. This complementary undertaking of what we affirm is a hallmark of purport to assert objectively. The same would go for endorsement-expressing sentences made collectively, as when an organization declares its support for something. Where we affirm without undertaking, we are stepping back to talk of something like further opinions. "Steve believes

Thetans are the source of all human suffering. I acknowledge his beliefs."
Placing things in the third person here does not involve any use the
action-guiding mode of endorsement; indeed, we put things this way to
express our reservations. But expressing endorsement need not be focused
on a course of action, nor even direct agents in general to perform one.
Endorsement-expressing declaratives frequently affirm outcomes or con-
ditions without thereby affirming a particular route to them, or attribut-
ing a commitment to pursue that outcome to anyone. For instance:

(18) You did the right thing by apologizing to Maria.
(19) Education is a social good.
(20) A *p*-value less than 0.01 strongly justifies accepting a hypothesis.

   All of these express endorsement, but they do not direct agents toward
particular courses of action. (18) is not a call to someone to apologize to
Maria again, or to keep apologizing regularly; (19) is not a call to anyone
to serve as a teacher, structure public schools in certain ways, and so on;
(20) does not call on us to become scientists. At most, they would seem
to direct our management of some of our own commitments. Should
someone perform an experiment and get such results, we should endorse
their hypothesis, undertake subsequent research with that as an assump-
tion, and so on. Such provisional guidance is different from the sort of
guidance that is given when a normative claim expresses direction. A
declarative that expresses endorsement will regularly have claims that
express direction for someone as consequences, but we should not con-
flate the two types of expression, nor think of endorsements as somehow
veiled directions of our actions.
   Normative declaratives can also express *repudiation* for their focus,
which inverts and complements endorsement. This is not just withhold-
ing entitlement to some party, but overtly refuting the action-type or
outcome in question.

(21) Flynn's comment to Liz was very cruel.
(22) This policy is oppressive toward women.

   Much like expressions of endorsement, expressions of repudiation have
a character that steers our conduct by focusing not necessarily on action

types, but rather by a type of valuing (or perhaps, *dis*-valuing) of the focus. An expression of repudiation does not tell us what to do instead; it tells us to disvalue its focus, leaving the courses of action we should take instead for subsequent determination. While an expression of repudiation will almost always imply a proscription on some action or action type, or at least direction to refrain from whatever is being repudiated, these are not necessary or expressively equivalent consequences. We can assert that something is a harm even in cases where we affirm the actions or action types that lead to it. When we face genuine moral dilemmas and other hard choices, we may find that harms are done, but refraining from imposing them would have been even worse. To call on a familiar example, the sacrificial victims of out-of-control trolleys (whether trapped on the tracks or thrown from footbridges) may truly be harmed, even in cases where it would putatively be far worse to refrain from imposing those sacrifices.[7]

Finally, claims can express *permission* for an action or outcome. We would characterize this as attributing entitlement to some party to take an action or let an outcome stand, without endorsement or rejection.

(23) Women are free to bear children.
(24) Students may opt to write a paper in lieu of taking the final exam.
(25) In the absence of strong evidence to the contrary, one may step out of the way of the charging bull.

(We take it (23) expresses a type of license here, rather than merely describing the absence of physical restraints.) Where we do express endorsement or direction, permission is also attributed. Not only *may* you do something or let something stand in such cases, you *are to do it*. But we can express permission for someone to take an action or let an outcome stand even where we do not express direction or endorsement of such things. We may think that your career choices, medical decisions, political positions, or many other matters involve commitments that we would never endorse, and still permit you to take them up. Such is the impact of treating agents as autonomous. This extends to theoretical commitments as well: at least on matters of leading-edge research where the jury is still out, all sides are permitted to pursue inquiry in

their respective veins. There is an old adage in deontic logic that "all that is not forbidden is permitted," and we would agree that there is an incompatibility in proscribing and permitting at the same time. We are wary of packing the logical notion of negation into the sort of practical import that we are describing, however. While this may be a useful way of approaching the formalization of deontic attitudes (as deontic logic aspires to do), we are of the view that the sorts of action-guiding content and management of commitments and entitlements are prior to formal logical tools in the order of explanation, as we said earlier. The role of negation, both in everyday use and in formal systems, should be articulated in terms of differences in practical import, not presumed by them.

A few quick general points to note in closing this section. It bears noting that there are complementary pairs at work among the modes here—direction and proscription, endorsement and repudiation. And while there is a sort of neutral mode "between" direction and proscription (permission) that serves as a license without endorsement or repudiation, there is no correlated mode "between" endorsement and repudiation. We suggested that the modes of endorsement and repudiation were about valuing and disvaluing (rather than urging action or avoidance). But to neither value nor disvalue would not seem to be say anything normative at all. More than one of the modes we have articulated here may also be expressed by a given sentence. Some sentences express both direction to act in certain ways and permission us to do so, for instance. Since action-guiding content is not akin to representation, this overlaying of different types of import need not create conflict in the way that two different representations of facts might. On the other hand, some combinations do give rise to incompatibities, e.g., we cannot simultaneously express endorsement of X in general and proscription of X in general. Most such incompatibilities will be local to the actions and outcomes involved. But, all expressions of direction commit us to expressions of endorsement and permission of their focus, and all expressions of proscription commit us to expressions of repudiation. We may also find the modes mixing by the inclusion of overtly normative subsentential expressions: "Murder is wrong" expresses proscription, but "murder" itself expresses repudiation; "Beneficence is a duty" expresses direction, but also endorsement. So long as the modes are compatible as we just described, doubling up in

this way poses no difficulties, though direction and proscription will take a kind of contextual practical precedence because they actually guide us to do/not do something, rather than just how to value it.

If a sentence directs its target audience to perform an action, or let an outcome stand, to assert it is also to express endorsement for doing it or its being so, even if that is sub-optimal. For instance, a doctor may say, "We should amputate the leg before the patient dies of gangrene." Sentences that express direction and proscription will imply endorsement and repudiation (e.g., "beneficence is a duty" implies "we value beneficence"), but statements that are only endorsement or repudiation will not necessarily imply thick direction/proscription statements. They are reports of statuses with respect to valuing non-normative foci, but not necessarily indications of what to avoid or do about it. Exceptions to all five types may be expressed locally in terms of general commitments to which present cases do not apply. Think of these as a category of "yes, but…" responses. We can generally express permission for people to eat peanuts, while proscribing it for a child with severe allergies. We can generally express proscription toward snowball sampling techniques,[8] but permit conclusions based on them when subjects are very difficult to find.

6.2   "Good" and "Right"

We have argued against representationalist accounts of normative discourse. If we are to give a thorough-going expressivist account of normative discourse, then it is incumbent upon us to explain the role of those normative terms which most plausibly seem to play a representational role in the language: terms like "good," "value," and so on. We also noted in Sect. 4.1.4 that there is a substitutional, expressivist interpretation of quantification (as opposed to the objectual, representationalist interpretation), and so quantifying over K need not ontologically commit one to the existence of K's. We note that a range of overtly normative vocabulary that we could call "goods talk" guides action not by the action-guiding mode of direction, but by privileging some outcomes. Its modes thereby tend toward endorsement and repudiation. We will find a range of "thin" normative terms, such as "good" itself, but also a range of "thick" terms

that apparently introduce more non-normative content into their mix, like "healthy" or "cruel."[9] We should elaborate why the latter does not commit us to a substantive property in a way that we have rejected. But we should also note that both thick and thin expressions in goods-talk appear both in adjectival positions and as singular terms that often appear to suggest reification:

(26) Maria looks healthy.
(27) Some forms of vegetarianism can be healthy diets.
(28) Health is our focus in this program.
(29) Charity is good.
(30) We now have good evidence.
(31) Goods are of primary importance in consequentialist accounts.

(We are not defending the truth of any of these claims, though they are all truth-apt.) We want to deny that sentences like (28) and (31) entail substantive ontological commitments by their apparent reification, so we will refer to such occurrences as "quasi-reification." Explaining the expressive role of this will be another major task in this section, and it will carry over to talk of virtues.

6.2.1 The Open-Endedness of Goods-Talk

When we talk about various goods, the best way to understand such talk is as expressing endorsement for an open-ended range of behaviors and outcomes, rather than naming a fixed object or property in the world. Consider a concrete example of a good, such as health. On our account, to say that health is a good is to endorse a range of behaviors and outcomes—eating a balanced diet, brushing one's teeth, and so on. Of course, the list is open-ended in various ways. We might have good evidence that a certain course of action (say, providing universal health care) promotes the good of health, but such a course of action might be opposed for other reasons; and it is open-ended in that we simply do not know all the ways to promote health (for example, we do not know of a cure for AIDS). But the list is open-ended in a more fundamental way. Health is not a property, like the utilitarian's pleasure, which can serve

as a fixed anchor for our moral theorizing. When we try to pin down what, precisely, health is, we find just another set of endorsements and social practices—empirically constrained, to be sure—but nothing like the objective fact or property the representationalist would like to find giving content to the term "health."

It is natural to think that "healthy" is a biological property, that even if it cannot be wholly naturalized, it can be give a representationalist treatment. But an expressivist treatment is more appropriate to the openness that characterizes our engagement with health. For starters, what counts as health is conditioned not only by various empirical conditions but also has a strong dependence on social practice. Consider one feature of concern in estimates of our health: urinary continence. Most contemporary westerners adapt to social settings in which indoor plumbing is a given and its appropriate use is expected. This requires training our muscles in early childhood, and various ailments can prevent us from exercising this control later in life. But for those in some other social settings, the plumbing and the expectations that go with it (for us) are not the norm. European colonists first arriving in the western hemisphere in the sixteenth century were dumbstruck by the casual, immediate way in which members of various indigenous groups would relieve themselves on the spot. This figured prominently in many images of the west at the time, and lent one further detail to their portrayal as not quite fully human (Rubies 2009, 123ff). The idea of waiting for an hour or more to relieve oneself (as contemporary westerners often do) would be alien to those in such social settings, and doing so on command in a new social setting would not be immediately possible for those not having trained their muscles to continence. But we should hesitate to call this a health problem, even if it has a biological basis. In our modern era, incontinence is treated as a health deficit, but notice again that the very problem itself is created by a set of practices and expectations about continence, practices, and expectations which are not inevitable or universal.

Also, various empirical conditions alter whether a person might appropriately be considered healthy: strength, endurance, mobility, and so on. Both authors of this book are healthy by the standards of the early twenty-first-century USA. But of course, if you were to transport us back 20,000 years, we would be considered disappointingly unhealthy because of our inability to participate in the persistence hunt of wildlife. Transport us

back 500 years, and again, we would be considered unhealthy for our inability to harvest wheat for 10 hours per day in late summer. (That is, if you transported us back as members of the peasant class. As members of a different class, we might be considered exemplars of fine health, even possessing all of our own teeth and not crippled by childhood disease.) To say that we are not healthy because we cannot perform the tasks of our distant forebears is to engage in an objectionable kind of essentialism, and probably to commit some form of the naturalistic fallacy. Neither of us needs to run down a kudu or harvest the wheat. Such things are not required of us in our social setting, with its complex division of labor (much of it mechanized). Nowadays, one requires much less in the strength, mobility, and endurance departments to count as healthy.

Goods like health are dependent not just on the historically contingent state of our social practices and various empirical conditions, but also on various other normative appraisals and endorsements we make. For example, whether heterosexuality is part of psychological health has been a hotly contested issue, and it is only in the past few decades that Western psychiatrists have come to see homosexuality as a normal variation on human sexuality rather than as a pathology. Various empirical inputs were involved in this decision. Particularly important was Evelyn Hooker's (1957) groundbreaking research on homosexuals. Instead of studying homosexuals who were already in contact with the psychiatric or criminal justice system, she recruited a group of 30 well-adjusted homosexual men and a control group of 30 similarly-situated heterosexual men, administered a battery of psychological tests to each, and had a group of experts blindly score the tests and try to figure out which test subjects were homosexual and which were heterosexual. (The experts did no better than chance.) But the American Psychological Association's 1973 decision to remove homosexuality as a mental disorder from the *Diagnostic and Statistical Manual of Disorders* (*DSM)* was a decision not based purely on empirical grounds, but also on evolving commitments on sexual ethics more generally and homosexuality in particular. It was a decision made in large part on moral grounds. And so our conception of health—mental and physical—is not separate from our evolving set of moral endorsements. Thus, the more we try to pin down the notion of health, the more we find more endorsements, more levels of contingency and fluidity.

What we have said of goods-talk here may be extended, with some modifications, to talk of virtues. One can *be patient*, wait *patiently*, and count *patience* among their virtues. Much of virtue ethics and virtue epistemology already are expressed in terms geared toward action-guidance. There may also be debates about how fundamental the virtues are in ethics or epistemology, but debates of those sorts do not undercut the expressive role that those distinctions play (aside from claims about which ones are fundamental, obviously). And there are corresponding issues concerning the correlates (if any) of the terms of virtue-talk. To what degree should virtue theories in ethics and epistemology take the ascription of virtues to agents to be ascriptions or posits of real, enduring psychological traits of the agents? Some theorists have embraced this question with enthusiasm and given thoroughly naturalized accounts. In epistemology, virtue reliabilists, such as Sosa (2007, Chap. 2, 4–5), Goldman (1992), and Greco (1999), have accepted that justified belief must be the product of a virtuous intellectual character, and identified that character with certain reliable belief-forming processes. These approaches have been complemented by virtue responsibilists, such as Code (1984) and Zagzebski (1996), who offer accounts that more deeply interweave virtue epistemology with virtue ethics, all of which has a more deeply normative character without purport to a naturalistic reduction of the traits. There have however been extensive challenges to the viability of virtue ethics in recent years by "situationist" philosophers and psychologists who challenge the reality and explanatory efficacy of character traits as ethicists have postulated them.[10] As we have said, it is not our aim here to defend any particular position in normative ethics or a general account of epistemology, but we are at least sympathetic to the responsibilists here, if there is to be worthwhile talk of virtues at all. That is, if they survive situationist challenges, but only as the reliabilists describe them, it would simply make them instrumental distinctions and shunt normativity off somewhere else.

6.2.2   Quasi-Reification

Thus far, we have offered a set of tools for understanding the action-guiding character of normative claims, as well as a set of reasons for

doubting that even "thick" normative vocabulary should require non-normative correlates. We face a sticky point, though: *we sure do talk like there are such things a lot*. Singular terms in normative discourse appear to function just like singular terms for things to which we more happily take ourselves to have ontological commitments. "Good," "evil," "cruelty," "justification," and so on do not appear that different as pieces of vocabulary from "electron," "testosterone," "molybdenum," or "Jaco Pastorius." As we said in Chap. 4, ontological commitment need not be read directly off of surface grammar, and we have made a case for not doing so with the expressive resources outlined here and systematic reasons elaborated in Chap. 3. We will call such cases *quasi-reification*—those in which surface grammar suggests designation, but we prefer to explain expressively and pragmatically. But this strategy creates an obligation for us. We need to say *why* discourse expressing the sorts of action-guidance we have described would so frequently take such forms.

An empirical hypothesis about linguistic competence does suggest itself here: we may do so simply because the mechanisms of linguistic competence we acquire suffice for these expressive purposes (with some qualifiers) and additional categories would be difficult to develop. There may just be a natural optimality or near-optimality to the grammars of our natural languages, whether we follow Chomsky in thinking they are innate or others in saying they are not. If so, we would be better served by using the grammatical tools at our disposal in different pragmatic styles (descriptive, normative, metaphorical, etc.). Singular terms in normative discourse would not be so surprising in that case. But as we said, this is an empirical hypothesis for linguists, anthropologists, and cognitive scientists to settle, not humble philosophers, so we mention it simply as an interesting proposal in passing. Even so, it would be an underwhelming response on its own. We want to say something more informative about why we use singular terms in precisely the ways that we do in normative discourse, not simply note that there is nothing else readily available. So instead, we will elaborate how different occurrences of cognate pieces of normative vocabulary (like (26)–(28) and (29)–(31) above) are inferentially related to one another in ways that are conducive to various purposes that normative discourse serves.[11]

We take it that the least contentious starting point for thinking about normative declaratives in expressive terms will be those cases in which overtly normative vocabulary occurs in predicative positions of atomic sentences, as in cases (1), (16), and (17) above. If we accept the account of action-guiding modes outlined in Sect. 6.1.3, we can understand the content of such sentences as declarations of statuses that guide action without committing ourselves to a substantial property or entity as an explainer of those expressions' meanings. "Murder is wrong" can express proscription in the necessary ways without our needing to posit a substantive property of *wrongness*. It may be easier to accept this for thin normative expressions at first, but thick normative terms occurring in such positions can be treated as expressing one or more of the modes while also committing and entitling us to additional restriction commitments to narrow the scope of actions and assertions to which the mode applies. For instance:

(30) Your comment to Flynn was very cruel.

Here, the thick normative term "cruel" plays a part in expressing repudiation, but to assert (30), one would incur a more specific commitment to confirm that the comment had intentionally caused some pain or distress, and if the assertion of (30) were warranted, one would incur an entitlement to sanction or punish whoever had made the comment.[12] If we could not affirm that there was pain or distress caused, and thus could not assert this, we could not call the comment "cruel." Thick normative terms can thus play a part in the action-guiding modes, but channel those expressions of endorsement and repudiation in more narrow ways that suggest which sorts of reasons would be deployed for the repudiation. Thus, even though "cruel," "sloppy," and "irresponsible" all play a part in many expressions of repudiation, they bundle in with that descriptive commitments that indicate the sort of reasons that would be deployed in the defense of a claim being made.

This notion of a restriction commitment can then show us how those overtly normative terms serve as adverbial and other adjectival modifiers (e.g., "a *cruel* comment" or "I gestured *appropriately*"). Those occurrences will express a mode, further restricted if the expression is thick, which is

then restricted still further by the expression being modified. So "…a good outcome…" expresses endorsement with almost no restriction commitments at all; "…a cruel action…" would express repudiation, restricted by commitments to confirm pain and distress, but without much restriction via the modified noun ("action"); and "…a *cruel* comment…" expresses repudiation with even greater restrictions (only to certain utterances) via both terms as we have elaborated.[13] And these complex expressions with some overtly normative vocabulary in them can, depending on the context of their utterance and the sentences in which they appear, involve commitments only to token actions and objects ("I made *a cruel comment* to Flynn") or to type-level ones ("It is just as wrong to make *a cruel comment* as it is to break a promise").

Many such restriction commitments are possible for any overtly normative expression, of course. Comments, actions, gestures, and many other things can all be cruel (or good, or correct, etc.). In many cases, consideration of all such possible variations with a given overtly normative expression is our aim; we sometimes want to say something about cruelty in general, rather than just cruel comments or the token cruel gesture someone just made. We can first begin to understand overtly normative singular terms occurring as stand-in expressions for quantified phrases. "Cruelty" can stand in for "everything that *is cruel*," "the good" can stand in for "everything that is good." That is, those overtly normative singular terms will sometimes be substitution-inferentially equivalent to the quantified phrases; making an inference from a sentence involving one to a sentence substituting the other in will be a materially good one, and vice versa. And some of those occurrences will even permit us to express counterfactually robust principles, such as moral laws. But it would be a mistake to think that these sorts of stand-ins for quantified phrases exhausted the roles that overtly normative singular terms can play. We say this is "sometimes" what overtly normative singular terms do because, depending on the context of their utterance, tokens may function differently. While "cruelty" is always going to express repudiation, it may express it for all possible actions, those in the actual world, some restricted subset, attitudes toward those ranges on the part of a particular person ("the cruelty of Vlad the Impaler"), hypothetical attitudes of hypothetical agents ("Someone who relishes their cruelty is immoral"),

or many other possibilities. Overtly, normative singular terms can thus appear in generalizations that have a counterfactual guidance analogous to natural laws in descriptive discourse (i.e., they can be used in principles like "murder is wrong"), but they can also have their scope restricted to actual circumstances ("*Harm* was done to my client, your Honor," "*Mistakes* were made in the analysis of the data," "*The cruelty of Vlad the Impaler* was well known").

So each occasion of these overtly normative singular terms could be made explicit as a more complex phrase with the expression occurring in a predicate and a further array of expressions acting as restriction commitments. We would thus make explicit whether we were expressing endorsement, repudiation and so on of all possible ranges of actions and outcomes, or of some limited range of them in the context of a given utterance. (We would speak of this as elaborating "what we mean by 'cruelty' here" in everyday terms.) But to speak with a singular term here is to use terms that leave those specifications to their contexts of interpretation, while carrying over the action-guiding modes (endorsement, repudiation, etc.) universally associated with the expression; all these variations on "cruelty" express repudiation, all the variations on "good" express endorsement, and so on. But we can then leave such restriction (or none at all) to interpretation in context, while invoking its action-guiding core commitments. And we can make inferential moves back and forth, from theoretical generalizations to more restricted claims, using the same relatively simple and familiar sorts of expressions and substitution-inference patterns. Thus, there is an inferential variety, fecund connections to draw, and a kind of abstraction accomplished by singular term usages that contributes to all of it.

6.2.3   A Last Note: "Right" and "Correct" and Expressions of Direction and Proscription

We close this chapter with one last set of elaborations of the expressive roles of normative declaratives. As the previous section was primarily devoted to "good" talk (and "bad" and other closely related notions), this last section is primarily devoted to "right" and "correct" (and "wrong"

and other closely related notions). Many of the same observations we made in the previous section apply to "right," and so on, as well. Just as we saw quasi-reification and singular term usages of good-talk, we can see it for right-talk (e.g., "a *right* to free speech"). But the real contrast to draw in making this last set of points is that we can turn our attention to expressions of direction and proscription to notice some important features of their finer textures. So, we turn to the ways in which we might say "x is *right*" or "x is *incorrect*" and other overtly normative expressions that express direction and proscription of actions. But this will also encompass expressions like (15) where we say "You *should* apologize" or "You *ought to* apologize." What we note is that where these action-guiding modes are in play, the focus of their direction and proscription will vary more greatly than one would first expect.

Here, we take a cue from Wilfrid Sellars, as we have several times in this work. In "Language as Thought and Communication," (1969) Sellars distinguishes two types of rules and rule-boundedness that are often run together. On the one hand, we may speak of rules of action, or as he calls them, "ought-to-do's." As the expression would suggest, these are expressions of direction or proscription on categories of action. Concern for these will be familiar in ethics, but as we emphasized, any issue surrounding them will extend beyond the ethical to include evaluations of correct action as well. "Thou shalt not kill" and "Drawing conclusions with a *p*-value over 0.05 is unwarranted" will be in the same boat. These are contrasted with rules of criticism or "ought-to-be's." These entail no particular course or type of action, though they do confer commitments on agents to bring about or preserve the conditions they include. Many courses of action may be satisfactory in addressing these commitments, though no generalizations about such actions will be entailed. So the demands of a rule like "Clocks ought to read the correct time" may be met in particular cases by building more precise clocks, checking them more regularly, switching to an online source, and so on, and some of these may be more satisfactory courses of action than others in different circumstances. (The online option would be a bad idea for anyone without a reliable connection, for instance.) A rule of criticism in a theory of distributive justice stating, "Every person should have access to adequate health care" might be addressed with a single payer system, a public/pri-

vate mix with considerable subsidies, a wholly private system with considerable regulation, or many other political and economic structures. The focus of a rule of criticism can have many possible satisfiers, and we often have some flexibility in how to pursue them. We typically do not *care* whether our properly functioning clocks are analog, digital, or streaming, so long as we can project one or another of those options as a reliable way of meeting our interests. Those projections might be wrong (we might find that electronic clocks depend on an environmentally unsustainable set of practices), and some ancillary features will certainly rule out some possible courses of action (killing off the sick and the old is not acceptable health care policy).

Rules of action lend themselves to a more direct reading off of social practices (you have to do X in this community, you cannot do Y, etc.) that are often made explicit in canonical forms. This is not to say that their justification thereby becomes relativistic, but that taking them as explicit codifications of practices we take to be justified is more straightforward. By instituting such rules, we ascribe commitments to perform or refrain from courses of action to a target audience bounded by our practices. A rule like "Thou shalt not kill" ostensibly commits all agents to a range of prohibited actions; a sign saying "Employees must wash hands before returning to work" confers commitments to a much smaller group. Many rules of action would also apply specifically to tasks or conditions under which commitments would be conferred ("Always look both ways before crossing the street"), whereas others would confer standing commitments across all contexts. One such as "Parents should seek medical care for their children in emergencies" wears its commitments on its sleeve. For each child, a particular set of people have a particular set of commitments to perform certain actions. But implicit in this is a set of entitlements, as well. Being the parent of a child entitles one to make certain decisions and take certain actions that it would be impermissible for others to take, all things being equal.[14] We judge an action right if it satisfies the commitments and makes permissible use of the entitlements conferred on agents by their standing in sets of social practices. Determining that they have done so may take a great deal of work, but this constitutes the expressive core of normative discourse using the language of right or correct action, and indeed anywhere the modes of direction and proscrip-

tion are in play. What bears attention here is that normative declaratives in this vein will generally depend on well-circumscribed conditions and categories of actions. While we can make a declarative like "You ought to do the right thing," this will strike most speakers as hopelessly thin.[15] Not all of our interests, nor our service of them, enjoy such specificity, though. The difference leads us to rules of criticism.

Rules of criticism thus occupy a practical role that straddles those of rules of action and the "good talk" described in Sects. 6.2.1 and 6.2.2. They express direction or proscription with respect to their focus, but do not include specific courses of action that will satisfy the call. We are exhorted to act to secure it, but with an open-endedness about how we might do so. "People with disabilities ought to have access to all public spaces and buildings" expresses something normative, but it does not direct us to any particular course of action. It may imply or at least strongly suggest some rules of action in context (ramps are to be built, doorways are to be widened, etc.), but we will not generally find specific, informative, "thick" rules of action to replace rules of criticism. Again, this carries over to non-ethical claims as well. If we assert that scientific theories should be coherent, or as simple as the data will permit, this does not entail particular rules for which posits to drop or how to rewrite laws. Rules of criticism incorporate normative distinctions that are not readily associated with simple sets of non-normative properties, or even static bundles of properties. What it means for a theory to be coherent or a community to be oppressed will involve a great many features in concert with one another in particular combinations, so rules of criticism will frequently require more context-sensitive handling than rules of action. Rules of action may also involve degrees of context-sensitivity, and we are wary of the suggestion that there will be any genuinely exceptionless rules of action. But the important point for the moment is the marked difference in the relation between the focus of the sentence and the markedly less restricted ranges of action to which we are directed or proscribed when it comes to declarative sentences that express rules of criticism.

What attention to the expressive role of talk of right/correct action shows is that these differences are in fact consistent deployments of a common pragmatic core. They vary in their specificity in directing action and coordinating behavior, but this variation is itself a response to the

conditions in which the discursive practices are embedded. The features of living as beings in the world, embodied as we are, make the commitment not to kill in cold blood very strongly and very narrowly associated with shared behaviors conducive to our interests. We can thus expect communities to adopt rules of action in cases where all or almost all instances of a type of action will have some feature that is strongly conducive (or contrary) to their interests. Thus, almost all communities have rules of action against cold-blooded killing. We enjoy more latitude in keeping time and distributing medical resources, even if it is equally urgent that we do so. We would expect rules of criticism where our projections of what is conducive (or contrary) to our interests do not include such rigid associations. This may make it sound as though we are simply biding our time with rules of criticism until a more precise declarative sentence can take their place. But this is not the case. The open-ended quality of rules of criticism will suit many ways that complex arrays of interests bear on complex arrays of conditions. Where projection is especially fraught, open-ended direction like this is appropriate. For instance, what would make for the right balance of sympathy, generosity, sternness, and countless other factors in the rearing of one's children will be effectively impossible for us to determine all the contours of the project ahead of time to declare rules of action. "Parents have a duty to foster healthy development in their children," directs us to seek these goals in relevant ways, whatever may come. Epistemic norms benefit in similar ways from rules of criticism; "We should favor simplicity in theory construction," does not direct us to make specific trade-offs in advance. This strikes us as especially important to understanding the irreducibility of normativity, and will figure prominently in our discussion of epistemic norms in Chap. 9.

Notes

1. DeRose and Grandy acknowledge this general pattern, but go on to argue for an account that unifies biscuit and normal indicative conditionals, albeit one that accommodates the differences noted here rather than denying them.

2. Apparently conjunctive complex commands also sometimes have a hint of ordered sequentiality about them. "*Give me 20 and drop*!" seems impossible, even though it includes the same conditions, because one needs to precede the other to be satisfied. This gets worse for longer series: "*Go to sleep* and *go to bed* and *brush your teeth* and *put on your pajamas…*"

3. The role is really more subtle, in that the first and second instances of "you" should be anaphorically linked. So it is closer to saying, "If *anyone* can't take the heat, then *he* should get out of the kitchen."

4. To get a sense of how such a norm might be defeasible, note that beloved entertainer Jim Henson left an explicit request that those attending his memorial should *not* wear black in an effort to make the event more consolatory and uplifting to those grieving the loss (Blau 1990). Adherence to the usual norm would be *disrespectful* on such an occasion.

5. This will require additional detail for some thick normative terms, which follows in the next section.

6. "Various bodily harms" obscures the fact that we would need to decide which effects to bodies are wrongful or inhibit normal functioning(and thus harms) in the first place, but we will set that aside for the moment.

7. As Sect. 6.2 will elaborate, we would be wary of any normative ethical theory that identified any non-normative category with harms, or even established one-way entailments (e.g., "all pains are harms"). But even so, there remains the possibility of the sort of horrible dilemma described here.

8. These are research methods in which you enlist the subjects to find other subjects. They are thus highly susceptible to bias, but may be valuable when a study's target is averse to being identified, for example, early epidemiological studies on HIV and AIDS in the 1980s.

9. This is a familiar point in the metaethics literature, but we were helped in formulating it by conversations with Steven Levine.

10. We discuss situationist challenges at some length in Chap. 9.

11. Numerous parts of the rest of this section draw from arguments made in (Wolf 2002), although the subject there was a very different class of expressions.

12. We might also have to confirm that the pain or distress was unnecessary, but there will have to be pain or distress in any case. The restriction commitment in the example is not intended to be exhaustive of all such commitments.

13. The restriction may also be "predicative" (e.g., to assert that something is a green table is to assert that it is both green and a table) or "attributive" (e.g., a good actor is good *as an* actor), as Mackie (1977) and others

noted. Overtly, normative vocabulary strongly tends to function attributively in this sense.

14. Exceptions abound here. Children may be looked after by guardians other than their biological or legal parents, the parents may not be present when emergency care is needed, and so on.

15. The exception here might be cases in which "the right thing" is already obvious, or specified in context, or used euphemistically. (Our grandparents' generation used this as a way of referring obliquely to a hastily arranged marriage.)

References

Austin, J.L. 1970. *"Ifs and Cans" in Austin, Philosophical Papers,* 2nd ed. London: Oxford University Press.

Ayer, Alfred Jules. 1952. *Language, Truth and Logic*. New York: Dover Publications.

Blau, Elanor. 1990. Henson Is Remembered as a Man With Artistry, Humanity and Fun. *The New York Times*, May 22.

Code, Lorraine. 1984. Toward a 'Responsibilist' Epistemology. *Philosophy and Phenomenological Research* 45(1): 29–50.

DeRose, Keith, and E. Grandy Richard. 1999. Conditional Assertions and 'Biscuit Conditionals'. *Nous* 33(3): 405–420.

Geach, Peter. 1960. Ascriptivism. *Philosophical Review* 69(2): 221–225.

Gibbard, Alan. 1990. *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.

Gibbard, Alan. 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press.

Goldman, Alvin I. 1992. Epistemic Folkways and Scientific Epistemology. In *Liaisons: Philosophy Meets the Cognitive and Social Sciences*, ed. Alvin I. Goldman, 155–175. Cambridge, MA: The MIT Press.

Greco, John. 1999. Agent Reliabilism. *Philosophical Perspectives 13: Epistemology* 273–296.

Hooker, Evelyn. 1957. The Adjustment of the Male Overt Homosexual. *Journal of Projective Techniques* 21(1): 18–31.

Horgan, Terry, and Mark Timmons. 1991. New Wave Moral Realism Meets Moral Twin Earth. *Journal of Philosophical Research* 16: 447–465.

Horgan, Terry, and Mark Timmons. 1992. Troubles on Moral Twin Earth: Moral Queerness Revived. *Synthese* 92: 221–260.

Horgan, Terry, and Mark Timmons. 2000a. Copping Out on Moral Twin Earth. *Synthese* 124(1–2): 139–152.

Horgan, Terry, and Mark Timmons. 2000b. Non-descriptivist Cognitivism: Outline of a New Metaethic. *Philosophical Papers* 29: 121–153.

Horgan, Terry, and Mark Timmons. 2006a. Cognitivist Expressivism. In *Metaethics After Moore*, ed. Terry Horgan and Mark Timmons, 255–298. Oxford: Oxford University Press.

Horgan, Terry, and Mark Timmons. 2006b. Morality Without Moral Facts. In *Contemporary Debates in Moral Theory*, ed. Jamie Dreier, 220–238. Oxford: Blackwell Publishers.

Kukla, R., and M. Lance. 2009. *'Yo!' and 'Lo!': The Pragmatic Topography of the Space of Reasons*. Cambridge, MA: Harvard University Press.

Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. New York, NY: Penguin Press.

Peirce, Charles Sanders. 1878/1992. How to Make Our Ideas Clear. *Popular Science Monthly* 12: 286–302

Rubies, Joan-Pau. 2009. Texts, Images and the Perception of 'Savages' in Early Modern Europe: What Can We Learn from White and Harriot. In *European Visions: American Voices*, vol. 172, ed. Kim Sloan, 120–130. London: British Museum Research Publication.

Searle, John. 1962. Meaning and Speech Acts. *Philosophical Review* 71: 423–432.

Sosa, Ernest. 2007. *A Virtue Epistemology, Volume 1: Apt Belief and Reflective Knowledge*. Oxford: Oxford University Press.

Stevenson, Charles. 1937. The Emotive Meaning of Ethical Terms. *Mind* 46: 14–31.

Stevenson, Charles. 1944. *Ethics and Language*. New Haven, CT: Yale University Press.

Wolf, Michael P. 2002. The Curious Role of Natural Kind Terms. *Pacific Philosophical Quarterly* 83: 81–101.

Zagzebski, Linda. 1996. *Virtues of the Mind: An Inquiry Into the Nature of Virtue and the Ethical Foundations of Knowledge (Cambridge Studies in Philosophy)*. Cambridge: Cambridge University Press.

7

Objectivity and Normative Discourse

We have already discussed a number of arguments for the view that normativity cannot be reduced or otherwise placed in the natural world, as many naturalists would insist it must be to remain legitimate. However, we have also argued that we should not thereby move to a form of non-naturalism that places the sources of authority for our claims outside the natural world. Such a choice was forced upon us by a set of false assumptions, and we can move beyond them with the sort of broadly pragmatist interpretation of normative discourse that we have offered in the last three chapters. On our account, normativity need not be placed as entities and properties in the world, but we also look to the world in which we are embodied in speaking and judging normative matters. We thus made the case that the incorporation of normative discourse with other forms of discourse about the natural world remained a worthwhile goal. This, we have argued, is a sense of naturalism worth embracing.

A parallel point could be made for the objectivity of our judgments. Objectivity has long been cast in representational terms that we have eschewed, casting the world as static, our representations as categorically removed from it, and the business of philosophy and science being to double-check the correspondence between the two. If we reject

such a picture of inquiry, it might be argued that there is no need for a notion of objectivity. Perhaps it would even be a dangerous holdover of Cartesianism. No philosopher has made this case more (in)famously than Richard Rorty, who urged us to embrace solidarity rather than objectivity. But we would make a case here that parallels the earlier one about naturalism. Even if we reject the form that a notion of objectivity takes in other philosophical paradigms, there will still be an epistemic distinction it expresses that is worthy of our concern. We have emphasized the world-involving character of normative discourse in the last three chapters with an eye toward this point. In this chapter, we elaborate how discourse in general and normative discourse in particular can be objective, even if we do not embrace representationalism.

7.1   Normativity, Objectivity, and Relativism

What was the motivation behind the drive for naturalism in the first place? It seems that for most naturalists, particularly of a reductive bent, the naturalistic basis was supposed to serve a critical purpose, satisfying important constraints: the naturalistic basis to which normative facts were to be reduced was supposed to provide an objective grounding for normative truths, while situating them within the prevailing scientific worldview of analytic philosophers. Thus, to give an overly simplified example, if it could be demonstrated that moral rightness were simply identical with utility maximization (where utility maximization itself were identified with some empirically measurable item such as preference satisfaction), then not only would we have provided a non-relative standard of moral rightness, but we would have situated our moral theory if not within the laws and models of the hard sciences, then at least within the empirically respectable and testable branches of the social sciences. Thus, moral properties and facts would not be queer, spooky, or supernatural, but would be ordinary natural facts of a kind that are perfectly well countenanced within the ontology of natural science.

As we discussed at the beginning of Chap. 3, most naturalists recognize that the entire content of a normative statement cannot be reduced to a natural fact. As we saw there, most philosophers (especially moral

philosophers) distinguish between the *extension* of a normative term (which they will happily naturalistically reduce) and the normative *conceptual* content of such a statement, which cannot be naturalistically reduced, due to the is-ought gap. As we saw, this troublesome normative conceptual surplus could not be reduced, and prevented the reductionist from being able to offer a fully naturalistic account of normative discourse.

The supervenience theorist is left with a similar normative surplus. Consider the argument of McNaughton and Rawling:

> For present purposes, we can agree that reasons are descriptive facts. What is always a *sui generis* normative fact, we maintain, is the further fact, on the occasions when there is one, that some fact is a reason. Suppose it is a fact that
>   (E) A would enjoy fell-walking.
> Then there is a further fact: the fact that
>   (F) (E) gives A a reason to fell-walk.
>   It is (F) that is a normative fact…[T]he fact that we have a (practical) reason to Φ, or a (theoretical) reason to believe *P*, is a *sui generis* fact distinct from the reason itself, which is our (descriptive) circumstance, since when we have a reason, this is *because* of our circumstance. (2003, pp. 30–31, 42–43)

Plausibly, then, even if normative facts supervene on descriptive facts, the situation is the same as with reduction: it is the descriptive, factual component that is located in the descriptive, subvenient base. The normative surplus is an additional fact, over and above the fact which is constituted by these descriptive facts. Thus, a naturalist like Brink (1989) gives a naturalized account of the extension of normative terms or utterances, but leaves the actual normative component unreduced.

But what is distinctive about normative discourse is its action-guiding aspect. And as we have seen (particularly in Chap. 3), for the naturalist, this aspect is an additional, surplus feature that is not readily accounted for by the part of the theory that seeks to explain how our theory fits into the natural world. The element in our normative theory that is naturalized (whether reductively or non-reductively) is not the normative

element, but the factual element; as Brink or McNaughton would say, the prescriptive, normative element is an additional fact, on top of that factual, descriptive element. But if only what is naturalizable is worth preserving, and this action-guiding part cannot be naturalized, then we as philosophers must either turn a blind eye to this "rogue" element of our theory, or we must strip from our theories that very element which makes them what they are in the first place. Indeed, the action-guiding element seems prima facie incompatible with the commitment to fit normative discourse into a thoroughly naturalistic world-view, since it would seem that normative discourse's fact-stating role (but not its action-guiding role) is that which is most easily assigned a counterpart in the descriptive world. Given that this action-guiding role is the sine qua non of the normative, our suggestion was that we should start with an account of the normative that favors action-guidance over the fact-stating or descriptive element of such discourse.

Does normative discourse even need to serve a fact-stating role? As we have noted, it is common among naturalists (particularly moral naturalists) to distinguish between the fact-stating role of normative utterances and their normative or prescriptive element. But the claim that normative claims state descriptive facts is significantly threatened by Harman's "explanatory challenge." Harman is writing about morality, but the argument generalizes to all forms of normative discourse. According to Harman, moral facts neither figure in the best causal explanation for any observations, nor are reducible to facts which are explanatory. Harman infers from this that there are no moral facts.[1] From this, Harman further concludes that moral nihilism must be true. But does this final conclusion follow? Do we need moral facts to have objective morality?

What follows from causal-explanatory eliminativism? At most, it shows that *ontological eliminativism* holds about a type of discourse. To be an ontological eliminativist about a particular subject matter is to claim that certain noun-like words or phrases belonging to that subject matter do not refer. Most of us are ontological eliminativists about unicorns and phlogiston; many are ontological eliminativists about holes, shadows, and the like. Thus, we are willing to concede (for the moment)

something like the following entailment: if K-talk is not involved in the best causal explanation of any uncontroversial phenomenon, then the terms of K-talk do not refer.[2] The mistake is in thinking that ontological eliminativism entails *doctrinal eliminativism* (the claim that all normative utterances are false, or that the discourse plays no legitimate assertive role). Numbers may fail the causal-explanatory test, yet doctrinal eliminativism about numbers is not a highly plausible view.

Timmons diagnoses the inference from the absence of moral facts (of the moral realist variety) to moral error theory as arising from a series of widely shared assumptions, chief among them one he calls the semantic assumption: "All genuinely assertive discourse is descriptive discourse" (1999, 130).

The semantic assumption leads the debate surrounding moral realism and anti-realism to have a characteristic shape. Those who wish to pursue an expressivist or conventionally non-naturalist line are forced into a variety of (in our view) uncongenial positions. Thus, Blackburn (a projectivist, although not an expressivist—see Blackburn 1993), arguing that descriptivism must be false, concludes that non-cognitivism must be the correct way to go. Copp (2001), who has one of the better-developed expressivist theories (influenced in large part by Gibbard's seminal work), nevertheless feels the need to yoke his expressivism to a realist metaphysics of moral properties in order to preserve the descriptive and realist character of moral discourse. But once one rejects the semantic assumption (a rejection that Timmons makes, and which we made most explicit in Chap. 6), one is free to argue for an account of normative discourse on which such discourse is genuinely assertoric, but which is not in the business of describing a non-normative reality. We think this is merely a rediscovery of a truth Sellars recognized half a century ago, and which we already cited in Chap. 1: "[O]nce the tautology 'The world is described by descriptive concepts' is freed from the idea that the business of all non-logical concepts is to describe, the way is clear to an ungrudging recognition that many expressions which empiricists have relegated to second-class citizenship in discourse, are not inferior, just different." (Sellars 1957, p. 282/§79)

Is a model of normative discourse in which this discourse is not fact-stating viable? As we see it, there are two chief conditions on the adequacy of such discourse:

1. The judgments made in this discourse (e.g., moral judgments in a moral discourse) must be non-relatively true, and must be subject to defense by giving and asking for reasons which are themselves held non-relatively.
2. The basis for (1) does not commit us to any non-natural entities, properties, or causal relations that cannot themselves be accommodated by physical or scientific theory.

The first adequacy condition we will call the Non-Relativism Principle; the second, the Naturalism Principle. Failure to satisfy (1) is particularly worrisome once you jettison the fact-stating role of morality: as we noted above, one chief motivation behind the move to naturalism was to anchor normative discourse in something objective (where, for the naturalists, this is most naturally understood as something independent of humans, on the model of a scientific fact). Once those anchor lines are cut, then the threat of relativism becomes acute. If, for example, morality is tied to social practice, but only by contingent or voluntary adoption by a community's members, then what is to prevent an intolerable diminution of the authority of our normative claims? ("*We* think genocide is wrong, but…")

The second principle is obviously important because even though we are rejecting reductionism and supervenience as viable ways to carry out a consistent naturalistic project, we are still committed to some version of naturalism as the best understanding of the world and what it contains; and a theory that commits us to non-natural objects, properties, or powers is clearly unacceptable.

One obvious tie to the natural that any normative theory will have is in criteria of application.[3] In order to defend against charges that normative theory just spins in the void, without a grounding in facts, natural facts (*that it caused her pain*, *that she wore a red dress*, etc.) must be among those that serve as the criteria of application for normative utterances. (We have already discussed, in Chap. 5, why we think these concerns

about "spinning in the void" are misplaced, and we will return in Chap. 9 to consider the robust relations between the non-normative and the normative.) None of this is to claim that criteria of application can be exhaustively stated in natural terms; no doubt they cannot. Indeed, a lesson from the arguments over moral particularism has been that you cannot formulate a set of principles of the form "Do A in C," because the characterization of C will itself involve moral terms (e.g., "That caused her *undeserved* pain," "Those people are *innocent*," etc.) and require moral understanding; and thus one would need moral understanding and mastery to apply the moral principles in the first place.

Although we agree that normative utterances should have the above-described tie to the natural world via criteria of application, we in no way regard ourselves as obligated to demonstrate the nomological derivability of normative utterances from any set of natural facts, in the way a classical reductionist might demand. Indeed, it should be clear from the above that we are doubtful of the prospects that such a project can be successfully carried out. As we argued in Sect. 3.4, it is altogether likely that the set of actions, events, and properties that properly license a particular normative utterance (such as "That was cruel") across a range of circumstances will appear gerrymandered and shapeless if looked at purely at the level of the descriptive. Recall that McDowell argues that this may be so *even if* the normative supervenes on the natural:

> Supervenience requires only that one be able to find differences expressible in terms of the level supervened on whenever one wants to make different judgments in terms of the supervening level. It does not follow from the satisfaction of this requirement that the set of items to which a supervening term is correctly applied need constitute a kind recognizable as such at the level supervened upon. (1981, p. 145)

Therefore, not only should we not expect to be able to give a set of entailments or reductions from the normative to the descriptive but as we argued in Chaps. 5 and 6, cruelty will not constitute a recognizable property at the non-normative level. As McDowell notes, the criteria of application for a normative concept might be shapeless at the level of the descriptive, and certainly will not support law-like gener-

alizations. Part of this problem stems from the above-noted fact: some of the criteria of application for a normative concept will essentially involve other normative concepts (e.g., "S is innocent" may be, in a particular set of circumstances, a criterion of application for "You may not intentionally direct harm at S"). The holistic interconnection of normative concepts complicates the attempt to draw any reductions or nomological entailments between the descriptive and the normative. But also, as we noted in Chap. 3, the simple variety of things and states—acts, omissions, objects, people, institutions, and so on—which can exemplify a quality such as "being unfair" makes it implausible in the extreme that when we make these attributions, we are attributing a property, all of whose instantiations share some non-normative feature (or even a related bundle of non-normative features). Other metaethicists who take a non-descriptivist outlook agree, although not necessarily for McDowell's reasons. Timmons, for example, writes that while we base our moral evaluations on features of the world, "I reject… any attempt to identify a realm of moral facts or properties with such FEATURES. To do so and then claim that moral statements function primarily to report or describe such FEATURES…would be to accept a version of descriptivism, which, I have argued, is not the way to go" (Timmons 1999, p. 154).[4]

As we have already noted, we are committed to an extremely weak version of the supervenience thesis. However, we understand this not as an attempt to explain normativity in terms of the non-normative, but as a principle of parity, an exhortation to treat like cases alike. Our weak version of the supervenience thesis is, in fact, not intended to establish normative naturalism, and is neutral among competing theories of the normative: if the criteria of application are not naturalistic, that does not in any way show the moral theory is not itself naturalistic; these are separate issues. A principle such as "Aid your friends when they are in need" can have non-natural criteria of application (and possibly non-natural consequences of application, too) if your friend Joe is being haunted by a ghost, but this does not show that our moral theory is itself non-natural. Morality could still have non-natural criteria and consequences of application even if moral rightness were, say, a natural property (like *maximization of preference satisfaction in humans*).

So to sum up Sect. 7.1, our suspicion is that most traditional naturalist accounts cannot preserve that which is the hallmark of normative language: action-guidance. We also think there are powerful arguments against the existence of normative facts, arguments which further push us away from traditional reduction and supervenience accounts. Both of these considerations press us to develop an account of normativity which starts with the distinctly action-guiding element of normative discourse, and downplays the fact-stating role of such discourse. The question that remains is whether such an account can be developed which satisfies our above-stated adequacy conditions. We believe it can, and it is to this discussion that we now turn.

7.2   Normativity within Social Practices

We have stated that an account of normative discourse must satisfy two adequacy constraints, the Non-Relativism Principle and the Naturalism Principle. We have also indicated that our preferred account of normative discourse is not committed to the existence of normative facts. Does that mean we have given up, right off the bat, satisfying our two principles? Can you have a non-relative conception of normativity that denies normative facts?

7.2.1 Normativity and Rule-Following

One of the chief arguments of this section will be that if an account of normative discourse satisfies the Non-Relativism Principle but does not involve any facts or truth-makers, then it also satisfies the Naturalism Principle. This may at first seem counterintuitive, but it gains plausibility on further reflection. One's ontological commitments are not and should not be a separate matter to be investigated independently of the theory one defends. Rather, one's ontological commitments grow out of what is required to make the theory work (and then, if such commitments cannot be sustained in the face of further investigation, so much the worse for one's theory). But if in the course of theory construction, it turns out

that one does not *need* to posit an array of facts and other supporting items, then one can proceed with theory construction in their absence. And in their absence, there is nothing to offend against the tenets of naturalism.

And what work were the facts doing in the theory? We have already argued, briefly, that the descriptive facts which were supposed to form the objective basis for normative discourse were not capable of doing the heavy lifting of the theory, but it is time that we returned at greater length to this point. The impossibility of grounding the normative in matters of fact has been convincingly argued by Wittgenstein and later commentators on Wittgenstein's work such as Kripke (1982) and McDowell (1984). What we wish to do in this and following sections is further argue that descriptive facts are irrelevant to the legitimacy of normative discourse because they cannot ground the normativity of such discourse. We will be led by this discussion into a social practice account of normative discourse which offers not non-normative truth conditions, but justification conditions on normative utterances, on the grounds that non-normative truth conditions (grounded in descriptive facts) would lead to the very kind of relativism or historicism that would violate the Non-Relativism Principle above. We will hold that normative statements are still *truth-apt*, but that this may be accomplished without appeal to traditional sorts of truth-conditions. We shall finally argue that this conception of normative discourse satisfies the Non-Relativism Principle without offending against the Naturalism Principle.

What is it to act under a (perceived) normative requirement? Kripke and Wittgenstein famously discuss this in terms of mathematics, although as we will see the lesson they teach can be extended to all types of discourse where fidelity to a norm is in question (which would mean all types of discourse). Wittgenstein gives the example of counting by 2s. One counts "2, 4, 6…" and the next number ought to be 8. Any other answer would be wrong. What do we say if someone goes on incorrectly?

> Now we get the pupil to continue a series (say +2) beyond 1000—and he writes 1000, 1004, 1008, 1012.
>
> We say to him, "Look what you've done!"—He doesn't understand. We say: "You were meant to add *two*: look how you began the series!"—He

answers: "Yes, isn't it right? I thought that was how I was *meant* to do it." (1953, 75/§185)

What makes one performance correct and another one incorrect? What determines the correctness of normatively constrained behavior? The answer most famously attacked by Wittgenstein is that it is a rule which determines the correctness of my behavior. In other words, I arrived at the answer 8 because I followed the rules of arithmetic. This answer is unsatisfactory, though. Wittgenstein's contention is that one must know how to apply the rule, and this would require another rule, if action under normative constraint were construed in terms of rule following. Since this second rule is itself open to interpretation, Wittgenstein argues that a vicious regress is in the offing, if one insists in construing acting under a normative requirement in terms of rule-following. Although Wittgenstein's aim is to discredit certain theories of meaning, his objection works equally well as an objection to various accounts of acting under moral requirements, or any sort of normative requirements. (How do I know I am correctly acting under a maxim of avoiding cruelty? How do I know I have correctly followed this epistemic injunction?)

The other side of the coin, for Wittgenstein, is that normative prescriptions cannot simply be reduced to descriptive regularities. There are two related reasons for this. First, there is no finitely possible truth maker for a normative rule or generalization, as any finite set of non-normative facts is compatible with going on in a way not licensed by the rule. To use Kripke's example, counting "2, 4, 6…998, 1000" does not determine a unique next member of the set. Again, the point generalizes: a set of objects in the world—mathematical objects, moral objects, colored objects, and so on—cannot determine the meaning of a term, cannot determine how to apply the term in novel cases, because no such set of objects necessitates any particular way of going on to future cases. (And this is so even if the cases in question are coupled with a rule, because there is no unique interpretation of the rule.) The second anti-reductionist reason is that the bare object or set of objects does not itself generate a rule or an interpretation. We will return to this point shortly.

How does Wittgenstein answer the puzzle of acting under normative constraint? The first part of Wittgenstein's solution is to claim that

"there is a way of grasping a rule which is *not* an *interpretation*" (1953, 81/§201). Consider the example of following a sign-post. Wittgenstein writes that when we follow a sign-post, we do not interpret it; we merely react as we have been trained to do: "I have been trained to react to this sign in a particular way, and now I do so react to it" (1953, 80/§198).[5]

This suggests that the correct following of a rule is determined by our causal dispositions to engage in certain sorts of behavior, a possibility discussed by Kripke. This solution faces a number of problems. First, we do not have enough dispositions to do the job. For example, I have no dispositions at all regarding the multiplication of 100-digit numbers. The natural response to this is to say that the correctness is determined by the dispositions I *would* have if my mental capacities were expanded sufficiently so that I could multiply such large numbers. This response, however, begs the question. The question of how to expand my mental capacities so that I had the correct dispositions is a question to be asked from *within* mathematics; no answer can be given which determines the correctness of mathematical calculations and which modification of my mental capacities would be the right one.

A deeper worry, however, is that a disposition is not the right sort of item to determine normative correctness in the first place. As McDowell writes,

> no doubt it is true that answering as one does is an exercise of a disposition that one acquired when one learned arithmetic, but the relation of a disposition to its exercises is in no sense contractual—a disposition is not something to which its exercises are faithful. (1984, 329)

For example, suppose that I have a disposition to brush off the seat of my chair before sitting down. Suppose that on one occasion, I sit down without first brushing off the seat of my chair. It seems clear that my behavior, while not corresponding to my usual disposition, was in no sense incorrect. Suppose, however, that I have a disposition to arrive at the answer 4 when adding 2 and 2. If on one occasion, I arrive at the answer 5, I have made an error, an error whose *incorrectness* cannot be accounted for merely as a deviation from a settled disposition. Thus, dispositions are not the appropriate sorts of things to explain normative correctness. Kripke (1982, 23–25) makes a similar argument.

Wittgenstein's solution is to move to the level of social practices. Recall the example of following a sign-post: Wittgenstein writes that when we follow a sign-post, we do not interpret it; we merely react as we have been trained to do. How, he asks, can this mere causal connection be the following of a rule? Wittgenstein's response is that "a person goes by a sign-post only in so far as there exists a regular use of sign-posts, a custom" (1953, 80/§198).

This is the first part of the solution, then: the mistake was to think that the correct following of a rule was an interpretation of the rule at all. It is not an interpretation, it is merely a causal response in the manner we have been trained. The correct response to the sign-post is not an interpretation, it is a trained response whose *correctness* rests on no interpretation or consciously followed rule. (This is not to say that we never consciously follow rules.) This would seem to obliterate the distinction between acting *according* to a rule (which machines and parrots can do) and *following* a rule, which only agents can do. Wittgenstein's response, his reinstatement of this important distinction, is that such a causal response is an example of following a rule (and not merely regular behavior) because it is part of a social practice. Most of us have a rough, intuitive idea of what a social practice is. It is, for example, what anthropologists posit to explain certain regularities in the behavior of a community. We attribute commitments (along with corresponding entitlements) to communities whose behavior we seek to explain. That a bit of behavior (staying inside the lines, stopping at stop signs) is caught up in such a web of appropriatenesses (commitments and entitlements), we will see, makes possible its classification as an *action* (and its performer as an *agent*), as opposed to merely behavior.

One advantage of explaining norm-governed behavior in terms of social practices is that such an explanation allows us to account for normativity without positing spooky non-natural properties or objects; we need only posit social practices, which are sets of commitments to social appropriatenesses (mainly implicit), commitments which we posit to explain the behavior of a community. But such an account leaves at least two questions unanswered. First, both Sellars and Wittgenstein argue that a conditioned response (such as responding to a sign-post as you have been trained to do) counts as an *action*, and is differentiated from mere rule-governed behavior (as the behavior of a parrot, or the behavior

of iron rusting in the presence of oxygen) in virtue of this behavior being incorporated into a social practice. How can incorporation into a social practice change behavior into *action*, and make *agents* out of those performing these behaviors? That is our first question. The second question is equally thorny: it would seem, at first blush, that our social practice account of normativity leads us straight into a rather nasty version of relativism. After all, if it is our community's practice which fixes the correctness of an action, then it seems straightforwardly contradictory to question whether an action which accords with our community's practices is correct. So the second question is this: can we develop our social practice account in a way that satisfies the Non-Relativism Requirement? Let us address these questions in reverse order, as answering the second question will give us the tools to answer the first.

7.2.2   Truth without Truthmakers

Our argument has already led us to embrace a social practice account of normative discourse. But as we have seen, such an embrace leads to worries about relativism. Kripke diagnoses these worries as arising from a conjoining of the views that practices underwrite the normative with the view that they must have truth conditions. Commitment to this conjunction of ideas seems to entail that the truth conditions for normative claims must be in terms of the community's social practices; therefore, "I ought to do X" is true if and only if the relevant social practice licenses X, or would in the long run, and so on. According to Kripke, Wittgenstein escapes relativism by rejecting the second of these two conjoined ideas. In other words, he denies that normative claims have truth conditions. Rather, there are instead justification conditions for normative assertions: "[W]e can say that Wittgenstein proposes a picture of language based, not on truth conditions, but on assertability conditions or justification conditions: under what conditions are we allowed to make a given assertion?" (1982, 74).

It may be thought that this solution only pushes the relativism worries back a step. After all, since we have concluded that normative utterances must be interpreted in terms of social practices, we must give a social

practice account of the community's epistemic norms, as well. And concerns about the ties between practices and the justifiability of normative claims will recapitulate many of our worries here. Consider an example to illustrate this point. Suppose that a community has a practice of slaveholding. On a naïve social practice account, we might think that the claim "Slavery is permissible" is true if and only if the community's practice permits slaveholding. However, following Kripke, we have rejected this naïve social practice account in favor of one which looks for justification conditions, not truth conditions. This leads us to examine the epistemic practices of the community. Suppose the community in question holds a certain religious text to be an epistemic authority, and consults it in matters as diverse as morality and science. This highlights the problem with which we are faced. We attempted to escape relativism by moving from truth conditions to justification conditions, but it seems as though justification will itself be a relative matter. X is justified if and only if it is supported more strongly than any other belief or course of action by the community's epistemic standards. If the community's epistemic practice is structured according to careful reading of this religious text, then that seems to be the only available standard of justification. Suppose that this religious text seems to endorse (or at least tacitly approve of) the practice of slavery; in our hypothetical community, the claim "Slavery is permissible" is therefore justified.

There is an escape from this form of relativism, as well. A typical community will have in place mechanisms for revising its practice. If you can justify the claim "We ought to do X," then the community is committed to revising its practice so that it includes doing X, if the practice does not already include the performance of X. But there is no reason in principle why the community could not also have in place mechanisms for the revision of the epistemic standards in question. One could, for example, point out internal contradictions in this religious text, or contradictions with archaeological or other types of evidence, and use these to undermine the epistemic authority of this text. Thus, we need not accept any direct entailment from what the practice regards as justified to what is, in fact, justified. These epistemic standards can be just as revisable as the first-order normative judgments (e.g., "We ought to do X") themselves.

Indeed, we need regard nothing as in-principle unrevisable. Some claims may be de facto unrevisable. For example, it is difficult to imagine what could license revision of the rules of arithmetic. Hypothetically, *entertaining* such revisions is not difficult, but imagining grounds for adopting those revisions will be. The important lesson, though, is that nothing is de jure unrevisable—not a community's moral claims; not its epistemic standards; and nor even its standards for revising these epistemic standards. (This point will be familiar from Quine (1953).) Indeed, there is good reason to think that we ought to treat nothing as de jure unrevisable, as we have ample reason to take ourselves to be fallible at any given turn, even after our best efforts. We argued in Chap. 1 that revisability is at the heart of rationality. Physicist Carlo Rovelli makes essentially this point, but casts it more in terms of the importance of the role of revisability in the confidence we have in our empirical knowledge:

> Science is extremely reliable; it's not certain. In fact, not only is it not certain, but it's the lack of certainty that grounds it. Scientific ideas are credible not because they are sure but because they're the ones that have survived all the possible past critiques, and they're the most credible because they were put on the table for everybody's criticism. (2014)

Left-Sellarsians often state this point by saying, "It's norms all the way down." This might mislead us if we imagine an infinite downward progression of levels of justification (or, more accurately, an upward progression through the object level, the meta-level, the meta-meta-level, and so on ad infinitum). But the Sellarsian point is not that there is a regress of levels of justification; the point is that at no point are we simply stuck with whatever the social practice endorses. The Sellarsian point is essentially to endorse a thoroughgoing form of fallibilism. Any claim is in principle subject to correction (something Sellars himself has explicitly stated [1956/1997, §38]). Thus, the move from truth conditions to justification conditions, and the further move toward the revisability of each facet of the social practice (from the object level to the epistemic standards governing which claims are justified at the object level, and so on) is not ad hoc, but, in fact, a natural consequence of a principled

commitment to fallibilism coupled with the disappearance of facts from the theory.

Having said that the theory does not commit us to an infinite regress of levels of justification, one might nevertheless worry that this is what we are committed to by the above. Let us see the emerging picture, and respond to this worry. The commitments and entitlements which compose the normative space of a practice come at a number of levels. At the first level, we have the practice, which consists of a series of communally sanctioned[6] behaviors and implicit normative appropriatenesses (e.g., responding to sign-posts with certain behaviors, responding to objects with certain color judgments, responding to actions with certain moral judgments, etc.). As Kripke and Wittgenstein argue, the solution to Wittgenstein's skeptical problem is to see this social practice as providing a place where agents exhibit behavior that is rule-governed, but which does not involve the interpretation of a rule.

An important feature of languages, though, is that they contain a second level, consisting of explicit normative judgments about the appropriateness of the practice (e.g., "One ought not hold slaves"). These utterances can serve to endorse the practice, or to call for its revision.[7] Such utterances must be adequately justified if they are to warrant revision of the underlying practice; the mere ability to assert a normative claim does not entail entitlement to that claim. This brings us to the third level, which are the epistemic standards (say, the authority a practice accords to appeals to sacred texts, or to prophets, or to double-blind studies, or to Tarot cards) that determine whether a second-level claim has or has not been adequately justified. Even these standards can be challenged, though; at the fourth level are standards for the revision of the given epistemic standards. For example, as we noted above, one might challenge the authority accorded a religious text by pointing out contradictions. Similarly, one might challenge the authority of Tarot cards by conducting a double-blind study which shows that the accuracy of Tarot card readers is at the level of pure chance. Even these standards can be challenged. Indeed, there is no level at which one must endorse the claim "My practice endorses X, therefore, X is right."

To head off any misunderstanding, let us make the following clarification: our picture of the normative has the following "levels":

1. Social practice
2. Explicit normative claims endorsing or revising practice
3. Epistemic standards
4. Standards for revision of epistemic standards

But the relation among these levels is not one of strict hierarchy. Rather, it is one of interdependence. If it were strictly hierarchical, then a claim at one level could *only* be revised by a claim at the next level. For example, you could only challenge a level 2 utterance ("Slavery is wrong") with a level 3 claim showing that we did not use the appropriate justificatory method to justify our level 2 claim. This seems to lead to an infinite regress of levels and rules, or else to an unjustified layer. But this is not how our game of giving and asking for reasons proceeds, or ought to proceed. After all, one can challenge a level 2 claim with another level 2 claim. For example, Smith might try to stop you from releasing his slaves, arguing that he signed a contract to buy them and paid money for them, and you cannot just liberate something he paid for, fair and square, against his will. Here, he is appealing to standards of ownership and private property. You can challenge this with numerous other level 2 claims: cruelty is wrong; people have the right to self-determination; and so forth. In this case, the moral claims you advance defeat the moral claims Smith advances. Thus, we have a case where a level 2 claim is used to defeat another level 2 claim. To use another example, one might use a level 2 claim to overthrow a level 3 claim. Consider a level 3 claim such as, "This religious text is the ultimate authority on matters of religion." One might use a level 2 claim—"Slavery is morally abhorrent"—to argue that, since the religious text endorses slavery, the text cannot be regarded as a moral authority. The same thing happens in science: one can use a level 2 claim (e.g., light travels at 3×10^8 m/s in a vacuum) to challenge an epistemic method, which is a level 3 standard. If this standard yields answers incompatible with what we know about the speed of light, then we have reason to question the authority of this

standard. Thus, we need not be committed to a strict hierarchy that would lead to an infinite regress of rules.

Another thing to notice is that social practice does not merely underlie our level 2 claims. The social practice underlies all the levels. To answer Wittgenstein's skeptical challenge, we must claim that any question of how to go on as we have before, or how correctly to apply a standard (at any level) will be answered in terms of social practice. So the picture emerging of our practice and its standards for revision is looking less like a rigid hierarchy and more like a web of interlocking beliefs, standards, and methods, with the social practice conferring meaning on all of these.

This picture may seem to confine all standards of correctness to within our practice. We hope it seems that way, because that is how things actually are. *Of course*, our standards will be our standards. Whose could they be, if not ours, if they were to be compelling for us? And these standards will be internal to our practice. It is not coherent to speak of judging in terms of standards that are not ours. We can *entertain* alien practices and standards in an anthropological or hypothetical manner, but to do so is just to *suspend* their force to actually compel us. We can revise practices, if we think it would be rational to do so, but the question of whether and how to revise them are questions that arise within the practice, as well. This may spark fears of relativism, but what this does not grant an agent or a community is the rational *immunity* that some associate with relativism. Our practices may (and should) still be open to challenges that *begin* outside their bounds in cases where alien practices and standards can be interpreted as directed toward interests that we recognize as well. Thus, another community's moral rules might bear on our practices when we can see those rules aimed at, say, promoting well-being for dependent groups (like children, or the elderly), an interest that we share. The Non-Relativism condition rightly commits us to saying that agents or communities immunized from reasons and challenges initiated outside their boundaries would be merely mimicking rational, truth-seeking discourse. Our account permits no such insulation, and does not undermine robust forms of objectivity; it only presumes that all such inquiry *begins* at home.[8]

So we have outlined an account of our normative practices which does not ground itself in descriptive facts, nor does it require substantive truth

conditions or truth-makers. The key insight to be had about normative discourse (an insight for which we do not claim credit) is that the point of normative discourse is not to describe some factual state of affairs, but to endorse a course of action or a way of going forward. That this is so can be seen by examining the pragmatic structure of a normative statement. As noted above, a normative utterance might have declarative-like criteria of application, but it has imperative-like consequences of application. Normative discourse serves to express direction, proscription, permission, endorsement, or repudiation for a belief, a course of action, a character trait, a linguistic usage, and so on. When we say, "That belief is irrational," we are not so much *describing* the belief as *expressing repudiation* of it, along a particular dimension. So normative discourse need not commit us to natural facts aside from the familiar ones we use in actions (just as an utterance like "Use a hammer to pound nails" commits us to no "imperative facts," only to familiar facts about hammers, nails, and the like). The appeal to truth makers is just a mistake in the first place.

Does this mean that normative utterances are not truth evaluable? Only if we presume that some relation between normative sentences and some set of facts constitutes their truth. We have already argued at length against representationalist conceptions of truth. There is no single word–world relationship which constitutes the truth. While *the fact that it caused her pain* is certainly a partial reason to take the claim, "Pulling her pigtails was mean and wrong" as true, we cannot merely point to a set of non-normative facts, and declare that it is a relation between these facts and a sentence in our language (the truth-relation) which makes this sentence true. (For starters, as we noted, such a relation cannot account for the action-guiding character of the sentence in question.) For also involved in the truth of this sentence will be our practices of normative evaluations, specifically of moral evaluation, without which it would not even make sense to speak of such things as wrongness, meanness, and cruelty.

Does this mean, then, that we merely need to add social features of our practice to the truth conditions of normative sentences? Again, the answer is "No." As we argued, in Chap. 4, such internalist theories of truth fail to capture important elements of truth, and so it is no advance simply to define truth (in whole or in part) in terms of elements of the

social practice. (We argued above in the present chapter that this is a bad idea.) The fallibilism to which we are committed means that any attempt to identify the truth of a particular normative sentence with any feature of our practice (such as some future or ideal state of our practice) would have to be open-ended in principle. But this move to the future or ideal state, no matter how we specify it, *identifies* truth with some fixed state of a community, violating the demands of objectivity and non-dogmatism, as we elaborated in rejecting internalism about truth in Chap. 4.

But the mere fact that we will not specify a truth-relationship in virtue of which all claims are true does not mean that normative claims are not truth-apt. On a deflationist view, to say "S is true" might serve some pragmatic role within the language, but it does not assert a relation between S and the world. But even in the absence of such a relation, there is every reason to think the sentence is true. For starters, as we have noted time and again, the absence of a word–world relation denoted by the truth-relation does not mean that our normative utterances are cut off from the world. Our normative practices are situated in, *and answerable to*, the physical world, and among the facts, we can appeal to in showing that S is true will be non-normative, physical facts. Consider the normative sentence "Setting cats on fire is cruel." Such claims will be caught up in other, normative claims which we can also appeal to, if (for some reason) someone were to challenge the truth of this. Setting cats on fire causes them pain, it shortens their lives, and so on. Pain is (ceteris paribus) an intrinsic evil, longer lives are, all things equal, better than shorter ones, and so on. And if we can demonstrate that setting cats on fire is cruel, then we have a fortiori (on the deflationist account) demonstrated that "Setting cats on fire is cruel" is *true*. So the deflationist accounts require no additional element for the truth of such sentences, over and above what it is to assert them, and the essential involvement of our evaluative practices in asserting them does not force us into a pernicious relativism.

As we have argued, any claim such as S (and therefore any claim to the effect that S is true) is defeasible and open to challenge. We might all believe at one time that a normative claim is true, and come to decide later that we were wrong. This does not mean that truth is relative, for in coming to a decision like this, we interpret our previous truth-attributions to have been *mistaken*, rather than simply endorsing new ones. So we do not now

judge the US antebellum practice of slavery to have been morally permissible at the time (it was not *then true* that slavery was permissible); rather, we judge now that people who then claimed that slavery was permissible (and thus that it was *true* that slavery was permissible) judged incorrectly.[9] This, we take it, is a great advantage that our account has over internalist accounts of truth. Truth really is external to what we think, or may ever think, because any claim or truth-attribution we make is always in principle defeasible, and could always be overturned with the presentation of new reasons (including novel *empirical* inputs). So we have shown that the justification conditions involved in normative discourse are always defeasible, and that no speakers can claim rational immunity (in virtue of no claim being de jure immune to challenge). Thus, our account of normative discourse satisfies the Non-Relativism Principle. Further, since our account does so without positing any non-natural facts, properties, or objects, it also satisfies the Naturalism Principle.

7.3   Agents, Prescriptions, and Reasons

We noted above that embracing a social practice account of normativity leaves us with two questions. We have already taken steps to address the first ("How can we embrace such an account without being led into relativism?"). But a second question still remains: how can incorporation into a social practice change behavior into *action* and make *agents*—both distinctions rich in prescription and responsibility—out of those performing these behaviors? Let us now address this important question.

7.3.1   From Social Practices to Agency

Sellars grappled with some of the same issues as Wittgenstein, and gave us some clues as to how to answer some of the questions created by Wittgenstein's skeptical solution. A problem often addressed by Sellars was the question, "What is it to learn or apply a concept?" For Sellars, the answer had something to do with rules, but for reasons that should be

familiar from our reading of Wittgenstein, it was clear that learning a concept (or applying a concept) could not simply consist of following a rule:

> Now it is obvious that acquiring the concept of red cannot be equated with coming to *obey* a semantical rule…If there were a semantical rule by learning to *obey* which we could come to have the concept of red, it would presumably be of the form *Red objects are to be called 'red'*…But, to recognize the circumstances to which the rule applies, one must already have the concept of red…One would have to have the concept of red before having it. (1963, 334)

So how do we (learn to) apply concepts, if not by following the rules which (for Sellars) are constitutive of the meaning of these concepts? The first part of Sellars's answer is reminiscent of Wittgenstein's answer that there is a type of rule-governed behavior which is not an interpretation of the rule in question, but wherein the agent in question simply responds as he or she has been trained to do. Sellars integrates the agent behavioristically into the causal order in a similar fashion. The first step in integrating mind into the causal order involves recognizing the importance of what Sellars calls "tied behavior." In explaining our ability to acquire and apply concepts such as *red*, we must consider this process as occurring not in judgment, but as a purely causal process of associating certain words ("red," "blue," etc.) with the deliverances of a pre-existing ability to respond differentially to objects in the environment.

This discussion of the basic element of Sellars's answer to the problem of concept application brings us forcefully up against our original problem: what makes a bit of tied behavior count as an empirical judgment, as opposed to a merely reliable (but not sapient) differential response to the environment? After all, parrots and iron filings respond differentially to their environments, but are not thereby considered agents. A clue to the second part of Sellars' answer can be found here:

> [A]bove the foundation of man's learned responses to environmental stimuli—let us call this his *tied behavior*—there towers a superstructure of more or less developed systems of rule-regulated symbol activity which constitutes man's intellectual vision…Such symbol activity may well be charac-

terized as *free*—by which, of course, I do not mean *uncaused*—in contrast to the behavior that is learned as a dog learns to sit up, or a white rat to run a maze. On the other hand, a structure of rule-regulated symbol activity, which as such is free, constitutes man's understanding of *this* world, the work in which he lives, its history and future, the laws according to which it operates, by meshing in with his tied behavior, his learned habits of response to his environment. (1949, 137–139)

There are two crucial elements in this comment. The first is the idea of the superstructure of inferential activity "towering over" the tied behavior. The second element is the Kantian notion that this superstructure is the realm of freedom (with the un-Kantian point about causation).

As we saw above, Sellars distinguishes between tied behavior and the superstructure of "rule-regulated symbol activity." Sellars writes that the tied behavior and the superstructure mesh with each other in that "certain intra-organic events…function as symbols in both senses, as both free and tied symbols" (1949, 141–142)."[10] Here is Sellars:

[T]he hook-up between rule-regulated symbol activity and the external environment rests on the *meshing* of rule-regulated symbol activity with what I referred to as 'tied behavior'…Now, what misleads these regulists who speak of the sense meaning rules of a language is the fact that in order for the above mentioned meshing of rule-regulated language with tied symbol behavior to take place, *certain intra-organic events must function as symbols in both senses, as both free and tied symbols*. Thus, as children we learn to understand the noise 'blue' in much the same way as the dog learns to understand the noise 'bone,' but we leave the dog behind in that the noise 'blue' also comes to function for us in a system of rule-regulated symbol activity, and it is a *word*, a linguistic fact, a rule-regulated symbol only in so far as it functions in this linguistic system. The noise 'blue' becomes a mediating link between what can suggestively be called a rule-regulated calculus, and a cluster of conditioned responses which binds us to our environment. (1949, 141–142)

It is in virtue of our tied behavior that words such as "blue" signify *empirical* concepts, and hence in virtue of such behavior that our theories

are about the world. And this superstructure allows us to conduct ourselves in what Sellars calls the "space of reasons."

We noted that our practices of following norms and discussing or arguing over their appropriateness can be separated into different levels. At the bottom level, we have the level of the practice. It is at this level of the practice that the object-level norms of our society are found, typically implicit in the tied behavior of the agents in that society. A person will see a light, and will judge that it is red (and this judgment will cause them to act a certain way, such as bringing their car to a stop before entering the intersection). A person will respond to a set of phonemes by producing a different, appropriate set of phonemes. Most of these norms are implicit in the behavior of agents (which is not necessarily to say that you could infer them from the behavior of agents, although perhaps you could). But if humans merely displayed tied behavior, what indeed would distinguish humans from very complicated differentially responding machines? This is where Sellars's superstructure becomes relevant. Agents are not only bound by norms, they can also talk about norms. They can make *explicit* the norms that are *implicit* in their practice, for the purpose of defending, challenging, recognizing, and supporting one another's conduct. They are bound by norms, but they can argue over whether they *ought to be bound* by any particular norm. They can argue over whether the standard for revising norms should be this or that. In short, agents are agents not merely because they respond differentially, but because their responses matter to them, and they attend to norms and appropriate manners of acting in doing so. They attend to their norms because they are the kinds of creatures who engage in "the game of giving and asking for reasons" (to borrow Brandom's phrase)—a game made possible only by the superstructure which allows for the explicit formulation and debate of the commitments and entitlements implicit in practice.

We can see now the sense in which our conceptual system is free. This is the sense in which our concepts, and the theoretical commitments made possible by our possession of these concepts are revisable. That is, our language use is free because we are not stuck with a frozen, unrevis-

able linguistic system of concepts. McDowell, for example, writes that "'Responsiveness to reasons' is a good gloss on one notion of freedom" (1994, xxiii). Sellars directly links a belief system's revisability and the epistemic status of the beliefs in that system:

> *Above all*, the [traditional picture of knowledge] is misleading because of its static character. On seems forced to choose between the picture of an elephant which rests on a tortoise (What supports the tortoise?) and the picture of a great Hegelian serpent of knowledge with its tail in its mouth (Where does it begin?). Neither will do. For empirical knowledge, like its sophisticated extension, science, is rational, not because it has a *foundation* but because it is a self-correcting enterprise which can put *any* claim in jeopardy, though not *all* at once. (1956/1997, 78–78/§38)

Thus, the two elements we have discussed (the idea of the superstructure and the revision of our conceptual system over time) come together in Sellars's picture. Sellars's picture represents, we think, a significant improvement over Wittgenstein's. Wittgenstein seems content, too often, simply to say, "One's spade is turned" and to exempt elements of the social practice from rational scrutiny. It often seems that on his account, the social practice just is and does, and that is the end of it (and where questions come to an end). But Sellars has a somewhat more nuanced account. The social practice makes possible agency, because the social practice makes it possible to give and ask for reasons. And it does this for all of the reasons outlined above—for its ability to make norms explicit, and make the norms implicit in practice the subject of overt discussion and criticism, and so forth. Thus, we regard Sellars' account of agency as an improvement over Wittgenstein's.

There might be some residual dissatisfaction with this answer. The answer we have given—agents are those who are engaged in the giving and asking of reasons, made possible by Sellars' superstructure, and so on—might look circular. If we want to explain what it is to be an agent—what it is to be governed by norms, rather than mere laws or regularities—do not we have to give an answer that is not itself stated in norms?

It should come as no surprise, given what we have said about the relation between the natural and the normative, that we should say no

The "Kantian" View

non-conceptual world   conceptual sense-experience   conceptual thought
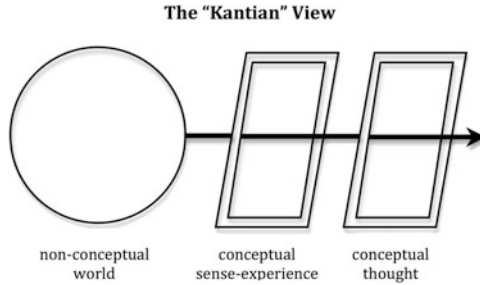
Fig. 7.1 The Kantian View

such answer is possible. To think that such an answer was even possible would commit us to the very reductionism we have rejected early on. And even if one could give necessary and sufficient conditions (spelled out naturalistically) for agency, one could not see the normativity of agency emerge from such conditions. Compare McDowell's instructive discussion of the "sideways-on" view of the world from *Mind and World*. According to McDowell, viewing the world in terms of a non-conceptual component rationally constraining (how?) conceptual thought arises specifically from trying to adopt a sideways-on view of the world, as Kant did(Fig. 7.1).[11]

The view implied here is that the non-conceptual world is something filtered through receptivity and spontaneity in our experience, but that we could then "step back" in a philosophical mode and grasp the non-conceptual world prior to, or independently of, our concepts. But McDowell rejects even the coherence of such a view. For McDowell, all thought is necessarily conceptual (Fig. 7.2). We cannot conceive of the world except as conceptual. So for McDowell, the only way of conceiving the world is "head-on" (Fig. 7.2).

McDowell's picture offers a helpful way of thinking about agents and their actions. We will never see the normativity of the practice from sideways on, and the demand that we do so is misplaced. The purely descriptive features of a living being's activity, viewed from sideways on, will just look like increasingly complex behaviors. Only from inside the practice do these behaviors look like what they are—acting in norm-governed ways, making explicit norms, defending and criticizing
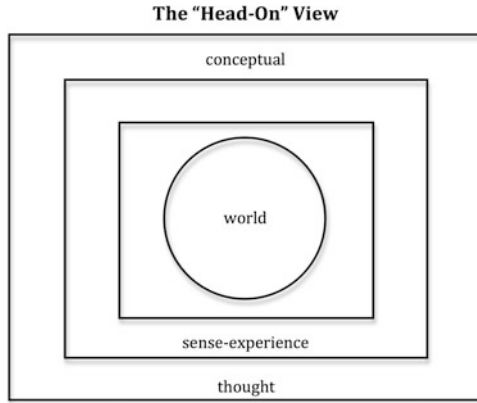
Fig. 7.2 The "Head-On" View

norms, and so forth. In other words, only agents, within a practice, can see other agents.[12]

The upshot of this argument is that we can imagine an extremely complex entity, engaged in very nuanced rule-governed behavior, sensitive to its environment, but which was still not an agent. Imagine, perhaps, a very complicated android of the sort we are as yet unable to build, one that can respond to verbal orders with complicated behavior, can even mimic a conversation, and so forth. Why might such a being not be an agent? It is not an agent if it could not exercise freedom and agency with respect to the rules governing its behavior, if it could not make the rules governing its behavior the object of the type of rational scrutiny agents are capable of. (We can imagine an android that could do such things, but this is just to imagine one that has become an agent.)

A machine can *learn* in a certain sense. Search engines can learn your search preferences, computer programs can learn your music preferences, or to recognize your voice, or your handwriting. But until an entity enters the space of reasons, where the game of giving and asking for reasons occurs, and one displays freedom and agency with respect to the norms governing one's actions, one cannot be regarded as an agent, and one's behavior cannot be regarded as consisting of *actions* in the full sense. As noted above, we doubt that there is a fully non-circular way of *describing* what it is to be in the space of reasons. One cannot give a set of neces-

sary and sufficient *descriptive* conditions for being an agent because such conditions would not capture the action-guiding aspect of agency and action. Again, the sideways-on view will never capture the normativity of agency. Only from within the perspective of agency can agency be seen.

We think that this conclusion is something that, in the end, even the strict naturalist must concede. Consider, David Copp, who is critical of "primitivist" accounts of the normative, which leave normativity as an unexplained primitive. According to Copp, any attempt to reconcile "our belief that there are requirements and prohibitions that are *categorically binding* and that there are corresponding normative properties, such as the property of *being obligated*" with "the scientifically informed view of what exists" (2015, 52) must "explain the nature of normativity in terms of naturalistic phenomena of which we have an independent account" (2015, 62). Copp holds that the only theories that have a decent prospect of accomplishing this are reductionist ones. Like the reduction-minded philosophers, we discussed in Chap. 3, Copp maintains,

> A reductive naturalist can agree with non-naturalists that normative *concepts* are unanalyzable even if she claims to provide naturalistic accounts of the nature of normative *properties*…[T]here seems to be a logical and conceptual gap between any description of a state of affairs in wholly non-normative terms and any normative evaluation of it – as an instance of wrongdoing, for example. For this reason, the most defensible form of normative naturalism is a kind of 'non-analytic naturalism' that rejects the further thesis [that the identity statements in question are analytic or conceptual truths]." (2015, 62–63)

Thus, we should not expect that a *description* of our social practices can be, at the same time, an account of agency and normative constraint. Any description will fall short of capturing the normative dimension of these practices. Thus, while normativity arises within the context of social practices, and social practices are things that are carried out by embodied subjects in a physical world, it is misguided to think that a description of these practices will somehow *show* the normativity of these practices, or show the constraints that are binding on the agents who are involved in these practices. This is not to say that the evaluative is spooky, or to be a primitivist about the evaluative. It is merely to make the point (which the

naturalist agrees with) that evaluating and describing are different activities, and no amount of description—no matter how detailed, or full of counterfactuals—will ever capture the evaluative dimension. This conclusion only leads to supernaturalism or an uncomfortable non-naturalism if you insist on reading this paragraph as making a series of *metaphysical* points about the nature of normativity, rather than as a series of *conceptual* points about the difference between descriptive and evaluative claims. We hope we have made it clear that while our normative practices are essentially world-involving, metaphysical questions regarding norms and normative properties are out of place; we are concerned instead with the function and structure of normative *discourse*.

7.3.2   Empirical Constraints On the Normative

Thus far, we have been offering a social practice theory account of normative discourse, and we see no need in such an account for substantive normative properties. To see such properties as necessary to our moral theorizing is to be overly wedded to a representationalist view of language, and a linguistic monism which takes only discourse which plays the role of positing causal explanations to be legitimate. A persistent worry that might be nagging the reader, though, is that we have thrown the baby out with the bathwater. In jettisoning substantive normative properties from our theory (indeed, in renouncing traditional forms of naturalism altogether), have we severed all connections between the empirical and the normative? Do our social practices spin free in the void, without input from the physical world? We have repeatedly stated that the empirical can constrain the normative, but given how we have elaborated our view—in terms of a social practice account where normative facts need play no role—can we now make good on that claim?

Even if the meanings of normative sentences are not reducible to natural facts, and even if supervenience does not capture what is distinctive about the normative, it seems clear that the empirical can still constrain the normative, and that this form of constraint is compatible with the theory we have offered here. Consider a pair of examples. A person living in the USA in the 1860s might, for example, have thought that slav-

ery was permissible because blacks were somehow less than human: he might have thought they were uneducable and fundamentally intellectually different from whites. However, upon reading Frederick Douglass's beautifully written memoir, *A Narrative of the Life of Frederick Douglass, an American Slave*, and hearing Douglass's eloquent oration, our hypothetical individual might find himself having to reexamine his moral commitments on the basis of this new *factual* information. In epistemology, empirical inputs might induce one to change one's epistemic methods. For example, empirical investigation into the various heuristics that people employ might reveal various systematic cognitive biases (such as certain availability heuristics), and upon learning of this result, one might make a systematic effort to correct for such bias in an effort to make one's judgments more reliable. Prudential judgments display a similar constraint by the empirical: if Smith takes zinc lozenges in an attempt to lessen the severity and duration of his colds, a discovery that studies have shown their ineffectiveness (and thus, that he is spending money and taking bad-tasting lozenges needlessly) can make him realize that his current policy is not the most prudent one. Thus, the various types of normative discourse are all empirically constrained. But saying this in no way commits us to the further claim that any of these types of discourse are somehow reducible to the natural, or that natural facts alone will deliver for us the prescriptivity that is characteristic of the normative. These seem to be examples of where one has a normative theory in place, which tells one how to act in particular circumstances. And when one finds out that circumstances are otherwise, one finds that another course of action is appropriate. So this kind of empirical constraint on the normative is not necessarily about how the empirical constrains the content of normative *theory*, but is more about how the empirical constrains normative judgments and *behavior*, once the normative theory is already in place. Thus, this type of constraint in no way entails that the normative must somehow reduce to the natural.

At this point, several objections can be raised against our account. An objector might argue, "Aren't you just lucky, though, that your normative theories allow empirical inputs from the world? For if you start with a social practice theory of the normative, surely it is a *contingent* matter whether this social practice will include norms regarding language-entry

moves from states of the world to normative statements." This is reminiscent of a familiar objection against coherentist theories of knowledge: If knowledge is a matter of coherence, then surely inputs from the world are not required, and it will be a contingent matter whether any system of knowledge requires observational inputs or does not. And a system of knowledge that does not require observational inputs is obviously inadequate. The standard coherentist response to this objection was that a belief system that did not allow for observational inputs would not even count as a system of empirical knowledge. And so the requirement for observational inputs acts as a sort of a priori condition on a belief system counting as a system of empirical knowledge in the first place. Our response to the similar objection is, if anything, even stronger. We already argued in Chap. 5 that the idea that the normative could spin freely rested on a fundamental misunderstanding of the normative, resting on the twin errors of dualism and intellectualism. We are embodied creatures, and norms must be understood in the first instance as implicit in our world-involving practices. Both dualism and intellectualism take the differences between normative and non-normative discourse as a gap between categories of entities that we must then bridge. (How? With what?) The key here is to recognize this as a fundamental error, rather than sharpening our responses.

But there is even more to say: Because we are embodied, the various normative practices in which we engage simply would not make any sense if they did involve this embodiment at a fundamental level. In moral theory, our embodiment creates vulnerabilities (and corresponding obligations), and a system of norms that did not take these into account would not be a system of *moral* norms at all. For example, if I did not observe that (say) the electrical shocks I was delivering to Smith were painful, and form on the basis of this observation that this pain is a reason not to continue delivering these shocks (and suppose this failure on my part is systematic), then I am not engaging in properly moral thinking *at all*. Similar comments will apply to other sorts of normative reasoning, such as prudential reasoning and epistemic reasoning. Our embodiment and involvement with the world will make the elimination of the empirical from such types of discourse impossible. A person who does not take into account empirical information will be

unable to satisfy even his most basic prudential interests (safety, nutrition, etc.). And because of our fundamental engagement with the world, a body of "knowledge" that did not accept inputs from the world, and hence was cut off from (and hence involved no information about) the world in which we are embodied would be so badly impoverished as to be unthinkable. It is doubtful whether such a system of beliefs is even possible for embodied creatures such as ourselves. But at the end of the day, our fundamental embodiment and engagement with the world means that the fact that our epistemic practice allows for empirical inputs is not a matter of luck—it is a fundamental feature of who we are and how the social practice functions.[13]

A second objection is this: do we really revise our normative beliefs or theories based on input from the world? Does the world really have this kind of evidential bearing on our normative theories? This objection stems from a certain conception of normative discourse, which has received its most lucid presentation in Harman (1977, Chap. 1–2). On this conception of normative discourse, the best explanation for any normative judgment we make is simply the psychological set of the person making the judgment. On this view, there are no normative facts in the world, and hence observation can never warrant revision of a normative theory (since, on Harman's view, a theory can only be mistaken, and hence in need of revision, if it somehow fails to conform to facts which exist independently of human belief systems).

We have already rejected the idea that there are substantive normative properties, or that these could play any useful role on our normative theorizing. Nevertheless, there are at least two different ways that the normative is constrained by the empirical, one of which is easy to establish, and one of which is more difficult. Consider the easy kind of constraint, which we have already discussed to some extent above. Suppose Smith believes that the death penalty is morally permissible, in part, because it is an effective deterrent; or that same-sex couples should not be allowed to adopt children because the children will suffer worse psychological outcomes (due to not having both a father and a mother, or due to teasing about their non-conventional family situation, etc.). In such cases, empirical evidence that (say) the death penalty is not an effective deterrent, or that children raised by same-sex couples have similar or better

psychological outcomes when compared to similarly situated children raised by opposite-sex couples would warrant a change in moral position. Similarly, a belief that a data set warrants belief in a conclusion, P, would have to be revised if it turned out the data was flawed, or systematically biased, or the result of fraud, and so on. We take it that it is uncontroversial that the empirical constrains the normative in this way. Agents may *refuse* to do so of course, and there is a growing (and philosophically dispiriting) body of empirical evidence that agents are especially recalcitrant in revising their beliefs on urgent matters such as morality and politics, even growing worse when corrective evidence is presented.[14] There is no denying this fact, unfortunately, but like many of the suspect heuristics and habits, we mentioned above, these are matters that will not receive a defense at the level of theory, and which cannot be explicitly defended.

But there is a more controversial way in which the empirical might constrain the normative. Might we make novel normative empirical observations, which can warrant theory change? Call this the "hard case" for empirical constraint on the normative. Let us consider an example, which is due to Richard Werner (1983). We are to imagine "that Fred has carefully read and considered the utilitarian literature and finds that utilitarianism conforms completely with his moral sensibilities and psychological set" (1983, 657). Fred has carefully considered the objections against utilitarianism, including the objection that utilitarianism would permit slavery under certain circumstances. Fred has concluded that slavery is almost always immoral, but would be permitted under a certain narrow range of circumstances. But then Fred watches the miniseries *Roots*, with its brutal depiction of the institution of slavery, and abandons utilitarianism, reasoning as follows: "The fact that slavery is obviously morally wrong and that it follows from my version of utilitarianism, shows that my version of utilitarianism is disconfirmed as a legitimate moral theory" (1983, 658–659). On Werner's telling of the example, Fred observed the *cruelty*, the *immorality*, of slavery, and this warranted a rejection of the moral theory he believed to be true.[15] On the epistemic side, a medical researcher might find evidence of systematic *bias* in unblinded studies, and realize that such studies are not justification-conferring. In both

cases, observation—input from the world—is warranting a change in one's normative theory or belief.

Several comments are necessary about the hard case. First, the hard case often suggests a model of observation which is different from the overtly inferential model suggested by the easy case. In the easy case, one observes some factual matter, and then infers some normative conclusion (say, that same-sex adoption is permissible). In the hard case, one's observation itself seems to have normative import—one observes that slavery is *cruel*, or that unblinded studies are *biased* (where the term "bias" has a clear normative flavor, and is not merely a description of some kind of statistical variance). Second, though moral theorists like Werner want to understand moral perception as an example of perceiving substantive moral facts or properties in the world, there is no need to understand the hard case in this way.

If we do not understand the hard case in terms of perception of moral properties in the world, then how should we understand it? A helpful route to understanding the hard case is by looking at McDowell's discussion of Blackburn's projectivism. On McDowell's account, Blackburn sees that the metaethicist has two options. On the one hand, she can "expand reality…to include an extra population of distinctively value-involving states of affairs or facts" (1988, 3), along with a special faculty of intuition with which to perceive these states of affairs or facts. On the other hand, she can deny that there are such states of affairs or facts, but the alternative then is to endorse some version of Humean projectivism: "According to Hume, when our 'taste' is projected on to the world, it 'raises in a manner a new creation'" (1988, 2). Thus, although there are, strictly speaking, no moral facts or states of affairs, we believe that there are as we project our moral sentiments onto the world. In saying this, Blackburn is not endorsing an error theory, though:

> Blackburn's proposal, in effect, is that this 'new creation' can be sufficiently robust to underwrite the presence of the trappings of realism, so to speak, in thought and speech which is correctly understood as projective; and that participants in such thought and speech need not be led by those elements of it into missing its projective nature. (1988, 2)

McDowell holds that these two alternatives—robust realism vs. projectivism—present us with a false dichotomy, and illustrates his point with the example of "funny," which is plausibly a concept which is at least as projective as any ethical concept. A robust realist account of "funny"—in which there are comic facts or states of affairs, existing independently of any human psychological propensity to find things funny—is completely implausible. But is a projectivist account of "funny" plausible? Can we say, for example, that "funny" is whatever makes us laugh? Surely not; for many things besides the comic (including embarrassment) can make us laugh. McDowell's suggestion is that mere exercises of our disposition to laugh are not enough to fix the extension of the concept "funny." Rather, the concept must itself already be at play to distinguish which of these exercises count as amusement (rather than, say, embarrassment). Thus, according to McDowell, we can say that neither is prior to the other: the concept of things that are funny, and the disposition to laugh at funny things. When we carry this lesson to the moral realm, we see that Blackburn's understanding of our options is too limited: "[The realist] holds that the moral features of things are the parents of our sentiments, whereas the Humean holds that they are their children" (1981, 165). But this is wrong; neither is the parent nor the child of the other, according to McDowell:

> Denying that the extra features are prior to the relevant sentiments, such a view distances itself from the idea that they belong, mysteriously, in a reality that is wholly independent of our subjectivity and set over against it. It does not follow that the sentiments have a priority. If there is no comprehending the right sentiments independently of the concepts of the relevant extra features, a no-priority view is surely indicated. (1988, 7)

McDowell even suggests ways in which the inculcation of moral sentiments is related to the promotion of our interests, which is relevant to the present account:

> No doubt reflections about the benefits of co-operation and social order go some distance towards 'placing' ethics—making it intelligible that we inculcate ethical sensibilities in our young, trying to give ethics the impor-

tance to them that we believe is proper. But we do not need to suppose that such 'placing' functions by allowing us to make sense of a range of subjective responses to a world that contains nothing valuable…What we 'place' need not be the sort of sentiments that can be regarded as parents of apparent features: it may be pairs of sentiments and features reciprocally related—siblings rather than parents and children. (1988, 12)

McDowell gives us the tools to understand the hard case. We have already emphasized the way in which we are embodied, and in which our normative practices involve a fundamental engagement with the world. However, this does not involve reading pre-existing normative properties off of the world, properties which exist independently of our interests and practices. On the contrary, our interests and practices do not exist independently of our engagement with the world; and for this reason (and in light of McDowell's argument above), it is not plausible to endorse a projectivist account of normative properties. But McDowell's argument should help us see that we do not need to embrace an account of our engagement with the world that involves normative properties in the first place. For on McDowell's account, even though there are not normative properties in the world, the projectivist account (whereby non-normative states of affairs elicit in us certain responses, which we "project" onto the world) is inadequate as well. Instead, on a McDowellian version of our account, we encounter a moral world, but a world that is not morally independent of our particular interests and projects. But again, there is no order of priority—we have the interests we do in part because of how the world is structured and because of how we are engaged with the world, and the world has the moral character it does because of our interests, and the way we engage the world to promote our interests. Thus, when we see a person on television *cruelly* abusing a slave and realize that no moral theory which countenances slavery could be true, this is a genuine instance of the hard case: we are encountering the world morally, but not in a way that presupposes a moral reality independent of our interests and our engagement with the world. When we realize that the patients' and researcher's expectations and beliefs *bias* the outcome of a study if they know who receives the placebo and who receives the medication, we are making an epistemic encounter with the world. But again, this is

not an encounter that presupposes that there is some mysterious realm of epistemic facts independent of our interests in understanding the world, and our practical engagement with the world pursuant to this interest.

To emphasize a point we argued in Chap. 3, and revisited earlier in this chapter, to say this is not to countenance the existence of substantive moral or epistemic properties, even of a socially constituted variety. As we argued in Chap. 3 (and reiterated in Sect. 7.1 above), it seems wholly implausible that cruelty will constitute a recognizable property at the non-normative level. One can observe of an action, or an institution, that it is contrary to various specific interests, and judge that the act or institution is cruel. But the action-guiding element of such an ascription is always going to be more important than the representational element. Although characterizing the act as cruel (rather than dishonest, or some other normative distinction) will specify in which *ways* the act is contrary to our interests (and thus categorize it to a degree, as a descriptive distinction would), the chief import of this ascription is that the action is contrary to the long-term exercise of our interests, and ought not have been done.

With this in mind, we can make sense of Werner's example. When Fred sees *Roots* and realizes that utilitarianism cannot possibly be a correct moral theory because it allows for the barbarous institution of slavery (under very narrow circumstances), it is implausible (given what we have said above) that Fred has perceived some independently existing moral reality, and concluded on this basis that utilitarianism must be false. Nor is he simply seeing a set of natural facts, and projecting onto them some kind of moral life, or inferring from them moral premises. Rather, he is seeing an action as morally relevant in a kind of way that is informed by his moral practice and his concrete engagement with the world. He can see that the institution of slavery is irreconcilable with a certain kind of normative commitment; that it is fundamentally opposed to general human interests in a way that cannot be reconciled within acceptable moral theory. This does not involve the perception of any one property, but involves a skillful ability to synthesize a very complex situation involving many factors—pain, humiliation, sorrow, lack of freedom, and forced separation from loved ones—and see in these features of the situation a moral crime incompatible with any decent moral theory.

Not every instance of the hard case has to be so dramatic. When a scientist sees *bias* in a particular research method (e.g., non-blinded studies of the effectiveness of medical treatments), and concludes that this method is not justification-conferring, we see a case that is parallel to the moral case. "Bias" is a normative term. For there are many different ways in which data can be arrayed, and such features of a data set or statistical analysis such as "representative sample" or "low *p*-value" do not have any inherent relevance, apart from our interest in acquiring understanding of the world. They play roles in sorting mere correlation from causation in a survey of data, discovering which medication works more effectively, and so on, all of which goals are related to, and serve, further interests. "Bias" implies that the data, or method of gathering data, are skewed in a way that is inappropriate to the interests which this data collection is intended to serve. Thus, the fact that we see the world as imbued with certain epistemic values is not independent of our particular interests. But we do see the world as so-imbued, and the epistemic observations we make can alter our epistemic beliefs (e.g., cause us to decide that non-blinded studies are not justification-conferring) or theories. But again, to say this is not to force us to be realists about epistemic properties and facts, any more than the moral case forced us to make this concession about moral properties and facts; and for the same reasons. We can see that a particular practice is not conducive to our epistemic interests, and form the proper conclusion on this basis. This in no way commits us to the claim that those practices which *do* serve our epistemic interests form any interesting kind which is not hopelessly gerrymandered, or which serves any useful explanatory purpose beyond what each individual case serves on its own. Thus, the case is parallel in relevant respects to the moral case.

A final objection against our conception of how observation and the empirical can constrain the normative goes in roughly the opposite direction from the last two objections. This objection suggests that contrary to our account, observational inputs serve as foundational inputs to normative theory. This, in turn, suggests an independent normative reality which we are cognizing when we make normative observations; and this picture of normative observation is fundamentally out of step with the picture of normative discourse we have been drawing in this work.

Not surprisingly, we find it implausible that observational inputs could be epistemologically foundational in normative theory (or in any other theory, for that matter). The chief reason for denying that observational inputs can be foundational is that far from being certain, or incorrigible, or having any of the other properties that foundational beliefs are supposed to exemplify, normative observational reports are always inherently defeasible, always fallible. A normative observation, such as "That experiment was justification-conferring," or "That action was unjust," or "That action was imprudent," is always fallible. Even if we grant, for the sake of argument, that traditional foundationalist sentences like, "I am in pain, now", or "I am being appeared to redly," are incorrigible, the same is clearly not true of the normative sentences cited. One might see a situation (say, store owners being compelled to serve members of all races equally) and non-inferentially form the belief, "That is unjust" (i.e., to the store owners, perhaps as a violation of their property rights). But over time, with more experience, and perhaps under the influence of moral argumentation, one will realize that one's earlier non-inferential judgment was in fact mistaken. Thus, it is implausible to consider these non-inferential reports as having the status of foundational beliefs or reports. This in itself does not prove that there is no independently existing normative reality. But to the extent that the argument for such a reality depended on the further claim that normative observation had some foundational status, this argument is undermined.

Notes

1. Koons (2000) argues in much greater detail that even though normative discourse (in particular morality and epistemology) fail Harman's explanatory requirement, such discourse is justified on pragmatic grounds because its linguistic and social role is not primarily fact-stating.
2. This assumption is controversial (as in the case of numerals and numbers that they apparently designate, for example), but again, we are willing to concede this claim, since it will turn out that little important for our account hangs on the question of whether there are moral facts. We would also adopt some more heterodox views on reference in general, but again, these would have little bearing on the present discussion.

3. Terms have (to use Dummett's (1993) terminology) criteria and consequences of application. The *criteria* of application for a sentence are those conditions which license utterance of that sentence. For example, both "A" and "B" are criteria of application for the sentence "A v B." Once we license utterance of a sentence, we become entitled to that sentence's *consequences* of application. For example, "A v B" has, among its consequences of application, "¬(¬A & ¬B)." The criteria of application for a moral term (such as "cruel") will in general be a mix of natural and normative facts (the fact that it caused X pain, the fact that X had done nothing to deserve this pain, etc.).

4. Timmons capitalizes words like "FACT," "PROPERTY," "FEATURE," and so on, when he wishes to make clear that such words purport to refer to mind-independent entities as these entities are supposed to play an explanatory role in our moral theory.

5. This suggestion is similar to Sellars's suggestion of the notion of tied behavior, which he discusses in his (1949), an article that seems to have been influenced by the work of Wittgenstein. We will return later to a discussion of tied behavior.

6. The community in question need not be the entire community, but may only be a subset.

7. This interpretation of second-level utterances is one of the novel features of the account given by Lance and Hawthorne (see (1997), especially Chaps. 1 and 3). Most thinkers regard the purpose of such utterances as descriptive. Brandom (1994, Chap. 1), for example, thinks that such utterances serve to codify or make explicit features of the underlying social practice.

8. See Wolf (2012) for more on this response to relativism.

9. We could play the moral language game in a historically reified and relativistic manner; but there are good reasons not to do so. For a more complete discussion of these issues, see (Koons 2003).

10. Emphasis removed from original.

11. We created these two images, but they are inspired by similar ones from Roderick Long.

12. This is not to say that we could not recognize as agents other beings whose practices we did not understand (for example, radically alien beings or cultures we did not understand). But until we understood their practices, their agency would always have the status, as it were, of a falsifiable hypothesis (which is not to say that it is something we explicitly inferred).

13. One can probably make an even more robust argument for the impossibility of a belief system without empirical inputs. Norms are essentially social;

to be a believer, a knower, is essentially to be involved in a social practice. But this means interacting with other embodied beings, and forming beliefs (*empirical* beliefs) about them and their activities and utterances.

14. See, for example, (Kahan et al. 2013).
15. It might be objected that Fred only observes a fictional representation or recreation of these properties in Werner's example, but we will assume the example could be adjusted to include in-person experience of oppression if readers found that troublesome.

References

Blackburn, Simon. 1993. Moral Realism. In *Essays in Quasi-Realism*, 111–129. Oxford: Oxford University Press.

Brandom, Robert. 1994. *Making It Explicit*. Cambridge, MA: Harvard University Press.

Brink, David O. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.

Copp, David. 2001. Realist-Expressivism: A Neglected Option for Moral Realism. *Social Philosophy and Policy* 18(2): 1–43.

Dummett, Michael. 1993. *Frege: Philosophy of Language*, 2nd ed. Cambridge, MA: Harvard University Press.

Kahan, Dan, Ellen Peters, Erica Cantrell Dawson, and Paul Slovic. 2013. Motivated Numeracy and Enlightened Self-Government. *Yale Law School.* Public Law Working Paper No. 307. September. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2319992.

Koons, Jeremy Randel. 2000. Do Normative Facts Need to Explain? *Pacific Philosophical Quarterly* 81(3): 246–272.

Koons, Jeremy Randel. 2003. Consensus and Excellence of Reasons. *Journal of Philosophical Research* 28: 83–103.

Kripke, Saul. 1982. *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press.

Lance, Mark, and John Hawthorne. 1997. *The Grammar of Meaning*. Cambridge: Cambridge University Press.

McDowell, John. 1984. Wittgenstein on Following a Rule. *Synthese* 58: 325–363.

Quine, W.V.O. 1953. Two Dogmas of Empiricism. In *From a Logical Point of View*, 20–46. Cambridge, MA: Harvard University Press.

Sellars, Wilfrid. 1957. Counterfactuals, Dispositions, and the Causal Modalities. In *Minnesota Studies in the Philosophy of Science*, vol. II, ed. H. Feigl, M. Scriven, and G. Maxwell, 225–308. Minneapolis: University of Minnesota Press.

Timmons, Mark. 1999. *Morality Without Foundations: A Defense of Ethical Contextualism*. Oxford: Oxford University Press.

Werner, Richard. 1983. Ethical Realism. *Ethics* 93(4): 653–679.

Wolf, Michael P. 2012. Boundaries, Reasons and Relativism. *Journal of Philosophical Research* 37: 205–220.

8

Unity without Uniformity: Cross-Discourse Contribution

In the preceding chapters, we have staked out an anti-reductionist account of normative discourse, albeit a less-than-conventional one. While we eschewed ontological commitments (not just to normative entities, but even to substantive normative properties), we defended an account of action-guiding content on which normative sentences could be true or false. *Ontological* eliminativism about the normative does not entail *doctrinal* eliminativism, as we said in the early going. This is an anti-reductionist account in the sense that what we say "at the normative level" will not be explained (and potentially displaced) by a more "fundamental" vocabulary such as psychology or physics, nor will there be entities and properties "at the normative level" identified with entities and substantive properties at a more "fundamental" one (trivially in our case, since we are not positing normative entities or properties at all).

While we endorse a form of anti-reductionism, we also recognize that it is a tricky position to occupy, and that its defense will have to be subtle and canny. The ramifications of our position on normative discourse and its relation to other theories strikes us as crucial because some approaches would insulate us from proper philosophical concerns about intertheoretic relations. As we said in Chap. 3, we take it that strong forms of

discourse pluralism are untenable, including strong discourse pluralism about normative discourse. While discourse regions can be distinguished for various metatheoretical purposes, the commitments in each one must ultimately be reconciled with the others. In the final estimate, there cannot be non-overlapping magisteria insulated from one another's reach. We are still pluralists in the sense that discourse regions will not be linked by an underlying uniformity that explains them all. But we will argue that there are metatheoretical relations between different regions that both help secure the objectivity of the normative and incorporate it securely within a larger account of the natural world. In this chapter and the next, we will demonstrate the possibility of a unity among discourse regions that does not impose a uniformity in their contents.

8.1   Varieties of Anti-Reductionism

Much of the anti-reductionist literature of the last 40 years, both pro and con, can be traced back to debates about the ontology of mental states and the autonomy of psychology from physics. In particular, Fodor's "Special Sciences" (1974) set the agenda for much subsequent discussion. Taking the positivists' views as a starting point, successful reduction would be the replacement of all predicates in the special sciences (i.e., every theory outside the physical sciences) with predicates from physics to thereby make the laws of the special sciences restricted cases of fully general physical laws. It is a matter of debate whether the special sciences have laws at all (and whether they are deficient if they do not), but the broader themes of Fodor's anti-reductionism do not turn on this. Even if we agree that psychology, biology, and other theories outside of physics proceed by carving their fields of study into kinds and searching for law-like regularities, the theoretically salient regularities that we find in those fields generally do not turn out to involve interesting physical similarities. Borrowing an example from economics, Fodor notes that something like Gresham's Law might be true even of wildly dissimilar physical events and objects: the currency involved might be pieces of paper, metal, or some electronic record of accounts; the mechanisms that overvalue or undervalue currency can involve untold numbers of agents, physical objects, and so on.

But such patterns need not violate more modest forms of physicalism; each token instance of a special science kind will be token-identical with some physical object or set of them.

Fodor's argument has been central to subsequent debates because it captures a particularly important intuition particularly well—the idea that patterns at different "levels" might be just as real and just as theoretically informative without aligning with patterns at whatever the level of their components might be. But Fodor's account is also steeped in the concerns and assumptions of computational theories of mind and the functionalist tradition that traces its way back to Turing, Putnam, and Armstrong. Functional properties exhibit the broadest possible substrate neutrality, often to their considerable explanatory advantage; you can define functions over anything there can be. If so, we could study functional properties with little or no attention to the physical inventory that realizes them, even while the inquiry itself remains an empirical one. Proponents often speak of the autonomy of the special sciences from the physical ones, and in many accounts, this permits an *insulation*, even an *isolation* from the physical sciences. One can speak or do research at one level while setting aside most (or all) concerns for its relation to others. This exemplifies one kind of anti-reductionist strategy: establishing different kinds of information (or patterns) and segregating them from one another to preserve what had been the reduction target. Much work in the philosophy of mind and cognitive science in the last three decades has pushed back against this segregating approach, arguing that some sort of attention to the physical (or neurological, or biological, etc.) details is necessary to any theory of the mental.

Another historical track of discussion runs through the philosophy of biology. The advance of mechanistic and later molecular explanatory models for many biological phenomena put the question of reductionism to biologists and philosophers of biology in the nineteenth and early twentieth century. Some even adopted vitalism as an anti-reductionist strategy. Vitalism has been soundly rejected, but in the wake of those earlier debates, numerous strains of anti-reductionism emerged, focused in some cases on ontological issues and in others on explanatory ones. A notable figure on the explanatory branch of the anti-reductionist literature is Philip Kitcher, particularly his work on classical and molecular

genetics in "1953 and All That: A Tale of Two Sciences"(1984). There, he argued for the failure of efforts to reduce classical genetics to molecular genetics on several grounds. In part, this was driven by doubts that classical genetics and other parts of biology could be correctly characterized with a syntactic covering-law model common to most reductionist accounts. Classical reductionist accounts also presumed that there would be means by which the vocabulary of a reduced theory could be systematically mapped onto the vocabulary of the target theory (e.g., the vocabulary of classical genetics could be mapped onto some portion of molecular genetics), and that a derivation of general principles in the reduced theory (classical) from the target theory (molecular) would explain why those in the reduced theory hold (1984, 339). Kitcher denies all three of these reductionist tenets—laws, mapping, and derivations. In some respects, his anti-reductionist account mirrors the features of the "autonomy" described earlier by Fodor:

> [A]nti-reductionism emerges as the thesis that there are autonomous levels of biological explanation… Explanatory patterns that deploy the concepts of cytology will endure in our science because we would forswear significant unification (or fail to employ the relevant laws, or fail to identify the causally relevant properties) by attempting to derive the conclusions to which they are applied using the vocabulary and reasoning patterns of molecular biology. *But the autonomy thesis is only the beginning of anti-reductionism.* (1984, 370–371, emphasis added)

Why not stop with some form of autonomy thesis? Left alone, it might imply the sort of segregation we described earlier. This will not suit many theoretical projects in biology that do make substantial use of physics or other "lower level" theories. But these intertheoretical relations are more subtle than reductionism would suggest. Instead, Kitcher suggests construing classical genetics (and other unreduced theories) as incorporating patterns of reasoning that underlie a core set of solutions to the theory's "pedigree problems" (1984, 353–354, 355–358). In some cases, molecular genetics provides an "explanatory extension" for classical genetics that resembles reductionism. That is, some set of molecular mechanisms are sufficient (or close) to explain some feature

of classical genetics, and once they do, that feature is thereby derivable from parts of lower-level theories in the physical sciences. And this constitutes a success. What had been a brute fact in want of an explanation now has one that unifies it with other more general explanatory accounts. But other features of classical genetics will not derive so neatly from the molecular, and may instead draw support from other theories, including even less "fundamental" levels of explanation within biology itself. Kitcher mentions limits on the expression of certain genes that are matters of the geometrical structure and relative distances between cells found in them; in such cases, no *purely* molecular derivation can serve as a sufficient explanation, and we actually find the direction of explanation going "downwards" from morphological descriptions of cells to their molecular components in the expression of a phenotype (1984, 372). Explanatory accounts in classical genetics (and other parts of biology) will thus involve regular shifts from higher levels to lower levels, and vice versa, and different features of higher levels of explanation will call upon various different explanatory accounts to address their problems.

> Anti-reductionists are not only able to contend that there are autonomous levels of biological explanation. They can also resist the weaker reductionist view that explanation always flows from the molecular level up… Because developmental processes are complex and because changes in the timing of embryological events may produce a cascade of effects at several different levels, one sometimes uses descriptions at higher levels to explain what goes on at a more fundamental level. (1984, 371)

This account would preserve some central features of Fodor's autonomy thesis, with higher levels of theoretical description modeling complex patterns that cannot be captured solely in the laws and vocabulary of lower levels of theoretical description. But in Kitcher's case, this does not entail a segregation of the higher levels from the lower ones. On the contrary, the various levels in play in this account are richly interwoven with one another in many ways, and this actually *strengthens* the various explanatory accounts. While the higher levels as wholes will have theoretical autonomy, their various features will not be brute facts and

stipulations, but instead will be integrated into other accounts in fruitful ways. That integration may be bottom-up, top-down, or perhaps in some lateral configuration where one higher-level account aids another higher-level account without the hierarchical ordering that reductionism presumes.

We take it that Kitcher's view also implies a rich local network of practical and theoretical supports for each higher-level account. That is, the ways in which something like classical genetics gets resources for explanatory extensions from other parts of biology, the physical sciences, mathematics, and others will be numerous, multifarious, specific to the problems at hand, and will demand an elaboration with attention to these local details as the account matures. Moreover, the combination that each higher-level account deploys will be unique to that account, rather than implying a general structure that all such intertheoretic relations must exhibit. How biological accounts make use of chemistry, physics, and so on, will differ from the ways that psychological accounts make us of chemistry, physics, and so on, and this is in no way a mark against those accounts. Their integration with other theoretical accounts—their unification, as Kitcher has called it elsewhere—is in itself a mark of their strength, even if we do not find any general, systematic way in which all such integration occurs.

So go these two sketches of anti-reductionist strategies. We should emphasize here that while we are interested in anti-reductionist approaches for our own purposes, we do not necessarily assert that either of these accounts are correct for their respective fields. Presumably, these are empirical hypotheses about the future course of these forms of theoretical discourse based on their pasts, and any such hypothesis should be open to challenge and refutation. We are sympathetic to anti-reductionism in both fields, but our present purpose is to take these accounts as illustrations of approaches and issues that might be at stake in countering reductionist challenges. Nor are we adopting or endorsing the details of other views such as Kitcher's. Rather, we are using them to illustrate themes in a response to reductionism, and how we might want any such response to take shape. (And even someone like Fodor sometimes sounds an integrationist note or two, as when he says that a better articulation of the impulse behind reductionism is that we should "explicate the physical

mechanisms whereby events conform to the laws of the special sciences"
(1974, 107).)

That is not to deny some forms of segregation, at least between pairs of
theories. We should probably think of mathematics as exhibiting a sort of
one-wat segregation from discourse about the physical, for instance, even
if physical details serve as illustrations and counterexamples, and many
parts of mathematics will be developed with the problems of modeling
specific physical phenomena in mind.[1] Physical discourse does not tell us
something about mathematical discourse in the ways that the molecular
accounts told us something important to genetic ones in Kitcher's exam-
ples. But even when such segregation is appropriate, it will be appropriate
with respect to some subset of other discourse regions, not all (e.g., mak-
ing claims in mathematics will not oblige us to *answer* to or be *constrained*
by physics as other sciences are). Any given discourse region will still have
deep intertheoretical relations with many others, as we argued in Chap.
4. Segregation should and will be local and somewhat exceptional in this
sense, and the development of different forms of discourse should be
open to the insights and constraints of others by default.

How can the lessons drawn from these non-reductive strategies be
ported over to the case of normative discourse? (We should be explicit that
our goal here is to draw a number of broad themes from these accounts,
not to defend any of them to the letter.) Given our worries about what
amounts to segregation in Chap. 4, we have good reason to seek some
form of integration for normative discourse with at least some (maybe
all) forms of non-normative discourse. This would also aid in the devel-
opment of a more promising form of naturalism, one that showed how
normative discourse arose out of an engagement with the world, rather
than appealing to something outside it. Part of what animates the inte-
grationist approach that Kitcher offers is a broadening and enrichment of
the notion of explanation. Unification is a more flexible, inclusive notion
than the strictly syntactic conditions in deductive-nomological models,
and more attention is paid to patterns of reasoning local to specific theo-
ries and their pedigree problems. But these are all still forms of *causal*
explanation, firmly ensconced in discourse regions whose primary roles
are descriptive and for which we have to make some array of ontological
commitments. We have been rejecting such readings of normative dis-

course, as they either descend into supernaturalism, queerness, or reductionism that does not address the normative surplus. So making a case for normative discourse on integrationist terms will require making a case for a broader set of legitimizing intertheoretical relations.

To approach this task, we should begin by noting some general features of the integrationist theme sketched earlier. Relations can sometimes hold between whole theories. Everything we say in biology will be constrained by all the laws of physics, even if many parts of physics are not pertinent when we do various parts of biology (e.g., cytologists do not generally worry about rates of cosmic expansion). Those sorts of whole-to-whole relations will generally be thin and less interesting, though. More informative insights will come from looking at how specific parts of one discourse can explain or otherwise assist parts of another. For instance, cognitive scientists working on models to explain visual perception will have call to appeal to accounts of optics within physics, physiological accounts of the muscular contractions that permit focusing, biochemical and neurological accounts of the various photosensitive cells and means by which they transfer information to other cells, and various parts of organic and physical chemistry that underwrite the biochemical and neurochemical accounts. But work in recent decades suggests that there will also be a place for accounts such as dynamical systems theory that cut across many domains, rather than solely to further hierarchically ordered levels of physical structures.

Note two features of the ways in which these different accounts assist someone working on vision. First, the relevance of some supporting account—like optics in physics or muscular structure in physiology—will depend heavily on the sort of question at hand in the account of vision. Appeals to optics are invaluable in explaining why visual organs exhibit structures with lenses, foci, and muscular systems that contract to change those optical features. But the features that make optics relevant here drop out when we then switch to questions about pattern recognition, or signaling through the optic nerve to the primary visual cortex where the mechanisms of information transfer are pertinent. In this way, the manner in which higher-level theoretical accounts integrate with lower-level ones will be *local to* the tasks at hand. How physics supports the explanatory work done in, say, cognitive psychology will be a unique

combination of borrowed models, laws, and other patterns of reasoning that serve as partial components of explanatory strategies like those mentioned above. Other theories will borrow from physics in this sense, but what they borrow, how they do so, and what explanatory strategies they pursue will differ from theory to theory. And the same will hold for integration in which lower-level properties are explained by higher-level ones, as in Kitcher's examples of classical genetics, where some genetic properties are explained by cytological properties.

Second, (as is somewhat explicit in this first point) there will be numerous types of explanatory assistance rendered from one theory to another. For some problems, our concern will be the composition of the objects of study: what are the cells in the eye, optic nerve, and so on *made of*, at the chemical level? In others, we need an account of causal mechanisms, for which the details of physical composition alone would not suffice (though in which they may figure). In others, those mechanisms will be woven into still larger models with other sorts of explanatory goals (e.g., the role of visual ability in enhancing an organism's reproductive fitness, explaining changes in its frequency in a population). We should also note that the roles played by theories such as mathematics, which seem fundamental to explanatory projects in physics, but not as ontological components in the classic reductionist molds. (We will elaborate a bit more on this contrast in the next section of this chapter.)

In our view, this diversity and locality of intertheoretical relations is no embarrassment to our larger metatheoretical understanding of our inquiry. To expect (or demand) uniformity across discourse regions is an ambitious mistake at best and dogmatism at worst. Instead, we should embrace this diversity. Doing so will require a broader metatheoretical notion to denote the array of different positive, constructive ways in which discourses can relate to one another. We suggest that the *contribution* that one type of discourse can make to another, suitably articulated as a metatheoretical expression, can serve this role. One discourse region may contribute to the projects and interests pursued in the exercise of another, even if neither of those discourse regions reduces to the other. We contend that such contributions are ubiquitous across different discourse regions, and indeed a paucity or even the absence of such relations would render a discourse region suspect. We will articulate a notion of discursive

contribution that accommodates the diversity we have discussed, both in the types of discourse in which we can engage and the relations of support that lie between them. This will also permit us to articulate the ways in which normative discourse can be integrated into a larger whole with non-normative discourse, with various discourse regions contributing to the normative and the normative contributing in various ways to others. Much of this was foreshadowed in Chaps. 4 through 6, and we will offer some concrete elaborations of this strategy in Chap. 9.

8.2   General Characteristics of Cross-Discourse Contribution

We can begin our introduction of the notion of discursive contribution by recalling the notion of a discourse region introduced in Sect. 4.2. There, we suggested that there were no fissures between regions of our languages that supported strong discourse pluralism (the view that regions of the language were isolated from other portions by their own internal structures), but that there were good reasons to adopt moderate discourse pluralism (the view that we could note coherent substructures within regions of the total discursive topography of our languages, depending on the metatheoretical interests at work when we draw such distinctions). Thus, we can say that there really are deep and metatheoretically significant differences between the theoretical discourses of physics, ethics, mathematics, and so on, and the differences in their structures reflect the interests served by those forms of inquiry. Each such region has a character of its own that distinguishes it, suggesting what it must achieve and why we pursue it. But for other purposes, we may cut the topography differently. For instance, much of physics, biology, and a host of other discourse regions could be characterized as descriptive, in contrast with much of regions like normative ethics. For still other metatheoretical purposes, we could distinguish subregions of these regions—cosmology versus quantum mechanics within physics, or cytology versus genetics within biology. But we suggested there that portions of different discourse regions overlapped one another in various ways, and that treating any region as wholly autonomous or isolated from all others either rendered strong

pluralism incoherent or simply recapitulated the problems it purported to resolve. Any way we slice it, different parts of our theoretical vocabulary *need* one another, and the commitments (ontological and otherwise) made in any one of them must ultimately be squared with those made elsewhere.

Where the substructures and achievements of one discourse region serve some need for another, we shall say that one region *contributes* to another. In the strongest of cases, one region may entirely replace another, as would be the case for successful theoretical reductions in the classical mold. But more often, contribution will be a matter of parts of a discourse region serving a partial role in the exercise of another. Paradigmatically, parts of mathematics contribute to the discourse of physics, such as geometrical terms in the articulation of models of a space-time manifold, or numerical terms for quantities of various properties, or in the expression of relations in physical laws themselves. Parts of chemistry contribute to biology, such as when molecular structures figure in explanations of processes such as metabolism or reproduction. These examples will be familiar from discussions of reductionism, but many others are possible.

As a first pass, let us say that where T_1 is some region of theoretical discourse, characterized by some set of interests, it contributes to another region of discourse, T_2, iff some of the elements of T_1 occur in T_2 in ways conducive to the interests pursued in T_2 and do so in ways that cannot be eliminated or replaced. This proposal is very open-ended, obviously, but this is by design. The aim here is to capture a wide variety of substantial practical differences that one discourse region can make to another, rather than a narrow logical, metaphysical, or semantic relation. (For the moment, we are also presuming the interests in question are sufficiently narrowly conceived to distinguish the discourse regions. T_1 and T_2 may both "contribute to our knowledge" or "help us get along in the world," but presumably they may characterized in terms of more specific interests that they do not share if we sustain them as separate sets of practices at all.) Greater insight into the integration between different discourse regions will come from closer examination of the particular forms of assistance that regions make to one another, and these will vary according to the interests pursued by the regions

in question. Degrees of resemblance will hold between different types of contribution, and there will be metatheoretical conclusions to draw from comparisons, but attention will be due to unique features of each instance's character. In these ways, our account of contribution may resemble a loose botany of types more than the rigid hierarchies of many reductionist and supervenience accounts.

Looking at these possibilities in a very broad fashion, note that when contribution occurs between a pair of discourse regions, it may exhibit a number of general forms. Where some contribution relation, R, holds between regions T_1 and T_2, that relation may be symmetrical or asymmetrical. We may say that a contribution relation is symmetrical iff there is a role played by some subset Φ of the elements of T_1 that contributes to the theoretical interests pursued with T_2, and there is some subset Γ of the elements of T_2 that contribute to the same interests pursued with T_1. In short, some of T_1 does for T_2 just what some of T_2 does for T_1. If that sounds like it would be exceedingly odd and rare, we agree. Why borrow from another theory to do what a theory can already do? Different discourse regions are rarely, if ever, devoted to the same interests, and in most of the rare instances in which they are, they would presumably be rivals rather than complements (e.g., caloric theory versus mechanical theories of heat). There is a trivial case in which symmetrical contribution holds: every discourse region can be said to contribute to itself. It would certainly be confounding if the vocabulary and practices undergirding T_1 were not conducive to the interests pursued with T_1. But this sort of trivial reflexive case will be true of many theoretical notions. All theoretical kinds are identical with themselves; all theoretical vocabularies can be reduced to themselves; all properties supervene on themselves, and so on. Setting aside the reflexive case, we suspect that symmetrical contribution relations between discourse regions will not occur.

So contribution relations will be all-but-universally *asymmetrical*. One type of discourse will contribute to another, but the same sort of contribution will not be made in return. This sort of asymmetry is commonly noted in the relations between different scientific discourses. Physics terms appear in chemistry (importing laws that help explain their laws and mechanisms), chemical terms figure in models of the mechanisms in biology, and so on up through neurology, psychology, and the social sci-

ences. So goes the orthodox picture in reductionist accounts. But even if one is not a reductionist about such matters, one may end up adopting the "weak reductionist" claim—as Kitcher (1984) called it—that explanation and composition always flow upwards from more fundamental theories to less fundamental ones. These familiar examples are asymmetrical contributions across different discourse regions, and it is fair to say that contribution does not flow *downwards* (or thus, *symmetrically*) in these cases. But these are not the only cases to consider, nor is symmetry the only possibility here. It may also be possible for two discourse regions to have two sets of asymmetrical contribution relations between one another— one in each direction, each of which is asymmetrical, but which are very different in character from one another. Thus, there might be some R_1 that holds asymmetrically from T_1 to T_2, and another relation, R_2, that holds asymmetrically from T_2 to T_1 (Fig. 8.1).

So we might further characterize the interrelation of two discourse regions as either *unilateral* or *bilateral*. For instance, physics contributes *asymmetrically* and *unilaterally* to biology. Physics helps explain certain biological phenomena, but nothing in biology explains anything in physics.

What would a pair of discourses that contributed to one another bilaterally and asymmetrically look like? Consider the relation between theories in psychology and economics as an example. In some cases, claims and models from cognitive psychology will contribute to explanatory models in economics. The clearest example of this would be what is frequently called "behavioral economics," much of which can be traced to Herbert
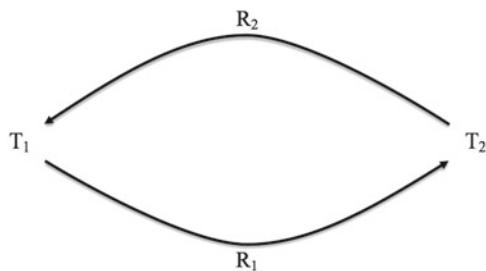
Fig. 8.1   Bilateral asymmetric contribution relations

Simon's work (1955, 1956). Frustrated with the persistently suboptimal performance of classical economic models, behavioral economists have argued that at least some phenomena can be better explained with more psychologically realistic models of the economic actors in play. Accounts positing limited self-interest and bounded rationality, which emphasize the roles of satisficing strategies, heuristic reasoning, and excessive aversion to risk have met with considerable success by many estimates. Here, we have a set of theoretical resources for explaining the behavior of individual agents serving as partial explainers for a presumably autonomous theoretical discourse that models larger dynamic phenomena. Note an important dimension of this shift to larger dynamic phenomena: these are economic models using psychological components, not psychological models. The heuristics and biases invoked there are models of individual agents' cognition, but the economic models will frequently generalize over frequencies in populations or collections of events. (By analogy, theorizing about the frequencies of genotypes—molecular sequences—in populations does not turn evolutionary biology into chemistry-in-disguise.) Now, these are empirical hypotheses that must be vindicated by future research, and juries are still out on their scope and utility. But any move to offer more realistic models of individual economic agents would follow similar paths, and to the extent that such strategies serve our theoretical interests in developing economic theories, we will see such contributions.

This gives us an asymmetric contribution relation from psychology to economics, as the explanatory contribution made there is not made in the reverse direction from economics to psychology. But there might be other sorts of contributions made from economics to psychology. To mention one such contribution, since the early 1990s, a great deal of research has gone into how persistent poverty during children's earliest years inhibits normal cognitive development. The literature here is vast, and we cannot do justice to its scope and variety in passing here. But useful recent works emphasizing poverty in particular might include (Hair et al. 2015; Jensen et al. 2015; Walker et al. 2007). It might be tempting to wave off the connection to economics here and insist that the pertinent causal links here are simply with biological or physical conditions (such as poor nutrition) during the critical period. But this will not do for

a number of reasons. Sub-poverty income levels are a strong predictor of these developmental inhibitions in longitudinal studies, but only against an array of background structural conditions that economics provides us; income below the US poverty level would go farther elsewhere and at other times, and so adjustments must be made for cost of living, remediation efforts by governments, government services, and so on. And even given those adjustments, the causal pathways through which persistent poverty can have these effects require those complex structural conditions as well (e.g., poor nutrition as a result of "food deserts" due to weak demand and poor public transport, increased stress load due to persistent economic insecurity, etc.).[2]

So rather than a simple "bottom-up" arrangement of discourse, with lower-level discourse regions explaining higher-level ones, or the higher-level ones reducing to lower-level ones, we have different regions contributing to one another asymmetrically and bilaterally (Fig. 8.2).

We should note that wherever this bilateral asymmetric arrangement holds, R_1 and R_2 will be substantially different from one another. When one discourse region serves another in these ways, it has something to offer that the other does not, and thus the other cannot offer in return. In our example, T_1 provides a finer-grained description of some of the elements in a much larger dynamic system, while T_2 provides a more structural description of conditions that will create causal pathways to affect those local elements, but which do not fit the sort of finer-grained
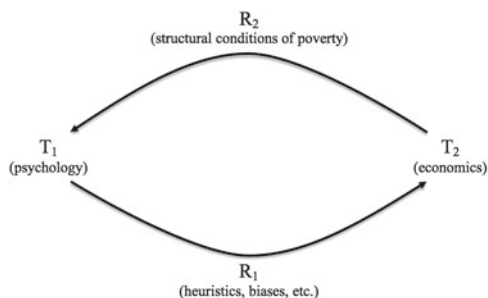
Fig. 8.2 Some bilateral asymmetric contribution relations between psychology and economics

theoretical description of individual agents we pursue with T_1. The fine-grained details of T_1 accumulate in ways that inform T_2 without then distributing back downwards to individual agents. The structural conditions described in T_2 serve as a kind of addition to the environment in which cognitive development will occur, increasing the likelihood of some causal pathways over others. Neither account is deficient in these respects, they are simply pitched at different sorts of interests. What satisfies those interests can be of further service to others, but to expect all such successes and all such relations between them to exhibit a single, uniform character—even at some higher level of philosophical analysis—is to impose demands for uniformity where we should embrace diversity.

Those who are wary of our notion of contribution may object here that the diversity we are promoting in intertheoretical relations results in a notion that is too lax. Without stricter criteria, any connection between discourse regions could count as contribution, and stricter criteria would likely push us back toward reductionism or eliminativism for some discourse regions. We could imagine an interlocutor objecting here:

> Quantum mechanics makes indispensable use of the term 'quark'; the term appears in Joyce's *Finnegan's Wake*, a classic of modernist literature; does modernist literature contribute to quantum mechanics? Given the looseness of the definition, it satisfies the conditions for contribution.

This would be doubly bad news, in that the only thing more inscrutable to most people than quantum mechanics is *Finnegan's Wake*. More to the present point, the flimsiness of this example would suggest that our notion of contribution permits too much. This objection can be met relatively easily, we would say. While Gell-Mann's introduction of the term occurred in connection with *Finnegan's Wake*, nothing about the novel itself or the literary practices in which we locate it serves the interests driving physical discourse involving quarks. He could just as easily have plucked some other neologism from comic books, the street slang of the early 1960s ska scene in Jamaica, or countless other sources. He could have simply stuck phonemes together until something sounded both alien and appealing enough.[3] Doing any of these, rather than dipping into *Finnegan's Wake*, would have made no difference to the phys-

ics, and so any purported necessity to the use of modernist literature in articulating the discourse of physics is easily eliminated. The sense that *any* discourse could contribute to any other emerges only if we pay insufficient attention to the role assigned to theoretical interests and the ineliminability condition we stipulated above. This contrasts with the psychology/economics example, in which one theoretical discourse really did need to borrow from the machinery of another, and failure to do so would leave us with an explanatory void.

This may point to a more trenchant objection, however. Suppose instead that the objection noted the use of everyday talk of physical objects. Thus, instead:

> Physics and chemistry regularly employ parts of "folk physics" (i.e. everyday talk of macro-sized objects); these figure prominently in research practices and no urgency is assigned to eliminating them; thus, folks physics contributes to physics and chemistry.

Examples of this would seem to abound. Schrödinger had his cats and bank deposits, Feynman had his indestructible bullets. One of us (Wolf) was taught a little bit of valence shell electron pair repulsion (VSEPR) theory by a professor as "various strangers in an elevator person repulsion theory," which illustrated the configurations that electron pairs in a valence shell would take by comparing them to bashful strangers in an elevator shuffling away from each other to maximize their personal space. But these are the stuff of popular and pedagogical repackagings. Much like our *Finnegan's Wake* example, it seems as though physics and chemistry could be done without these items, even if that made them harder to teach to undergraduates. To make the stickiness of the objection much more evident, it bears remembering that we have argued that the discourse of physics (and other such theoretical discourse regions) should not be limited to their central formal theoretical statements, but should also include discourse about the conduct of the practices themselves by which such central theoretical commitments are determined. This entails that physical discourse should include not only its statements of laws as equations, but also about the equipment by which its research is conducted, and the methods and standards by which hypotheses are offered and results evaluated. That

would seem to entail commitments to medium-sized concrete objects and agents playing around with them in more substantial ways than mere pedagogical illustrations. This might conflict with the ontological priority of the scientific image that many naturalists and those who undertake inquiry into physics take it to have.[4] Rejecting object naturalism, red in tooth and claw, is one thing, but does our notion of contribution leave us with a position on ontology that has no teeth at all?

We think the right response here is simply to defer to ongoing inquiry in any number of fields. We have presumed a relatively generous ontological pluralism within the natural world (physical, biological, etc. entities) at many turns, so long as our commitments arose out of theoretical need. Maintaining all those levels would involve appeals to supervenience, which would not trouble us. (Our objections in Chap. 3 were to accounts of normativity steeped in supervenience, not to such accounts in general.) Our aim in this book has been to carve out a way to understand normative discourse without ontological commitments, but with an integration into our other forms of inquiry and coping with the world that makes them truth-apt. Whether we can or should save other forms of discourse such as everyday talk of macro-sized objects is a great philosophical question, but it is not the one we are tackling here and we do not purport to have a novel or easy answer up our sleeves. It does matter to us that there remains the possibility of action-orienting conceptual content—we have to have some way of table-talking and book-talking to say "Put this book on the table." And a lot of discourse about artifacts is not eliminable without considerable loss in explanatory power. Consider, "The car won't start because the fuel pump is busted" and try saying that—or explaining to someone how to fix the car—without talking about fuel pumps. Implicit here is a kind of Harman-style assumption that essential involvement in causal explanation implies the need for genuine ontological posits. That presumes these are genuine explanations, and that our need for them will persist. But as we have shown in the preceding chapters, we can sometimes articulate the content of some forms of discourse in ways that preserve their intelligibility and truth-aptness without thereby accruing further ontological commitments. Even if we feel the need to defend these forms of common sense discourse (or some fragment of them), our general commitment to the importance of cross-discourse contribution

need not entail positing such objects and properties. Maybe everyday objects *will* be replaced or eliminated in the future, but recasting all of common sense talk about objects in expressivist terms without commitments to objects and properties strikes us as both a monumental task and one with a somewhat shaky prognosis. We have no dog in this fight, at least in this book, so on this count we will defer to future inquiry.

8.3 Two Genera of Discursive Contribution: Causal-Functional Mechanisms and Mathematical Formalization

We take it that each type of intertheoretic contribution is its own beast in the end, and the basis for its inclusion and our ongoing commitment to it will be articulated in terms of the unique fit between theoretical interests pursued in the two related discourse regions. The central assumption of views we oppose here is that there must be a single type of relation reaching upwards from a fundamental discourse (or its items) toward any other putative legitimate one, and that any discourse region that cannot be determined by or derived from more fundamental ones is thereby suspect. Property-identity and law-derivation forms of reductionism would be such views, as would requirements of supervenience. But to stake out this sort of moderate pluralist view is not to deny that there are sometimes such relations between discourses, or at least parts of them. There *can* be theoretical identities! The laws and models of the social sciences *cannot* violate the laws of physics! We simply take it that these sorts of relations are a slice of a much larger field of possibilities.

Our move to the notion of contribution is made with an eye toward preserving unity among discourse regions without expectations of uniformity among them, as reductionism or strict physicalism would demand. We have made indispensable use of the notion of an interest in this account, suggesting that our discourse about our world must be understood as a type of practical engagement, not a passive representation. In Chap. 5, we suggested that our theory-building interests would coalesce in ways that distinguish different discourse regions—there are goals we

adopt and methods we pursue building biological theories that are different from the goals and methods in building physical ones. Our pluralism and emphasis on the local in much of our discussion of contribution so far grows out of this general view. But our sort of localism about contribution does not preclude noting similarities and even some general qualities across different contributions that can usefully (if only partially) characterize the work done between different discourse regions.

An examination of some of these general features would shed light on how one discourse region can serve the interests of another without subsuming it. Contrasts could then be drawn with other types of contribution for further illustration. In this way, we might shed some light on how an irreducibly plural landscape can still submit to useful metatheoretical examination. We cannot do this exhaustively here, sketching out every general type of cross-discourse contribution. That would be to cover the entire landscape of human inquiry, and our publishers are (quite reasonably) not that generous with their paper. But more importantly, we do not see any a priori case to be made for the derivation of some complete list of general categories of contribution (i.e., something like Kant's categories). A more productive strategy here would be to cast some crucial and familiar types of intertheoretic relation in the sort of terms we have used to characterize cross-discourse contribution. We want to see what something like a theoretical identity or a causal mechanism could *do for* the higher-level discourse—in very explicit practical terms—without replacing that higher level's pursuit of a set of interests.

With this goal in mind, in this section, we will offer a sketch here of a familiar species of intertheoretical relations, which we will call *causal-functional mechanisms*, and a short historical sketch of how this actually played out in the field of endocrinology. (To avoid making this account even more cumbersome, we will sometimes speak of endocrinological substances themselves and their place in various mechanisms. The account however, is ultimately one about how regions of theoretical discourse function.) We then draw some contrasts with the ways in which mathematics might contribute to other fields, emphasizing some salient differences rather than an exhaustive account of how mathematics may be woven into other discourse regions. Both of these will be what we have called *unilateral asymmetric* contribution relations. We then close

the chapter with some lessons to be drawn for a broader account of how normative discourse may exhibit bilateral asymmetrical contribution relations with other discourse regions.

8.3.1 Causal-Functional Mechanisms as Vertical Contributions

To emphasize that we are examining causal-functional mechanisms as a form of contribution is already to introduce a type of specification to a more general notion. To offer a functional account in the broadest sense is to give some means of mapping sets of input elements onto sets of output elements. That mapping may be purely stipulative. By adding the "causal" distinction here, we are restricting our focus to those that are physically realized and discovered, presumably by observation, but for which there is some variety in the physical conditions that could realize the functional pattern. So in the present sense, just to name a few, various enzymes play causal-functional roles of catalysis in metabolism; various neurological structures and components play causal-functional roles in cognitive processes such as memory retrieval; various actions and conditions play causal-functional roles of increasing inflation in the economy of a country or region. But in each of these examples, there are nomologically possible alternative realizations of the functions played as we actually find them. We can at least imagine other molecules catalyzing reaction in a very different sort of organism, for instance. So go the details of multiple realizability, familiar in functionalist accounts since the 1960s. Note also here that the discoverable character of causal-functional mechanisms as we are describing them contrasts with items that we develop with functional constraints in mind. Hammers, mousetraps, and computers have causal roles to play that can be characterized functionally, and can be realized in various ways with different materials, but what makes them such items is generally a matter of our crafting them with designs in mind and taking them up as such.

Multiple realizability give us grounds to assign discourse about such mechanisms a degree of autonomy from lower-level discourses, but how much? And what should it permit (or prevent)? As we have said, we con-

tend that there are good general reasons to favor the integration of discourse regions where possible and conducive to the interests they serve, but if a theoretical discourse has this sort of autonomy, why should we be concerned with integrating it with lower-level theoretical accounts? Perhaps a functionalist account of the mental could suffice without any appeal to biology or neuroscience at all. One might even think that such an attitude could apply, *mutatis mutandis*, in any case where a discourse region turns out to have autonomy. There are a number of reasons to reject this. Autonomy for a discourse region will actually reveal itself to be a patchwork affair, with one region sufficing on its own for some purposes while needing support in others. A causal-functional account of a cognitive process like long-term memory retrieval can be described in terms that treat memories as internal representations manipulated by various coding and retrieval mechanisms, all of which are amenable to functional descriptions without attention to their physical substrates. But for other questions, such as why certain encoding strategies are more effective than others, or why various memory disorders emerge in some subjects in just the manners that they do, the answers will lead us back to biochemical and neurological accounts of the systems that realize those causal-functional mechanisms. Further appeals to the higher-level causal-functional mechanisms on such questions just will not do.

We can make the integrationist case here in a more positive manner, however. Think of the sort of contribution we are considering here as *vertical* contributions (coming "up from below," as it were). With these, we account for some feature of a higher-level discourse in terms of a lower level one, as when molecular properties help us explain features of classical genetics, or neurological properties help us explain features of cognitive psychology. A theoretical identity statement ("water is H_2O") would be one such vertical contribution, and it would establish a license to substitute each of those expressions for one another, though the more fruitful substitution inferences would be into sentences that expressed laws or other claims in the contributing discourse (chemistry, in this example.) A mechanism that explained some higher-level phenomenon could thus be seen as a number of such identities, structured temporally, plus particular instantiations of laws to account for the pathways the mechanism followed. Not all contributions will be vertical in this sense, contrary to

reductionism. (In particular, note that the bilateral asymmetric contributions described in Sect. 8.2 will not be vertical; the pertinence of that set of distinctions will become more clear in Sect. 8.3.3 and Chap. 9.) But there are still considerable practical payoffs from success in establishing vertical contributions. We think there are four loosely characterized types of theoretical benefits that are immediately noteworthy.

First, vertical contributions often add to the *plausibility* of a causal-functional mechanism. Given the expectation that these patterns will be physically realized, their occurrence would amount to accidents or mere correlations if their regularity were not the outcome of some features of their substrates. This will be the case more often when the vertical contribution is made from a discourse region in which our confidence is higher, as when the biochemistry deployed in molecular genetics makes clear how the distribution of traits already observed (but not yet explained) in classical genetics is possible.

Second, vertical contributions will frequently open up the possibility of directly manipulating instances of the mechanism. To use a medical example, the biochemical mechanisms of immune response, better understood in chemical terms now, are actively manipulated in the process of vaccination to modify response to pathogens. More recently, our theoretical account of such mechanisms has facilitated the development of antiviral and antiretroviral medications that manipulate immune response even more directly. For instance, the discovery that retroviruses replicated by reverse transcription (a causal-functional mechanism, in our sense) allowed researchers to prevent this mechanism to a degree by developing compounds that would selectively inhibit the enzymes needed for it. These processes are often facilitated or improved by a more precise account of the molecular structures that those pathogens exhibit.

Third, a vertical contribution may provide ways of projecting from observed patterns to new ones, thereby extending and refining our account of the mechanism, or directing us to others. If we have ample explanatory and predictive capacities in the discourse making the contribution, projecting how the mechanism will behave under different conditions may be feasible, whereas without vertical contributions, it will often be impossible. Medical research will often proceed this way: determine the underlying conditions that manifest as a pathology (which will

sometimes be a causal-functional mechanism in our sense), then search for measures that will alter those underlying conditions.

Fourth, vertical contributions will frequently serve as explainers (if not necessarily good predictors) of deviations from ideal formulations of a causal-functional mechanism. Some discourse regions also include these sorts of corrective resources internally for their own causal-functional mechanisms, (e.g., analytic chemistry, which allows us to correct for errors due to impurities in materials, the background noise of equipment, etc.) The relation between psychology and economics outlined in Sect. 8.2 would be another such example.

Consider a somewhat extended example of two discourse regions that have extensive vertical contributions between them: endocrinology and chemistry. (Many other examples would do here; no greater reason than familiarity lurks behind our choice. For simplicity's and brevity's sake, we will occasionally speak directly of the endocrine system and its parts, bearing in mind that our point here is how all of these items and relations are described and explained discursively.) Most notably for our purposes, endocrinology begins to mature as a scientific discourse during the early twentieth century as the molecular structures of several key hormones are determined. Chemical terms have become the coin of the realm, and every key concept and model in endocrinology will have a dimension that is expressed in chemical terms. But like many other cases favored by anti-reductionists, this pervasive role for chemistry and the consistent discovery of type-identities does not somehow render endocrinology otiose by exhausting its content in another theory's terms.

Endocrinology enjoys a degree of autonomy from chemistry and physics in ways that will be analogous to other anti-reductionist examples. Candidates for reduction targets will partially describe some features of the endocrine, but higher-level phenomena would appear gerrymandered if we tried to squeeze them into the laws of chemistry or physics. And in some cases, even where there are very similar chemical structures between two hormones, they will have very different roles in the system (e.g., oxytocin and vasopressin [ADH]). Even the notion of a hormone itself does not admit reduction to a single chemical structural property; instead, we find three major classes (peptide, amino acid derivative, and steroid) with numerous subclasses within them. Nor are there simple macro-level dis-

tinctions such as organ size or tissue type that capture the distinguishing features of the system. Some glands produce hormones, but so do some individual cells. The hormones themselves also exhibit different scopes, with some acting on almost every tissue in the body (e.g., insulin), and others targeting only very specific tissues or the glands of the endocrine system itself (e.g., thyroid-stimulating hormone (TSH) and adrenocorticotropic hormone (ACTH)).

A more general characterization of the endocrine system would include the *signaling* dimension of its causal-functional mechanisms, and their role in the control and development of other systems, organs, and processes. The signaling or communicative dimension should not be thought of as fully linguistic, at least not at the level of one with a grammar of the sort in the natural languages spoken by humans or formal languages employed in mathematics, or computer science. Rather, the role played here is one of regulation of one part or process of an organism by another (e.g., testosterone stimulating sebum production in the sebaceous glands or terminating bone development in late puberty). Glands in the endocrine system signal, and thereby regulate, such conditions by secreting hormones—chemical agents that stimulate, or inhibit those outcomes. That abstract level of description—the endocrine system as a signaling system to regulate biological processes—may be too general to capture the right level of organization, though. To regulate biological processes, the components of the endocrine system monitor their environments, adapt as needed to maintain homeostasis, and do it all with chemical means. But we can say just the same of the nervous system. We also find cells devoted to communication *between* the nervous and endocrine systems, particularly in the hypothalamus. Cells also sometimes signal to themselves in order to self-regulate (*autocrine* processes) or do so to neighboring cells (*paracrine* processes). The crucial distinction is that processes of the endocrine system involve the secretion by glands of hormones into the bloodstream (distinguishing the system from neurological processes that do not) to affect glands and cells other than themselves (distinguishing it from autocrine and paracrine processes).

The functioning at the core of endocrinology centers around homeostasis and extends to the regulation of processes in reproduction. Control functions (negative feedback, positive feedback, inhibition, endocrine

rhythms) are sometimes directed toward stabilizing some ideally static feature, but in many others, it is about stimulating or supporting a dynamic process (e.g., reproduction, lactation) that must unfold within some narrow range of possibilities in order to secure outcomes (e.g., offspring, nutrition) that chemistry would quite properly be indifferent toward. Ignoring this level of organization would be to amputate our theoretical interests rather than serving them, so a degree of autonomy is due to endocrinology as a theoretical project. And yet, there is a clear sense in which there is vertical contribution from chemistry in that pursuit. It tells us *how* the endocrine system can pull it off and just what it uses to do so. That is to contribute to our theoretical pursuits, even if it is not to replace one with another.

So let us return to the four payoffs of vertical contribution mentioned above and consider each of them with some details from the theoretical form and history of endocrinology. How does the molecular account increase the plausibility of causal-functional mechanisms and their posits in endocrinology? Postulates of something like sex hormones have been around for centuries, and presumed to be part of gonadic function. Concentration—relatively good extraction of the hormone, but not nearly perfect isolation—was done with dog and guinea pig testicles, which were then injected into human subjects with noticeable effects consistent with those earlier posits. Later, larger amounts were isolated from bovine testicles, and when those were injected into capons (castrated roosters), they had the effect of reactivating masculine features such as the growth of their combs (Gallagher and Koch 1929, 495–496). This provided fairly strong evidence for the presence of what we would come to call androgenic hormones and their causal-functional role in a variety of other biological mechanisms. But even at this stage, whether to posit a single androgenic hormone or several was a provisional matter.

> Whether one or more active principles are concerned is at present mere speculation… [U]ntil more is known of the chemical nature of the hormone no name should be given the extract. As yet any name would be valueless and not at all descriptive. Too often a name gives a false sense of security as regards the purity of the product… (1929, 500)

In 1935, Ernst Laqueur's research group would successfully isolate testosterone, which permitted a determination of its molecular structure. With this in hand, molecular explanations of the generation and mechanisms of action for androgens would demonstrate how the observed effects could be produced by the kinds of secreted substances being postulated. The vertical contribution from chemistry thus supported the posits of items in the causal-functional mechanism and accounted for the reliability with which the input conditions would generate the outputs.

How does the molecular account make it possible to manipulate endocrinological processes? The isolation of testosterone and subsequent determination of its molecular structure opened the possibility of synthesizing it from other substances. Laqueur's isolation of testosterone in 1935 was followed quickly by two research groups who chemically synthesized it successfully (Nieschlag and Nieschlag 2012, 8–9). That transition was fairly rapid, in fact; the structure was published and synthesis achieved within three months. This opened a wide range of medical applications almost immediately, leading to both oral and injectable forms of testosterone for regular clinical use. Success in establishing the molecular structure of testosterone also permitted a variety of projections from observed patterns to new ones. Once its characteristic ring structure had been determined, existing theoretical accounts facilitated hypotheses and eventually successful synthesis of other anabolic steroids; at present, there are over a thousand described in the literature (Nieschlag and Nieschlag 2012, 9).

How does the molecular make it possible to explain deviations from ideal patterns? Consider explanations of androgen deficiency in younger males. (Some decline in testosterone level is normal with age.) The most common explanations are diagnoses of trauma to the tissues that secrete testosterone; damage to those tissues simply prevents the chemical processes that would generate testosterone. Disorders and illnesses such as Kleinfelter syndrome and childhood mumps have among their effects the prevention of normal development of tissues that secrete testosterone or its precursors. So again, in such cases, the deficiency will have a simple molecular explanation: the absence of the conditions that normally lead to the generation of testosterone. Once trauma and other illnesses are eliminated, the most common cause is a pituitary tumor—a microade-

noma or macroadenoma (only difference is size). The deviant effect of the tumor is actually mechanical: it physically blocks the release of pituitary hormones by compressing other cells and either preventing the secretion of luteinizing hormone (LH), or not allowing it to travel normally. And the simple molecular fact of lower levels of that compound explains the presence of less of the compounds whose production it stimulates.

So in this extended example, we can see some echoes of features common to other anti-reductionist accounts of intertheoretic relations. We have in the notion of signaling a distinguishing feature of the higher-level theoretical discourse, with patterns of reasoning from chemistry filling in blanks in the mechanism by which signaling takes place. That much will be familiar from the Fodor and Kitcher strategies discussed in Sect. 8.1. Filling in those details matters only within the contours of the theoretical project however, and those contributions do not supplant the higher-level discourse or satisfy the theoretical interest we pursue with it. In the pursuit of other theoretical interests, those sorts of vertical contributions may offer only very occasional benefits, or none at all. The ubiquity of vertical contributions from chemistry in endocrinology reflects the utility of those patterns of reasoning to the goals pursued in the higher-level discourse, which are themselves bound up with other bundles of interests both theoretical and non-theoretical, such as the practical bearing of this theoretical knowledge on health and welfare.

8.3.2   Mathematics as Structural Contribution

So it goes for one familiar sort of non-reductive contribution. But this might seem a paltry response if we seek to expand naturalism's horizons and integrate normative discourse with its non-normative brethren. If all we had as a non-reductive option were causal-functional relations, that might only add a variation on old themes. Contribution does not have to end here though, and we should draw some contrasts with other types whose roles are not apparently causal at all. We suggest a good interim step here would be to consider the discourse of mathematics. While most scientists and philosophers would argue that true mathematical claims can be made and that true claims can be made about physical items and

their behavior involving mathematical vocabulary, we could not point to any serious claim that physical discourse *reduced* to mathematics. A strongly realistic position on the ontology of mathematics will insist that mathematical objects do play something like a compositional role in physical states of affairs. On such accounts, they are abstract objects instantiated in the physical: some cardinality really does inhere in a set of objects, some value really does inhere in the acceleration of a particle as it travels through a field, and so on.[5] We will not take any stand here about the reality of mathematical objects, though as Sect. 1.2.2 would suggest, we would be comparatively hard-nosed about admitting abstract objects in general. What seems clear however is that even if we are realists about the ontology of mathematics, these items will not play the role required for a reductionist account. Even if we accept real cardinal properties in the world, threeness does not compose a set of three photons in the way that hydrogen and oxygen atoms and their bonds in a molecule compose water molecules. The mathematical does not compose or supplant the physical (biological, etc.), even if nothing in any other domain can violate the laws of mathematics.

We spoke of vertical contributions in the previous section, and now we can speak of mathematical discourse as providing *structural* contributions. These would involve types of structures that could recur across different discourse regions in the same fashion: cardinality would be such a structural property, as would geometrical properties like angularity, orthogonality, perpendicularity, and many others. Note that this is not yet an ontological claim about whether or not mathematical objects and properties exist (we make no claim to settle such matters), but a claim about the overarching goals of mathematical theorizing and the development of licenses and restrictions on assertion and inference across the entire field of discourse, including itself. We also sometimes speak of molecular structures (among others); in our current set of terms, those would require structural contributions from mathematics, as when we indicate the *number* of atoms of some type in a molecule, or its *valence*, or the *angles* between bonds. So note that the vertical contributions described in the previous section will more often than not incorporate some structural ones, as well. (You cannot have "H_2O" without "2.") The intent here is to get at implicit commitments baked into richer notions

like molecular structures. To put matters more precisely in the terms we have employed, structural contributions from one discourse region to another will involve concepts and expressions from a contributing region (e.g., mathematics) adding additional patterns and relations among the elements of the receiving region. Their inclusion will thereby allow us to make further claims and inferences by drawing on the rules and principles of the contributing region, which apply without restriction to the terms, concepts, and so on of the receiving region. A paradigmatic example would be a law of physics that made essential use of various mathematical terms in its expression. This differs from vertical contribution in that a structural contribution does not purport to tell us what composes the items in a discourse region, but how they are arranged. So in our laws of physics, force, mass, and acceleration may be put into a particular set of mathematical relations, but mathematics does not tell us what composes the physical items involved. We can then make further inferences drawing on the rules and principles of mathematics, as when we use the methods of algebra to solve an equation for a different variable (e.g., deriving "$a = F/m$" from "$F = ma$").

Accordingly, the distinguishing feature of mathematical discourse is its *maximal entity-type neutrality*. The claims and inferences that mathematical distinctions license, permit, and prohibit should not be restricted by the types within other discourse regions to which they contribute. For instance, the members of any set in any discourse region should be countable in just the same way wherever they are countable at all. A set of 17 items will be greater in number than a set of 11, regardless of what the elements of the set may be. By contrast, the terms, concepts, laws, and so on of chemistry do not make (as far as we can see) general, systematic contributions to, say, economics. (At most, they will have peripheral explanatory roles, such as explaining why some commodities rise in value at a specific time given their chemical properties and a wealth of other background assumptions.) In this way, mathematics can be woven into important structural claims in every other discourse region. Not all of mathematics can, and we choose "maximal" above for good reason. Geometrical terms will not contribute significantly to discourse regions that do not make use of spatial distinctions, for instance. So while we may see the terms of one discourse region recurring in different vertical

contributions (e.g., the individual elements in chemistry may contribute to many other discourse regions), each such contribution will be distinct from the others (e.g., the molecular structures that figure in neuroscience will be different from those in molecular genetics, polymer engineering, etc.). But the *same* structural contributions may be made to many different discourse regions (in some cases, all) in just the same fashion, (e.g., the resources of calculus play central roles in physics, neuroscience, economics, and many other theories).

Given the emphasis we have placed on the practical and the pursuit of theoretical interests, the contrasts between vertical and structural contributions laid out here suggest that what makes mathematical discourse so deeply interesting to us is not that it tracks or represents a particularly esoteric set of entities. Rather, it marshals and unites methods of reasoning that enjoy the widest scope and the least irregularity possible. We trade off some of the rich detail of other forms of discourse, but the payoff in developing methods that can then be redeployed across other regions is immense. This is not to deny that there could be real mathematical entities; it is to make a point about the integration and application of methods of reasoning. Without that practical story, their existence is irrelevant, while with it, their non-existence might not be problematic. (While ardent nominalism or fictionalism might eliminate mathematical entities, no viable version should eliminate mathematical discourse altogether.) That, in turn, suggests an alternative to the deeper logic of reductionism. The very image of many "levels," each composed of another's entities until we bottom out in the physical, each with its laws derivable from those below it, is a misconception projected from some types of cross-discourse contribution (admittedly, important ones) to all of its regions. There are "bottom-up" contributions in the classical reductionist mode, but there are also cases that move "top-down," structural contributions that suffuse others without regard to those hierarchical orderings, and pairs of theories that relate bilaterally and asymmetrically, as we shall see. The task before philosophers and scientists is not to find the theory from which everything else descends, but to map overlapping regions of the pragmatic terrain where these different projects support and inform one another.

8.3.3   Lessons On the Role of Contribution for an Account of Normativity

As we emphasized in the opening to Sect. 8.3, the descriptions of causal-functional mechanisms and mathematical formalizations are only initial sketches of types that would ultimately be fleshed out with greater levels of local detail. They are also not intended to exhaust the potential variety even at the very general level at which they are pitched. We will demonstrate in Chap. 9 how a number of bilateral asymmetric contribution relations characterize the connection between normative and various types of non-normative discourse, making good on the more complex arrangements described in Sect. 8.2. These will not be the sort of unilateral, asymmetric arrangements we see in vertical or structural contributions, and we are now in a position to spell out how Chap. 6 and Sects. 8.3.1 and 8.3.2 demonstrate the need for a different approach to normative discourse.

The familiar point to most readers will be that the vertical and structural contributions we have described here account for properties, regularities, and states of affairs, whereas normative discourse guides action. (Given our pragmatist commitments, we think *all* content guides action in a sense, but we elaborated a more specific sense of this in Chap. 6.) To "account" for regularities here will entail having paradigmatic ways of describing them that may then be translated into specific positions in the problem-solving strategies and other explanatory approaches of other discourse regions (e.g., the composition accounts of the input and output states in our endocrinology example, coupled with laws and models to explain the regularity of the transitions). Vertical contributions are thus interpretations of a salient isomorphism between parts of one discourse and parts of another in terms of the theoretical interests guiding the supervening discourse region. But in such cases, both sets of theoretical interests are satisfied by the discovery of patterns and regularities; successful contribution occurs when the two sets of patterns line up in ways that are fruitful to those types of theoretical inquiry. There is a fundamental disanalogy here with normative discourse, in that the direction of action that it provides anticipates conditions which may not yet obtain, or which

have not yet been confirmed. There is no normativity without the potential for our actions to make a practical difference; if something would happen or be so anyway, or would never happen no matter what we did, then normative claims make no sense. We have norms to do X because we would expect not-X if we did not act in certain ways. Normativity kicks in when something is not the case or will not be the case *unless we act in the ways prescribed*. They are thereby unlike predictions, which we take to be claims about conditions that will obtain, however we attempt to act. In this way, even predictions about our own actions are fundamentally different from normative claims. If my friend is an alcoholic and I predict that he will relapse several more times before recovering (if he does at all), then my claim is not *directing him* to drink where he otherwise would not have; it is reporting (with dread and regret) that this pattern is to be anticipated of him in any case.

And in this respect, non-reductive supervenience accounts of normativity were always going to be just as troublesome as reductionist accounts. Both these types of account take *identifying* some non-normative property or pattern to be their task, then expect those properties to somehow determine normative ones. Presuming that vertical determination between levels was the only sort of contribution available and imposing that demand on normative discourse makes these two versions of the same error. Historically, we often find some form of an is/ought or description/prescription distinction invoked when this contrast is noted, and assurances that we cannot derive one from another. That is to make the present point about vertical and structural contribution at hand. The expressive roles normative discourse plays, particularly the expressive modes outlined in Chap. 6, simply do not align with the sorts of patterns of reasoning expressed in non-normative discourse regions. To guide action and to describe it (or its outcomes) are simply very different types of linguistic engagement with the world.

So the purport of normative and non-normative discourses is fundamentally different. But this is often read to imply that normative discourse is then conducted in isolation from the non-normative. This is autonomy run amok, and it threatens the integrity of normative discourse, as we have warned. Instead, we suggest that there will be bilateral asymmetric

contributions between various non-normative discourse regions and normative ones. What we must look to are the ways in which the character of non-normative items bears on the articulation of interests in normative discourse because they shape the embodiment conditions under which those interests are pursued. In this way, the harms considered in normative ethics will be informed by the biological and psychological conditions in which we are embodied. Likewise, physics, biology, and cognitive psychology will contribute to our discourse on epistemic standards; what we can demand or reward in epistemic agents reflects what is possible for beings like them.

This suggests a few things about the contribution that non-normative discourses can make to paradigmatically normative ones, and we will elaborate these themes further in Chap. 9. What about the other direction? How do paradigmatic forms of normative discourse contribute to non-normative ones? Here, we shall look to the ways in which the exercise of various kinds of non-normative theoretical inquiry needs to be self-correcting, noting that the means of this sort of self-correction are not the objects of study in those discourses. Physics requires a concern for how to do physics better, but the subject of study in physics is not *doing physics better*. Non-normative discourses will need resources to regulate their own exercise, and this gives us grounds to consider standards in an abstract fashion that will then be articulated more fully and precisely within the non-normative discourses themselves. A paradigm of such metatheoretical contribution from normative discourse to non-normative discourse would be epistemic standards being brought to bear on some form of non-epistemic discourse. We can think and speak about how epistemic concerns operate in general with some autonomy from the details of other theoretical projects. Those insights can then be brought back to other types of discourse, and their local details and application fleshed out. A fuller account of how the interests we express in normative discourse shape inquiry would thus elaborate different epistemic norms and the development of inquiry-worthy problems in other discourse regions. A more thorough presentation of such epistemic and ethical norms will thus be a focus of our discussion.

Notes

1. And even this does not imply that mathematical claims might not be invited, initiated, or inspired by empirical observations. Topology may begin with Euler thinking about bridges in Königsberg, for instance, but this does not ground topology in civil engineering.
2. Two things should be noted here. First, there will be other discourse regions involved in an explanation of inhibited cognitive development under persistent poverty besides economics, such as education and social psychology. But all the more grist for anti-reductionist mills. Second, which pathways account for developmental inhibitions and to what degrees are still fiercely debated. Food deserts seem to be on the wane lately, while others are on the rise.
3. In all historical honesty, the last one here is not far from the truth. Gell-Mann has said that he had a rough sense of the sound of the word he wanted to coin, and stumbled across the ideal spelling in *Finnegan's Wake* (1995, 180–181).
4. For instance, see Sellars (1962) and (1959/1991) and Churchland (1979) in their full-on, take-no-prisoners scientific image mode.
5. Exceptions here would be authors like David Armstrong and Bertrand Russell, who took physical particulars to be bundles of universals. But even those who take up this position do not generally assert that physical particulars are composed solely of mathematical universals, in the way that physicalists would say that biological particulars are composed solely of physical entities and properties.

References

Churchland, Paul. 1979. *Scientific Realism and the Plasticity of Mind*. Cambridge, MA: Cambridge University Press.

Gallagher, T.F., and Fred C. Koch. 1929. The Testicular Hormone. *The Journal of Biological Chemistry* 84: 495–500.

Hair, N.L., J.L. Hanson, B.L. Wolfe, and S.D. Pollak. 2015. Association of Child Poverty, Brain Development, and Academic Achievement. *JAMA Pediatrics*, July 20. doi:10.1001/jamapediatrics.2015.1475.

Jensen, Sarah K.G., Erin W. Dickie, Deborah H. Schwartz, C. John Evans, Iroise Dumontheil, Tomas Paus, and Edward D. Barker. 2015. Effect of Early Adversity and Childhood Internalizing Symptoms on Brain Structure in

Young Men. *JAMA Pediatrics* 169(10): 938–946. doi:10.1001/jamapediatrics.2015.1468.

Kitcher, Philip. 1984. 1953 and All That: A Tale of Two Sciences. *The Philosophical Review* 93(3): 355–373.

Nieschlag, Eberhard, and Susan Nieschlag. 2012. The Medical and Cultural History of Testosterone and the Testes. In *Testosterone: Action, Deficiency and Substitution*, ed. E. Nieschlag, H. Behre, and S. Nieschlag, 1–14. Cambridge: Cambridge University Press.

Simon, Herbert. 1955. A behavioral model of rational choice. *Quarterly Journal of Economics* 59: 99–118.

Simon, Herbert. 1956. Rational choice and the structure of the environment. *Psychological Review* 63: 129–138.

Walker, Susan P., Theodore D. Wachs, Julie Meeks Gardner, Betsy Lozoff, Gail A. Wasserman, Ernesto Pollitt, and Julie A. Carter. 2007. Child Development: Risk Factors for Adverse Outcomes in Developing Countries. *The Lancet* 369(9556): 145–157. doi:10.1016/S0140-6736(07)60076-2.

9

Weaving the Normative and Non-Normative Together

Given that we admit that there are distinct causal-explanatory and normative discourse regions, we owe a story about the interrelations among these discourse regions. We have already denied that that normative discourse receives vertical contributions from a lower-level non-normative discourse, or that normative things are somehow composed of non-normative things (as, say, endocrinological things are composed by chemical things). We wish to urge that the lack of causal-explanatory links between non-normative and normative discourse regions is only problematic if you are in the grip of a certain picture, a picture according to which all discourse is meaningful in virtue of representing facts (hence the moral realist's attempt to separate moral claims into a fact-stating element and a normative element). We have consistently rejected representationalist assumptions, and with them, the assumption that all language must in the first instance serve some explanatory (particularly causal-explanatory) role, or be explicable in terms of such roles.

But if there are no explanatory or causal-explanatory relations between non-normative discourse and normative discourse, then are we failing to satisfy one of our own criteria? Does this mean that nor-

mative discourse is isolated from other discourse regions, and therefore illegitimate? Again, the answer is no, because not all such relations among discourse areas need to be causal-explanatory relations. We will demonstrate that a variety of non-normative discourses do contribute to normative discourse in ways that cannot be eliminated. But the contribution will be multifold, and will connect normative discourse to virtually every region of non-normative discourse. Moreover, we will find the sort of bilateral asymmetrical relations between different normative and non-normative discourses that we described in Chap. 8. There will be ways in which non-normative discourse regions (particularly psychology and the social sciences, but not necessarily only these) shape normative discourse, but there are just as surely ways in which normative discourse will shape those non-normative ones. (We spoke in Chap. 2 about how normative discourse necessarily permeates our descriptive and other projects that are not explicitly normative; we will revisit this theme in Sect. 9.2.) Symbiosis will be the rule for normative and non-normative discourse regions, rather than the exception. Thus, we will show that normative discourse is not an isolated island, cut off from the main body of theory, but a piece of the mainland, firmly connected to other parts of our theory by numerous contributory strands.

We should add in passing that one area in which there will be substantial cross-discourse contributions would be our accounts of intentionality and the mental. We are sympathetic to the view that many dimensions of our ascriptions of intentionality are irreducibly normative, while also seeing psychology and neuroscience as important contributors to our understanding of such ascriptions. We considered devoting space in this chapter to this point, but it quickly became apparent to us that doing the matter justice would take far more than that, and dramatically expand our project. So for the time being, we have set discussion of intentionality aside, with the thought that extensions to include it would be possible given the resources we have offered, and that we might undertake such extensions in the future.

9.1   Non-Normative Disciplines Contributing to Normative Ethics

None of the preceding should be surprising; as we have noted, the normative is, at bottom, practical, and the practical is about our engagement with the world—how we cope with the world, with artifacts, with other people, with institutions, and so on. So the way the world is must have a fundamental bearing on the shape of our practices, and on the norms that we owe allegiance to in carrying on these practices. The task that lies before us, then, is to elaborate various ways in which that relevance (of the non-normative to the normative) can be manifested. We do not claim that the examples we give will exhaust all of the various ways in which the non-normative can contribute to the normative. But we do hope to flesh out some of the field of possibilities.

Let us begin with some simple and obvious ways in which the empirical bears upon the normative. We are embodied creatures, and our practices are essentially engaged in the physical world, and so our moral system will have to take into account observations about what is happening in the world and the effects of people's actions ("My action is causing Jeff pain," "Bobby hit Sue," etc.). These are uncontroversial: that my action is causing Jeff pain is a prima facie reason to discontinue performing this action; that Bobby hit Sue is relevant to determining whether Bobby ought to be punished. Although these kinds of relations between the empirical and the moral are commonplace, they are nevertheless complex. They are caught up in a network of interests (our interest in avoiding pain, in inculcating good behavior, etc.). These connections are always defeasible, in that there is never a *necessary* connection between the empirical observation and any particular correct course of action (perhaps I am filling Jeff's cavity; perhaps Sue is a crazed ax murderer; etc.). The defeasibility of these empirical-normative connections merely demonstrates the ways in which various interests, as well as a variety of other non-normative features, combine to prevent a given moral situation from yielding to any straightforward analysis in terms of entailment. This is one reason, as we noted at the end of Chap. 5, we are not tempted to embrace any kind of foundational account of the role of observation in normative discourse.

Thus, various facts about how the world is will contribute in obvious ways to the normative claims we are entitled to make, but even these obvious ways can be complex and subtle.

9.1.1 EJ-Contribution and Everyday Non-Normative Discourse

What kind of contribution is the non-normative world making to normative discourse here? Let us start with a look at everyday empirical descriptions before moving on to more substantive sorts of theorizing. We have argued at length that there is no relation of reduction or constitution. As we argued in Chap. 3, if you consider all of the various non-normative properties that instantiate a circumstance we deem "unfair," it is exceedingly implausible that all of these things have something in common at the natural level. It is not even terribly plausible that there is a relation of token identity. If, for example, Johnny is close friends with Sue, but betrays Sue's trust in a way that causes Sue pain, it is not the pain that is cruel (although that is certainly part of the story)—it is the *betrayal* that is cruel. And that is a much more complicated story than one of a moral property simply being token-identical with some unpleasant state being experienced by Sue. (It is even less plausible that we could identify some feature of large systems or institutions, or global economic arrangements, that is token-identical with a moral property such as "fair" or "unjust.")

Nor is the relation between the non-normative and the normative one of causal explanation. When we give a normative explanation, we explain how an action or feature alters or contributes to a particular normative characterization of an action, person, or object. In a moral case, we might say, "That was cruel because it caused her pain." Her pain does not *cause* cruelty; nor does the cause of her pain cause cruelty. The same goes for epistemic relations: factual evidence may cause a belief, but it does not cause the justification. So what kind of relation obtains between the natural and the normative in the kinds of cases we have discussed here?

We hold that it is a particular type of inferential relation, an *explanatory-justificatory (EJ) contribution*. The term "explain" is crucially ambiguous in a way that is important for our project, as it is ambiguous

between a descriptive and a normative reading. We provisionally ceded some ground in earlier chapters to those who would insist on calling only causal explanations genuinely explanatory, but we will retake some of that ground now. In one common sense of the word, the explanatory relation is purely a descriptive one. Thus, we can *explain* someone's belief by delineating its etiology, while in no way endorsing this belief, or even committing ourselves to its rationality. But there is another kind of explanation that is normative in character. For instance, one can say that the wrongness of slavery is explained by (among many other things) the suffering it causes to its victims. This sort of distinction was at least alluded to above, when we separated out normative explanations from other sorts of explanations (e.g., causal ones). In a normative explanation like this, the empirical facts do more than just explain (descriptively) why slavery has the moral status that it does. They serve in the context of a moral practice as premises to justify a particular kind of normative stance—one of condemning slavery—which itself has various practical and inferential consequences. Thus, we can distinguish this particular type of explanatory relation from purely descriptive explanatory relations by denoting the former as EJ-contributions, to note their special normative character.

In introducing EJ-contributions, we are not introducing a kind of inferential relation that is odd or unfamiliar. Indeed, EJ-contributions seem to share major important characteristics with inferential patterns in non-normative disciplines. For example, EJ-contributions are defeasible. Thus, while one might normally infer "A was wrong" from "A caused Jeff pain," one also might not, as when A is "giving Jeff a root canal which he requested and which he badly needs." Similar comments apply to EJ-contributions between descriptive statements and epistemic statements: "The subjects in Smith's study demonstrate a correlation between obesity and hair color with $p < .05$" might justify our believing that there was a causal connection at work, but this justification could be defeated by learning that the sample was unrepresentative. The defeasibility of these inferential relations is in large part a product of the holistic dependence of such inferences on a network of theoretical relations. Thus, although pain is prima facie wrong-making, moral permissibility is not merely a function of this, but bears complicated relations to consent and autonomy, perceived benefits to the agents, and a raft of other considerations

which can outweigh or simply negate this wrong-making valence of pain. Similar comments apply to non-normative material inference. A material inference such as "The match is struck; therefore, the match will light" necessarily depends on a host of background hypotheses about the role of oxygen in combustion (and the presence of oxygen in the environment, etc.). The reliance of the inference on these background theories and hypotheses makes the inference subject to defeat from a number of directions. Again, then, we see broad structural similarities between non-normative inference and the EJ-contribution.

It is tempting to say that empirical facts only EJ-contribute to normative statements in the context of a normative practice. In Chap. 6, we argued that while the truth of any normative claim cannot be reduced to the consensus of a community, normative statuses are in the first instance instituted by social practices, and would not exist in their absence. Does this distinguish EJ-contributions from non-normative inference? We hold that it does not. For surely, any kind of inferential practice—even one whose subject matter is descriptive—must be socially instituted. Quine taught us that no isolated descriptive statement has any observational consequences taken in isolation from the rest of theory. The lesson is much more general: No descriptive statement has any inferential consequences *at all* taken in isolation from the rest of theory. Thus, to infer anything at all from any descriptive statement, one must possess a body of theory. But a theory is not something which is whispered into our ears by the mouth of God. It is something that develops from a practice of inquiry, investigation, and revision over time. In short, it is something that has its life in a particular kind of theoretical practice, a practice that is carried on by a community, governed by norms that one can only make sense of communally. Thus, only in the context of a social practice can one make even descriptive inferences. Indeed, only in the context of a social practice, do descriptive statements even have inferential consequences (and hence semantic contents) in the first place.

One might argue that we are being unfair to descriptive inference, and papering over a legitimate difference between descriptive inference and EJ-contributions. Normative statuses like *cruel* are instituted by social practices, and would not exist without them. Thus (goes the objection), without the social practice, these relations described by normative

inference—between, say, pain and cruelty—would not even exist in the first place. By contrast, the subject matter of a descriptive inference is not dependent on social practice at all: if H_2O, at 101.325 kPa of pressure, is cooled to 0 °C, it will shift from liquid to solid. Thus, the basis of the corresponding inference (inferring "The H_2O will turn from liquid to solid" from "The H_2O is cooled to 0 °C at 101.325 kPa of pressure") does not depend on a social practice, but has (as it were) independent existence.

While the above may be true, it does not point to any structural difference between non-normative and normative *inference*. It is certainly true that H_2O turning from liquid to solid may *causally* follow from its being cooled to a certain temperature, at a certain pressure. But this shows nothing about what is required for an agent to *infer*, from her knowledge that a particular cause has occurred, that a particular effect will happen. Our ability to make causal inferences is no different from our ability to make any other inference—and as we noted above—one can only infer one descriptive statement from another in the context of a whole body of theory, which depends on a practice of inquiry. Inference of any kind is an activity that is itself governed by norms. One can infer well or badly, correctly or incorrectly, and we cannot make sense of the notion of inferring (as opposed to the mere association of ideas, without committing ourselves to their being any logical or rational relations among these ideas) without a background of social practice to ground the normativity of these inferential processes.

So far, our examples of EJ-contribution from non-normative to normative discourse have been simple examples of empirical inputs (e.g., *that my action is causing Jeff pain*, *that Bobby hit Sue*, etc.) to normative statements drawn from everyday discourse. The examples are potentially misleading by their very simplicity. First, such examples (if these are understood as paradigm instances of such contribution) might tempt one into thinking that the contribution of the empirical to the normative followed a certain, simple model on which all such contribution amounted to a set of details that we then apply a set of principles to. That is, we have a set of principles that operate as major premises (e.g., "Causing pain is *prima facie* wrong"), and empirical details serve as minor premises ("My action is causing Jeff pain") which then yield normative conclusions ("My action is wrong"). This view is objectionable for a number

of reasons, most notably that it suggests that a system of norms can be understood in *isolation* from the world, and then *applied* to the world. But this was precisely the view we saw (at the beginning of Chap. 5) was untenable. It results from the twin errors of dualism and intellectualism, and understates the fundamentally practical, world-engaged nature of the normative. To be sure, the above are legitimate examples of empirical contributions to the normative—but to take these as paradigm instances, or as the only legitimate kind of contribution, is to miss out on myriad ways in which the empirical and normative are fundamentally enmeshed.

Another comment is necessary at this point. We take it that normative discourse serves to express these norms implicit in practice in order to make them subject to defense, revision, and so on. Thus, in the first instance, the empirical bears on normative statuses which are *already implicit* in practice. That Sue's action caused Jeff pain EJ-contributes to the status of Sue's action as wrong, and thus also changes other people's status with respect to Sue (say, they might become entitled to chastise her). But given the expressive role of normative discourse, it follows that an empirical fact that EJ-contributes to a particular normative status (such as the wrongness of Sue's action) can *also* contribute to someone's entitlement to make an explicit claim about such a status (e.g., "Sue's action was wrong", as uttered by Paul). This is true even if a normative status is not recognized within the practice, but we are arguing that it should be (and so we are, in essence, calling for revision of the practice): that slavery is cruel EJ-contributes to the wrongness of slavery, and hence entitles an abolitionist (in a slave-holding society) to claim, "Slavery is wrong," even if the practice does not generally accord slave-holding this status, and the truth of this explicit claim is not generally recognized in the society. This is a relatively minor point, but we should emphasize the priority of practice in the order of explanation when it comes to the normative.

9.1.2   EJ-Contribution and the Sciences

Moving on, we can look to the ways that more substantive empirical theories EJ-contribute to normative discourse. We note that epistemology

is an area in which EJ-contributions from psychology and the social sciences are easy to find. In general, states of the world play a large role in the justificatory story for our beliefs, particularly our beliefs regarding descriptive subject matter. Some of our examples in Chap. 8 could be reintroduced here. The systematic link between economic conditions of poverty and long-term effects on cognitive development described in Sect. 8.2 might EJ-contribute as empirical premises to arguments against various policies and practices, for instance. Demonstrating a causal pathway between some of our coordinated social activities and their empirical outcomes offers a compelling reason to alter the activities when the outcomes are undesirable. It does not determine *which* strategies to take up in order the remediate the effects, and reflective equilibrium will require us to accept some tradeoffs of undesirable outcomes in most cases. But deliberating "what justice demands" here can and should be informed by systematic empirical theorizing about the large-scale effects of our practices, which we can draw in various ways from different scientific theories.

   Many moral disputes turn on factual disputes in these ways, rather than disagreement about distinctly moral premises. We would argue that a lack of attention to empirical results often discredits many arguments in which normative ethics and the social science should be working hand-in-hand. Consider the debate over euthanasia. While many bioethicists express their general concerns in terms of sanctity of life, the focus of debate frequently turns sharply to the vulnerability of those who are not in a position to protect themselves. For example, if voluntary active euthanasia is allowed, will this inevitably lead doctors to begin practicing non-voluntary active euthanasia? (That is, practicing euthanasia on patients who are not in a position to give or deny consent, but for whom doctors think euthanasia would not be a harm.) Such questions already admit of some degree of empirical resolution. Voluntary active euthanasia is permitted in several countries, and detailed records are kept on these sorts of cases. So one can investigate whether or not these worries have been borne out (indeed, such studies have been carried out), and which measures are effective in inhibiting such drifts in practice. To put the point more generally, the sort of systematic causal link between policy and outcomes invoked here is where our best theories about such effects are most pertinent, and their very absence renders any normative theorizing suspect. Similar concerns arise about adop-

tion of children by same-sex parents. To the extent that one's opposition to such adoption is rooted in the belief that children need both a mother and a father to grow up fully psychologically healthy, this opposition is based upon a *factual* belief—one that can (and has) been put to the test, given that there exist a large number of children who have been raised by same-sex parents, and who are available for study. Notice that this is a close cousin to our earlier examples of everyday particular facts about the non-normative world being appealed to license particular moral claims. Thus, "Studies show that children raised by same-sex couples do not differ in their psychological health from children raised by comparable opposite-sex couples" is a fact appealed to as part of our justification for the normative claim, "Same-sex adoption ought to be allowed."[1] As we noted in Chap. 5, notions like "health" in general, and "psychological health" in particular, are concepts that are not entirely devoid of normative content; so this might not look like an example of a factual claim being used in support of a normative claim. But even if psychological health is a normatively loaded notion, the judgment that different children meet this standard to an equivalent degree will still predicated upon numerous *factual* judgments.

Moving to a more complicated connection between factual discourse and normative discourse, we have been arguing that goods and interests are essential to any account of moral or epistemic discourse. As we argued at length in Chap. 5, interests essentially depend on (but are not reducible to) facts about our physical makeup and environment. Thus, whether someone is healthy will depend on a variety of social and normative factors (does the person's social situation require him to be able to work in the field all day? etc.), and obviously various facts about one's environment and physical makeup will be relevant to one's health (e.g., that one has been kicked in the ribs by a mule, and that one's ribs are not mule-hoof-proof, are physical facts that explain why one is not healthy enough to work in the field this week). But these physical facts are embedded in a set of practices and expectations, and hence serve as part of an EJ-contribution for why Jim is not healthy enough to work in the field this week: broken ribs, in light of certain *demands* and *expectations*, are seen as *excusing*, whereas (say) a splinter or a bruise would not, and a person who tried to beg off field work on such grounds would be seen as shirking.

Consider a more complicated way in which the empirical can bear on the larger terrain of moral inquiry. To cite a recent example, where much work in bioethics has been driven by empirical challenges to traditional conceptions of personhood and its biological basis. The subfield emerges in large part because certain fundamental components of the normative theory can be upended by theoretical and technological change. For centuries, criteria of death had been articulated in terms of the cessation of certain crucial biological functions, particularly heartbeat and respiration. However, by the mid-twentieth century, various technological innovations arose which required a re-evaluation of what it means to be dead. Machinery was developed to keep patients alive long after they ceased being able to respirate independently. The ability to transplant organs made it necessary to sustain, on life support, patients who otherwise would have ceased all bodily functions,[2] until decisions about transplant can be made and the transplant team can be mobilized. Electroencephalography developed and improved to a degree that previously undetectable levels of neural activity could be measured, opening questions of whether the flickers of a mental life and promise of a return to some level of normal function might remain. Thus, questions arose as to whether a patient, whose brain had little or no function, but who could be artificially respirated for some indefinite period of time, was still "alive." Empirical inquiry and new technology spawned new debates about the endpoint of human life, and the ethics of ending that life. Questions that might have been treated as unremarkable extensions of our general obligations (e.g., beneficence) instead developed into novel extensions of normative inquiry with a life of their own (no pun intended). Bioethics is just one field, but it provides an example of the way that a normative discipline can be altered and informed by contributions from scientific discourse regions.

This may strike some readers as very broad in comparison with the detailed contribution relations described in Chap. 8, so let us move on to some more concrete examples in which different empirical disciplines can contribute to and constrain normative ethics. One type of result that has received much attention in recent years is the situationist challenge to virtue ethics.[3] According to many versions of virtue ethics, virtues are stable dispositions to ethical behavior, and thus the sort of substan-

tive properties we discussed in Chap. 3. But do we have such stable dispositions? Situationism denies that we do. Doris (1998) reports on the results of a study (Isen and Levin 1972) which seems to show that our willingness to help others is conditioned by seemingly irrelevant situational influences. In this study, an experimental confederate drops a folder full of papers, scattering them everywhere, just as a caller leaves a phone booth. Virtue ethics tells us that those callers who are kind and caring will strongly tend to help, whereas those who are selfish and indifferent will not. As it turns out, the strongest determinant of whether the caller will help pick up the scattered papers or not was whether the experimenters left a dime in the coin return slot of the phone before the caller entered the phone booth. (Thus, some subjects' affective states were primed positively by the minor windfall.) When they did so, 14 subjects helped (vs. 2 who did not help), whereas when there was no dime in the return slot, only 1 caller helped (vs. 24 (!) who did not help). Other research has shown that various other factors such as the presence of bystanders (Schwartz and Gottlieb 1957), ambient light levels (Zhong et al. 2010), and so on, affect people's moral behavior in ways that are difficult for the virtue theorist to account for. Much work in experimental philosophy in recent years has also extended this sort of empirical contribution, looking more closely at how biases (such as order effects) alter our judgments situationally. It is open to the virtue theorist to attempt to refute these attacks (as they have attempted to do), and in the final analysis, we might not find these attacks convincing. But the moral of the story is that whether our moral theory is built in terms of the right or the good, in terms of rules or character, ultimately, we must be able to give an account of how it is that agents can act in accordance with its dictates. Every moral theory is committed to moral agents being embodied creatures that can act in accordance with the normative element of the theory. Thus, if our moral theory has certain consequences for moral psychology (or certain facts about moral psychology must be true for our moral theory to work), and empirical investigation tells us that this account of moral psychology simply cannot work, then this is at least prima facie evidence that our moral theory is not in order as it stands. The interests that we pursue with normative ethics are constrained by empirical conditions, and we should expect other

explanatory accounts of those conditions to contribute to that work. Conflict or incompatibility here would be a mark against a normative theory, and well-established psychological accounts would inform future directions for normative theory, even if they are not replacements for it.

This type of relation between the psychological and the normative might seem different from the sort of relation between the natural and the normative we have been exploring up until now, which might make it seem as though this is a different type of relation from the EJ relation we have been discussing so far. This is unsurprising, really. A common theme of our anti-reductionist argument (especially in Chap. 8) has been that a reason you cannot reduce certain types of discourse to other types of discourse (like the biological to the chemical) is not that there is not a relation between the former and the latter. The reason is that there are too many relations, and reduction can only "succeed" by privileging one while ignoring others. It is an inherently Procrustean solution. Thus, just as the relation of "contribution" can be instantiated in many different ways, EJ-contribution itself admits of multiple types of instantiation, from the relatively simple (that A caused her pain is evidence for A's wrongness) to the relatively more complex (that moral theory T is underlain by a faulty moral psychology M is evidence against T; that a cogent moral psychology M' supports some T' is evidence for T'). But the important thing that these relations have in common is that they all allow us to *justify* a particular normative stance (we should reject this normative theory; we should condemn A, and adopt other ancillary positions attendant to condemning A, etc.), while at the same time offering a partial explanation for the adoption of this stance. Thus, these multifarious relations of contribution are all EJ relations, at the end of the day.

Consider another example of empirical investigation constraining the structure of moral theory. Utilitarianism enjoins us to maximize the good. But the familiar question is, "How should we measure utility?" Familiar answers are that we should measure utility in terms of desire- or preference satisfaction. But there are empirical reasons for thinking that neither of these is a good measure of well-being, and to the extent that we think that utilitarian theories ought to measure well-being and not just some subjective quality like "happiness," we should reject these accounts of utility.

The example is complicated by the interplay of empirical and normative elements. On the empirical side, economists and other social scientists have documented a phenomenon known as "adaptive preferences," whereby an individual's satisfaction with her situation is conditioned by what she has come to expect. Writing on this phenomenon, Martha Nussbaum (2000, 139) notes that according to a 1944 poll, 45.6% of widowers in India rated their health as ill or indifferent, but only 2.5% of widows did so. This result is despite the fact that widows were, as a group, in manifestly worse health than the widowers. Since they had come to expect poor health, they came to accept their poor health as satisfactory. According to Nussbaum, the problem of adaptive preferences presents a serious difficulty for utilitarian accounts of welfare. Amartya Sen echoes this sentiment, writing, "Quiet acceptance of deprivation and bad fate affects the scale of dissatisfaction generated, and the utilitarian calculus gives sanctity to that distortion" (1984, 309).

On the normative side, a utilitarian can simply bite the bullet and say that one way to maximize utility is to lower people's expectations, and take advantage of adaptive preferences. But if you think utility should measure objective well-being, rather than something more subjective (like happiness, construed as a subjective state), then this empirical result demonstrates that certain accounts of utility are inadequate. But notice here that the empirical result only refutes certain accounts of utility when combined with a separate normative premise (that subjective accounts of utility are inadequate). This normative premise can be established independently—perhaps by Nozick's experience machine, or some similar argument. But this is a different, more complex, way in which the empirical can contribute to our normative theorizing.

We have primarily been discussing various ways in which empirical discourse can contribute to moral discourse, but similar contributions can be made to other types of normative discourse. If we discover a perceptual phenomenon such as simultaneous contrast (that how a color appears to us is affected by other colors that are adjacent to it), such an empirical finding might be problematic for various versions of sense datum theory, for example. Just as morality is tied to action (and so, as we discussed above, implies facts about human psychology which may or may not be true, on investigation), epistemology is tied to belief; and so

if an epistemic theory requires or entails norms which empirical findings in cognitive psychology tells us humans cannot follow, then (on the plausible assumption that "ought" implies "can"), such a theory would need to be revised. (To illustrate this point, the coherentist theory of epistemic justification has been criticized on the grounds that humans do not have the intellectual ability to uncover incoherence in their belief set, as an entire belief set is far too complicated to investigate in even the most cursory way. Many empirical findings on, for instance, working memory would support this criticism.) Epistemology presupposes, just as morality does, that we are agents engaged with the world in fundamental ways (or else what point would morality and epistemology have?), and so to the extent that these theories make false assumptions or impossible demands about this engagement (and to the extent that empirical investigation can reveal these infelicities), there will be numerous pathways for discourse regions reporting empirical facts to EJ-contribute to normative discourse, in epistemology just as in morality.

9.2   How Normativity Contributes to Metatheoretical Discourse

One sort of objection against normative discourse imagines it as isolated from non-normative discourse, not accepting non-normative inputs, floating free and unconstrained. We argued, in the first half of Chap. 5, that normative discourse as so conceived is, in fact, impossible, and this conception represents a fundamental misunderstanding of the nature of normativity, and how normativity arises out of our practical engagement with the world. We have also shown, in Sect. 9.1, various ways in which the non-normative can contribute to normative discourse. Thus, even given our non-descriptive account of normative discourse, the normative is firmly anchored in the non-normative.

The more pertinent objections might be from philosophers of science who contend that there could be scientific practice conducted by beings without normative entanglements at all. Thus, this objection goes in roughly the opposite direction: While the worry of Sect. 9.1 (and Sects. 5.1 and 5.2) was that our normative practice could carry on without

non-normative inputs, the worry here is that our non-normative practices could carry on without any normative constraints. By laying to rest this objection, we can show that the normative is firmly entrenched with the non-normative: The two discourse regions are solidly mutually interdependent in a way that should show, for once and for all, that our normative discourse exhibits myriad connections to our non-normative theories, and is not an isolated theoretical island.

To a significant degree, this sort of case for normative contributions to non-normative discourse regions has already been made in this book. We argued in Chap. 2 that robustly normative epistemic claims were essential to the exercise of even the most fundamental forms of scientific inquiry; we can now see that our earlier argument was about a type of contribution between discourses. The hope that we could have a scientific practice without normative entanglement or engagement is, in essence, the hope that we could have a pure theory of the world, pure representation, without the involvement of characteristic human interests. While we do not want to assert, offhand, that such a theory is in principle impossible, we do want to deny that such a theory would be an extension of *our* theories or projects of inquiry. There are at least two reasons for this. First, as we argued at some length in Sect. 2.2, scientific inquiry requires for its very conduct the existence of epistemic (which Putnam calls "methodological") norms. Putnam points out, for example, that a number of "value judgments…are internal to scientific inquiry itself: judgments of *coherence*, *simplicity*, *plausibility*, and the like" (2004, 67). We argued that scientific theory involves implicit adherence to epistemic norms: It involves evaluation of certain methods as entitlement-conferring, certain theories as rational or credible (and others as not), certain figures as authoritative (and therefore their judgments within a certain field as entitlement-conferring), and so on. Further, even to recognize the authority of science in the first place is a normative judgment, and requires privileging some forms of inquiry over others (and, indeed, requires a thoroughgoing commitment to being truth-seeking, which as we also saw is a normative attitude). Thus, as we saw throughout Sect. 2.2, to conduct oneself as an inquirer is to take a stance that is thoroughly imbued with normative significance. So we will not rehearse these arguments at length here. Instead, we will devote the rest of this chapter to further ways in which concern

for interests and the normative discourse that articulates them could play a role in shaping theoretical inquiry.

9.2.1 Interests and Metatheoretical Discourse

In imagining the gathering of information divorced from norms and interests, one is not imagining an ideal theory, or an ideal researcher—one is imagining simply an ideal storage device. For inquiry, not guided by interests or norms, is blind. Simply gathering a greater and greater number of truths would not count as a success; what counts as a success would be gathering a greater and greater number of relevant, interesting, and/or important truths. If inquiry is not guided by interests, we are left with no sense of what would make for a *theory*, rather than a random agglomeration of information and idle correlations. To explain and understand the world is to explore it in a purposive way, in a way that is guided by considerations of what is relevant to us—our needs, wants, the way we are embodied, our particular point in our historical narrative, our environment (both physical and social), and so on. After all, what makes the Spurious Correlations website (which, despite its name, generally records actual correlations, but spurious *relationships*) entertainment, rather than an interesting source of theoretical knowledge? For example, why is not it theoretically interesting that, between 2000 and 2009, the divorce rate in Maine was strongly correlated (correlation = 0.992558) with per capita consumption of margarine in the USA? (Vigen n.d.) To be sure, this is a mere correlation, and there is no causal relation between the two trends. But the fact of their correlation is just that: it is a *fact*; it is *true*. However, it is clearly an *uninteresting* (if amusing) truth. We can record it, and innumerably many other reports of empirical details, but information alone, of any sort, is idle in the absence of guiding interests for a theoretical project. (The amusement comes from the incongruity of such a strong correlation that is so devoid of import, we would say.) We can simply collect all the information we like, but without having interests to drive and guide us, we would not even know where to begin our inquiry. We would not know which questions were worth investigating, which questions were important. Why not investigate potential cor-

relations between margarine consumption and divorce rates rather than investigating something else instead?

If, by contrast, we found a correlation between divorce rates and unemployment rates, or the ages of the spouses at the time of their marriage, or another familiar economic or psychological variable (and we could rule out confounding variables to establish causation), then we would have a piece of information that was of value. The significance would lie in this explanation's unifying different elements that each has some type of practical urgency for beings like us, at least in our current social settings. But, again, the fact that this piece of information is of value can only be understood against the background of our interests and norms—our marital norms, our interests in promoting healthy marriages, and so on. (In this sense, the margarine example above could prove worthy of inquiry, if the link proved to be causal and its effects tapped into our interests in having healthy relationships.) The value of this piece of information cannot be explained without this additional background.

We can see, then, that even if it is possible to think of a practice of inquiry not guided by a metatheory—not guided by norms related to our interests—this practice of inquiry would not in any recognizable way be an extension of our theory of inquiry. Nor would it be an idealization of our theory of inquiry. For the vast majority of facts which are available are simply irrelevant. Indeed, there are an infinite number of facts one can seek to uncover, and a correspondingly uncountable number of potential lines of inquiry one can pursue. The choice of which line of inquiry to pursue (should I attempt to determine how many atoms are in this grain of sand, or should I attempt to discover what is predating the endangered turtle eggs on this beach?) must always be decided by which questions are relevant, important, non-trivial, and so on. This is always a question of values and interests that cannot be decided in a value-free way. If one's inquiry were not guided by any interests, one would not have any reason to develop a *theory* of anything. One might acquire a fact here, and a fact there, but not be concerned about the relation *among* such facts.

But it is our interests that drive us to develop comprehensive constellations of interlocking claims, practices, techniques, and models that enable us to understand (and perhaps also manipulate) some aspect of the world. For example, consider the history of iron production. At least

6000 years ago, humans discovered how to fashion weapons and other artifacts from iron meteorites. The discovery of how to extract iron from ore was discovered much later. Iron had various deficiencies as a material, particularly in the production of weapons: it was brittle, but insufficiently hard, and thus unable to maintain a sharp edge. Iron is also subject to corrosion. Though lacking a sophisticated knowledge of chemistry, metalworkers discovered (no doubt by trial and error) various forging techniques by which iron could be made both stronger and harder. Tempering this forged iron could further increase its toughness, strength, and hardness. With advances in both chemistry and metallurgy, modern steelmakers are now able precisely to adjust the carbon content of steel to achieve the desired properties, whether they want steel that is cheaper and more flexible (lower carbon content) or stronger and capable of greater hardening (higher carbon content). Also, other alloys (e.g., chromium) can be added in various proportions to achieve other desirable qualities (e.g., corrosion-resistance). Such intensive investigation into the properties of iron—how to forge it, how to refine it, how to temper it, and how its various qualities (hardness, toughness, strength, etc.) are affected by various processes and techniques, the effects of alloying iron with other compounds, and so on—only makes sense given the understanding that iron is useful for various purposes, and that this investigation all serves the purpose of allowing us to utilize this material for making weapons, tools, machines, houses and buildings, jewelry, cookware, and countless other artifacts which themselves serve manifold other interests. Interests serve as the point around which theories coalesce. They are the reason we have, for example, a theory of metallurgy, why we have a theory of different methods of tempering (tempering to different temperatures for different purposes, differential tempering, etc.), a theory of various methods of forging (hot vs. cold forging, drop forging, open forging, die forging, etc.).

We can see, again, the relation between discourse regions and interests. Various phenomena are unified into a theory not merely because of considerations of explanatory coherence and unity. The entire point of developing a theory in the first place is that some purposes are served by explaining this aspect of the world, and at this particular level of detail. And an explanation of some aspect of the world means developing not

just a set of isolated facts, but developing a theory—a set of interrelated, mutually-supporting and -explaining claims, practices, models, and techniques. Some purposes are served by having a theory of chemistry, giving us an account of these phenomena, at that level of explanation. But there is also a set of purposes served by having an account of some aspects of the world at a different level of explanation—the level served by physics, for instance. And also biology. And economics, and psychology, and so on. But again, if these accounts were not theories, they would not explain, and would not augment our understanding. Thus, there is a fundamental connection between interests and explanation/understanding. Interests provide a direction for inquiry, a center around which collection of claims and practices can coalesce into theories. Inquiry not guided by values and interests would not result in theories, but merely, as we have said, an agglomeration of facts, with no way to sort out which ones are important, or which connections among facts are important.

How, precisely, do interests structure the practice of inquiry? We must be careful not to conflate different ways in which inquiry and value are interrelated. We have already discussed one such relation: inquiry is guided by methodological or epistemological norms. These norms relate to theory choice, justification of belief, and so on, as we also discussed at length in Sect. 2.2. But the issue at hand in this subsection directs us to a different set of questions about the relation between value and inquiry. What makes a particular fact, or line of inquiry, worth pursuing or worth knowing in the first place? And how does this relation shape theory formation? As it turns out, in addition to the ways in which inquiry is guided by methodological or epistemic norms, there are at least three distinct ways in which interests contribute to the guidance of inquiry.

9.2.2   Interests and the Shaping of Theories

One set of disciplines in which it is easy to see the sorts of contributions described above from the normative to the non-normative is the social sciences. What is considered an important question is a question of normative import; and these normative questions serve, in turn, as points around which theories coalesce, as a grain of sand forms the nucleus for

the accretion of matter in the formation of a pearl. Thus, this first type of contribution is really two intimately related types. First, our practical concerns (including our moral concerns) give direction to our inquiry in the social sciences. Without these concerns, as noted above, we would have no conception of what line of inquiry to pursue, of what was important. Again, there are an infinite number of lines of inquiry one *can* pursue; without guidance by practical considerations, there is no way to choose among them the most fruitful and constructive lines of inquiry. And this is a juncture at which the contribution of various normative theories, both epistemic and particularly moral, can contribute with explicit statements of how (and why) different focuses of concern in social scientific inquiry are to be prioritized. Second, as illustrated with the example of metallurgy above, these interests serve as points around which theories coalesce. Since they drive inquiry in a practical direction—toward the solution to a particular practical problem, or the answer to a particular question—these interests will thus require not just the gathering of facts, but purposive inquiry. They will require the gathering of facts and development of techniques that illuminate and support each other with the end of answering this question, or solving this problem. Thus, because inquiry is interest-driven, inquiry results in *theories*—bodies of explanatorily related facts, models, techniques, and practices which provide understanding and perhaps also the ability to control some element of nature.

Consider an example from the social sciences regarding the first point. Economics, as a discipline, has both applied and theoretical orientations. (Even insofar as economics is a purely theoretical discipline, it is interest-driven. We will say more about this shortly.) But what drives particular lines of inquiry in applied economics is driven by practical, and again, often moral, concerns. For example, why study poverty, when doing economics? The reason is obvious—but it is obvious to *us*, with our characteristic interests and concerns. We are concerned with human well-being, flourishing, and autonomy, and perhaps also with justice and equity (or perhaps even just with efficient use of social capital), and we realize that poverty can impede all of these important values. Thus, poverty is something we care about; it is an area of human concern; and it is an area that appropriately drives research in some subfields of economics. Thus, these researchers might ask, "What causes

poverty to persist over time? What are key causes of social mobility? What effect does social inequality or poverty have more generally on a nation's or community's economic growth? What policies address or fail to address poverty?" Thus, we see the first sort of relation mentioned above: a particular set of normative concerns points inquiry in a certain direction, toward the understanding and resolution of certain issues or problems.

But this example also illustrates the second sort of relation: our interests serve as a point around which theories coalesce. For example, in an attempt to give economists the tools to address poverty, various economists (most notably Amartya Sen), dissatisfied with various attempts to measure well-being (such as the utilitarian approach) and with welfare economics more generally, developed the capabilities approach to measuring how well-off people (particularly people in poverty) are. But this account (driven by an interest in fighting poverty) is a comprehensive theory, giving not just an account of human well-being, but also containing a criticism of rival accounts (particularly utilitarian accounts). Thus, the need to address a particular interest (How do we measure well-being in the context of fighting poverty?) serves as the nucleus around which a particular theory forms; this illustrates the second facet of this type of normative contribution: not only do interests direct theory, but they also organize inquiry *into* theory.

As we have emphasized, these connections between interests and inquiry are ineliminable. Without interests, there is no way to decide whether to investigate this rather than that; there is nothing to motivate inquiry. But also, there is no theory, nothing coordinating the collection of facts into a project aimed at understanding *this* or controlling *that*. And these two relations between interests and inquiry obtain not merely in the social sciences—as our example of metallurgy displayed above, they also obtain in the hard sciences, where interests not only direct inquiry but also serve as a focal point around which theories coalesce. This is one of the most fundamental, ineliminable, and crucial roles that interests and normative concerns serve in inquiry. It is (aside from methodological constraints) perhaps the most fundamental type of contribution obtaining from the normative to the non-normative.

9.2.3 Epistemic Norms of Counterfactual Robustness and Non-Generalizing Subjects of Inquiry

The value-driven shaping of social scientific inquiry we described in Sect. 9.2.2 may very well be conceded by others thinking about such questions. But, some might object that to make such appeals to interests *so* central to our conception of scientific inquiry is to obscure or omit an even more important dimension of non-normative theoretical inquiry. We often articulate our attention to some questions (or to inquiry in general) in terms of our love of knowledge in general, or as the pursuit of knowledge for its own sake. Work on fundamental questions in, say, cosmology might give us a more profound understanding of the origins of the physical universe. This might serve a concrete interest like those mentioned above in a very indirect way (perhaps a more complete physics would eventually pave the way for new technologies), but it might not, and it would certainly mischaracterize the way in which many forms of inquiry are conducted to impose such interests on it. (Why did we build the LHC? *Because we just want to know how the universe works.*) We would agree that this is true of many forms of inquiry, perhaps even all of them for some questions or to some degree. Thus, to concede that all of our practices are all ultimately pursued for (and justified by) some interest or other is not to concede some crude instrumentalism, on which every practice must have some concrete payoff in terms of improvements in the material welfare of humanity. Our interests are manifold, and many of them do center around more abstract goals like the acquisition of knowledge and understanding for their own sake. But this is still a type of practice driven and shaped by epistemic and moral concerns. To value the exercise of our epistemic potential, to love learning and knowing some things for the sake of learning and knowing them is to take up a stance on who we are and what we should do with ourselves.

One very general way in which there will be such normative contributions to our non-normative discourse is through what we might call norms of counterfactual robustness. These have rarely been made explicit, but they would be almost painfully obvious to anyone doing any form of theoretical inquiry. Some types of results are simply more

epistemically robust: that is, they project across more possible worlds, they generalize to more cases, allow us to make projections or predict future results, and so on; and theory-building inquiry is shaped (often implicitly) by valuing both questions and results that have this sort of projectibility. Thus, for example, information about Sears mail-order catalogs does not generalize or project across possible worlds in the way in which, for example, information about laws of physics does. There are many very close possible worlds (and many time periods within the actual world) in which there are no Sears mail-order catalogs at all, or in which they have very different properties; but again, results about gravitational force, or the charge of electrons and protons, project much better across near possible worlds. It is a superior kind of epistemic result. Thus, whereas the contribution discussed in Sect. 9.2.2 was primarily between practical (including moral) norms and scientific (and social scientific) inquiry, the norm of counterfactual robustness is an epistemic rather than a practical norm. To be sure, if one discovers a law-like regularity, one can exploit this in practical ways. Human technology relies on the discovery and exploitation of such regularities. But what makes, in the first instance, a discovery about the mass of electrons interesting (vs. a discovery of the mass of this year's Sears catalog) is precisely the epistemic status of the former compared to the latter. It is projectible, it is connected in fundamental ways to other various theoretical results (whereas the latter result is an uninteresting theoretical outlier, without interesting explanatory links to other phenomena we are interested in explaining), and so on. That is to say, the importance of the former result is in the first instance an epistemic rather than a practical matter.

Consider interest in the question as to whether *Acanthodes bronni* is really a common ancestor to both humans and sharks,[4] or what the chemical makeup of Titan's oceans is,[5] or whether there is (or was ever) some form of life on Mars. These are not examples of interests directing inquiry in the social sciences, and so do not fall into the type of interest-inquiry relationship described in Sect. 9.2.2. But also, this is not the type of interest-inquiry relationship described at the beginning of Sect. 9.2.3, because there is no norm of counterfactual robustness present here. Indeed, the characteristic feature of these subjects of inquiry

is that the answers are generally not projectible across a wide range of possible worlds: the discovery that the oceans of Titan are made of water ice is not a law-like, projectible result in the same way that (say) the discovery that the elementary charge is $1.60217662 \times 10^{-19}$ coulombs. (The latter fact varies across possible worlds, too, if we admit that there could be logically possible worlds without our laws of nature; but it projects across more possible worlds than does the former fact, and this projection is "non-accidental" in a way that the former projection would not be.) Similarly, while the fact that sharks and humans are both descended from *Acanthodes bronni* is interesting, it is intuitively a far more contingent, and far less projectible, fact than a law-like regularity would be. On its face, this might sound like a devaluing of biology or some social sciences, which essentially involve some historical features; but this is not our claim. What distinguishes these as *theories*, rather than mere *catalogs* of historical details, is their systematic efforts to characterize those historical details by laws or mechanisms (e.g., natural selection) that do exhibit such counterfactual robustness. As a piece of the larger picture of the origin of our species, and thus a window into the origin of *all* species, our relation to *Acanthodes bronni* is of theoretical interest; without norms of counterfactual robustness, it is merely one more piece of information.

We might put the point of this subsection in overtly normative terms by saying that curiosity is a virtue. We wish to know things about the world in which we live, and we want to know our place in the world; and even if various bits of information do not generalize or project, they can still be informative, satisfying, and important. And we would gleefully concur. We should refine our understanding of that commitment just a bit though. For whatever reason, and despite the fact that so much inquiry is motivated by curiosity (and philosophers do seem to care so much about the methodological norms *governing* inquiry, at least), curiosity itself—and specifically, what norms govern what is *worth* being curious about and what is not—has not been a subject of very much philosophical attention at all.[6] This in itself is puzzling, as curiosity—which provides this final link between interest and inquiry—is, in a way, the most distinctly human of motivations for inquiry.

Recent philosophical writing on curiosity has noted that it has not always been regarded as a virtue. Christian writers, in particular, often held that curiosity might lead one to explore areas where the temptation to sin would arise. Augustine, for example, writes that curiosity might drive one to

> search out the secret powers of nature—those which have nothing to do with our destiny—which do not profit us to know about, and concerning which men desire to know only for the sake of knowing. And it is with this same motive of perverted curiosity for knowledge that we consult the magical arts. (1955, 234)

Plausibly, though, curiosity as such—a drive to know and to learn—is more properly treated as a virtue. While one could perhaps have a well-lived life while not indulging in curiosity, or actively seeking out knowledge,[7] it certainly seems like such inquiry is a form of enrichment, and one possible element of a well-lived life. However, even if curiosity is a virtue, not all objects of curiosity are equal. Curiosity may be directed at some objects which are inappropriate. Some objects might be morally inappropriate objects of curiosity. Consider two (of several) examples offered by Neil Manson:

(1) Tom is a nosy, inquisitive, busybody. He hates his neighbor Jerry. He breaks into Jerry's house, reads Jerry's diary, and finds out that Jerry has cancer.
(2) Tom is a nosy, inquisitive, busybody. One hot evening Tom notices that Jerry has left his window open, and is talking in an agitated way on the phone. Tom quickly goes out, stands in his own yard outside the open window, in order to overhear the phone call. He learns from this that Jerry has cancer. (2012, 253)

Inappropriately directed curiosity can (if acted upon) violate a moral obligation to respect people's privacy. (Manson, incidentally, distinguishes between the content of what is learned via curiosity, the process by which it is learned, and the purpose for which it is learned. The interplay between these is complex; a content can be neutral, but the process

by which this content is acquired immoral. Or the process neutral, but the content inappropriate. Or the morality of inquiry may depend on the purpose for which the information is intended.)

Objects of curiosity may also be epistemically inappropriate. They may be trivial, or uninteresting to various degrees. As Manson points out, a person who neglects an important project to satisfy his curiosity about the lives of celebrities has made a poor tradeoff (at least if that pursuit is extended and consumes time and other resources), given the triviality of this object of curiosity. Some subjects of inquiry are more trivial than others, and some are trivial to the point of not being (epistemically) worthy objects of curiosity. To use a well-worn example, we would judge that someone who devotes his day to counting blades of grass on the lawn has wasted his time, because such information simply is not worth having. The chemical composition of Titan's moon (although, in a sense, equally useless, from a practical standpoint) is less trivial. It is not a waste of time to investigate this question.

As we noted, though, neither bit of information is of particular practical use. What makes one a proper object of curiosity, and the other a waste of time? The answer cannot just be that we find one question *subjectively* interesting, and the other not. The inference from "we are interested in X" to "X is interesting" is just as risible as Mill's inference from "we desire X" to "X is desirable." We think there is a principled distinction that can be drawn. Ultimately, the notion of epistemic importance within this third category is tied back to our notion of fallibilism. There is a sense in which being a fallibilist means that one has a standing commitment to try to fail. That is to say, aware of the limitations of one's knowledge, and aware not only that one has much left to learn but also that much of what one thinks we know could in fact turn out to be mistaken, one should always be on the lookout for new ways to put our current body of knowledge to the test. Curiosity is, in large part, a virtue because it compels us to do precisely this—it compels us to go investigate, increase our store of knowledge, and in doing so, put our existing knowledge (whether wittingly or unwittingly) to the test. Thus, for example, the measurements of the radial velocities of galaxies performed by Vesto Slipher in the early twentieth century ended up providing the empirical evidence to support the hypothesis of the expanding universe. Investigating the orbits of the

planets (using the tools of Keplerian astronomy and Newtonian physics) led to the discovery of a perturbation in the orbit of Uranus which, as is well-known, allowed for the discovery of Neptune (the only planet whose location was mathematically predicted before it was directly observed). Stumbling upon something new while investigating something different is a familiar way of progressing in science.

Here, then, is where curiosity intersects with fallibilism. Our pursuit of a wide range of varied research projects is, in an important sense, a counterbalance to the sort of epistemological conservatism we advocated earlier in the book (such as in Chap. 1). Conservatism, we have argued, is a necessary feature of any epistemological system, but it has its risks, and it can be carried too far; we noted that a necessary counterpart to conservatism is fallibilism, and a diachronic conception of rationality. We must be willing to treat these "prejudices" as at least de jure subject to revision. This shifts the focus of any investigation into the nature of rationality not from the arbitrary starting point of investigation, but to the nature of how beliefs systems are revised over time. On this conception of rationality, curiosity becomes a critical epistemic virtue. Curiosity for curiosity's sake is a crucial counterbalance to epistemic conservatism. As an epistemic norm, curiosity bids us to investigate the world, and in doing so, we might add to our knowledge, or upset previously settled areas. So what, then, makes the composition of the moons of Jupiter and Saturn a more worthy subject of inquiry than (say) the number of blades of grass on the courthouse lawn?

The number of blades of grass on the courthouse lawn is a question that involves an arbitrarily staked out plot of land, a social kind whose boundaries do not coincide with any interesting graminaceous facts about the world. Thus, the number of blades of grass in this arbitrary space is not going to tell us anything interesting about the biology of grass, or about anything recognizable as an ecosystem, or even anything interesting about grass. But learning about, say, the surprising abundance of liquid water in various satellites in our solar system—on Europa, Enceladus, and Ganymede, and perhaps several others—tells us exciting things about (for example) possible locations where life may have developed. It might even suggest that there are more M-class planets in the universe. Thus, it is a more revealing, more interesting result than the blades-of-grass result, which is trivial and has a

strong arbitrary element to it (as determined by the more or less arbitrary size of the courthouse lawn).

This is not to say that pursuit of trivial knowledge never has any place at all, in an individual's life. Manson (2012) presents us the case of Tom (who spends all of his time on social networking sites, catching up on news of his friends), and Mary (who spends all of her waking hours reading celebrity gossip in tabloid magazines). The problem is not that they are engaged in epistemically trivial pursuits, but that they are doing so to the *exclusion* of other (potentially more valuable) activity. As Manson writes,

> If we transform our examples above so that Tom and Mary engage in their trivial epistemic pursuits in their spare time, in order to relax, would we hold that they ought not to do so? Indeed, our conception of a well balanced life is one that has a place for relaxation and idleness. (2012, 251–252)

Plausibly (and we think this is implied by Manson), the value of what Tom and Mary are doing does not really come from the knowledge that they acquire—at least in the case of Mary's fascination with celebrity gossip, it is not really knowledge worth having. (Tom's case is less clear: one should have some concern for what one's friends are up to, but there are also diminishing marginal returns on such inquiry.) But the real value of the activity comes from the fact that Tom and Mary find the activity relaxing and diverting. While we have above rejected the Millian inference from "X is desired" to "X is desirable," it is not entirely implausible that if Mary finds a gossip magazine interesting, then in the absence of countervailing considerations, this does provide her with some (weak) prima facie reason for reading it, just as the fact that one feels like going for a walk might (again, in the absence of countervailing considerations) count as a reason for doing so.

And so in this chapter, we have seen an outline of the sorts of bilateral asymmetric contribution relations that will hold between normative and non-normative discourse regions. Normative discourse regarding morality and epistemic matters will be deficient without importing concrete empirical details that they cannot provide themselves. But non-normative

projects such as scientific theories will be shaped by the interests and standards at play in normative discourse as well. Our case for a more moderate naturalism has thus come full circle. To engage in normative discourse is to find ourselves enmeshed in a natural world and to integrate such self-understanding with our non-normative articulation of that world—in some ways, even to defer to such theoretical projects. But to engage in the explanatory projects that so many naturalists favor is always to bind and be bound by norms and to guide and be guided by interests and a conception of what is important to us. There is no normativity available to us prior to our engagement with the world, and there is no engaging with our world without normativity.

Notes

1. This is a broad conclusion about a great deal of social science, stated in exceedingly simple terms. More precise, detailed claims might serve the same roles, particularly in more circumscribed disputes.
2. For more on these two points, see Randell (2004).
3. Important starting points in the situationism debate would include challenges to virtue ethics from Doris (1998, 2002) and Harman (1999, 2000), with responses from Adams (2006, Chaps. 8–9), Kupperman (2001), Miller (2003), and Sreenivasan (2002, 2013). Upton (2009) also offers a nice overview of the debate.
4. See Davis et al. (2012).
5. Water ice, apparently. Anticlimactic, we know.
6. Inan (2012) provides an in-depth discussion of the philosophy of curiosity, but mostly from the perspective of philosophy of language. Inan is chiefly concerned to develop an account of "inostensible reference" (reference to the unknown) and discuss how curiosity, and its satisfaction, involves a move from inostensive to ostensive reference.
7. It seems like one could lead a good life while being unwilling, or even unable, to partake in some good or range of goods. Has Stephen Hawking, whose health and mobility have been significantly curtailed, not lead a good life? This claim is suspect to us.

References

Adams, Robert Merrihew. 2006. *A Theory of Virtue: Excellence in Being for the Good*. Oxford: Clarendon Press.

Davis, Samuel P., John A. Finarelli, and Michael I. Coates. 2012. Acanthodes and shark-like conditions in the last common ancestor of modern gnathostomes. *Nature* 486: 2447–250.

Doris, John. 1998. Persons, Situations and Virtue Ethics. *Nous* 32: 504–530.

Doris, John. 2002. *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.

Harman, Gilbert. 1999. Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error. *Proceedings of the Aristotelian Society* 99: 315–331.

Harman, Gilbert. 2000. The Nonexistence of Character Traits. *Proceedings of the Aristotelian Society* 100: 223–226.

Inan, Ilhan. 2012. *The Philosophy of Curiosity*. New York: Routledge.

Isen, Alice M., and Paula F. Levin. 1972. The Effect of Feeling Good on Helping: Cookies and Kindness. *Journal of Personality and Social Psychology* 21: 384–388.

Kupperman, Joel. 2001. The Indispensability of Character. *Philosophy* 76: 239–250.

Manson, Neil. 2012. Epistemic Restraint and the Vice of Curiosity. *Philosophy* 87(2): 239–259.

Miller, Christian. 2003. Social Psychology and Virtue Ethics. *The Journal of Ethics* 7: 367–395.

Nussbaum, Martha. 2000. *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge University Press.

Randell, T.T. 2004. Medical and Legal Considerations of Brain Death. *Acta Anaesthesiologica Scandinavica* 48: 139–144.

Schwartz, Shalom, and Avi Gottlieb. 1957. Bystander Anonymity and Reactions to Emergencies. *Journal of Personality and Social Psychology* 39: 418–430.

Sreenivasan, Gopal. 2002. Errors About Errors: Virtue Theory and Trait Attribution. *Mind* 111: 47–68.

Sreenivasan, Gopal. 2013. The Situationist Critique of Virtue Ethics. In *The Cambridge Companion to Virtue Ethics*, ed. Daniel C. Russell, 290–314. Cambridge: Cambridge University Press.

Upton, Candace. 2009. Virtue Ethics and Moral Psychology: The Situationism Debate. *The Journal of Ethics* 13: 103–115.

Zhong, Chen-Bo, Vanessa Bohns, and Francesca Gino. 2010. Good lamps are the best police: Darkness increases dishonesty and self-interested behavior. *Psychological Science* 21(3): 311–314.

References

Augustine. 1955. *Confessions and Enchiridion.* Trans. Albert Cook Outler. Philadelphia: Westminster Press.

Blackburn, Simon. 1981. Reply: Rule-Following and Moral Realism. In *Wittgenstein: To Follow a Rule*, eds. S. Holtzman, and C.M. Leich, 163–187. London: Routledge.

Boghossian, Paul. 1989. The Rule-Following Considerations. *Mind* 98(392): 507–549.

Clark, Andrew, and David Chalmers. 1998. The Extended Mind. *Analysis* 58(1): 7–19.

Copp, David. 2015. Explaining Normativity. *Proceedings and Addresses of the American Philosophical Association* 89: 48–73.

Dennett, Daniel. 2013. *Intuition Pumps and Other Tools for Thinking*. New York: W.W. Norton.

Dewey, John. 1916/2008. The Nature of Method. In *The Middle Works of John Dewey, 1899–1924, Volume 9: 1916, Democracy and Education*, ed. Jo Ann Boydston, 171–187. Carbondale, IL: Southern Illinois University Press.

Dotov, Dobromir, Nie Lin, and Anthony Chemero. 2010. A Demonstration of the Transition from Ready-to-Hand to Unready-to-Hand. *PLoS One* 5(3), e9433. doi:10.1371/journal.pone.0009433.

Dreyfus, Hubert L. 1990. *Being-in-the-World: A Commentary on Heidegger's Being in Time, Division I.* A Bradford Book. Cambridge, MA: The MIT Press.

Fodor, Jerry. 1974. Special Sciences. *Synthese* 28(2): 97–115.

Gampel, Eric H. 1997. The Normativity of Meaning. *Philosophical Studies* 86: 221–242.

Gell-Mann, Murray. 1995. *The Quark and the Jaguar: Adventures in the Simple and the Complex.* New York: Henry Holt.

Goldman, Alan H. 1977. Plain Sex. *Philosophy and Public Affairs* 6(3): 267–287.

Harman, Gilbert. 1986. Moral Explanations of Natural Facts – Can Moral Claims Be Tested Against Moral Reality? *The Southern Journal of Philosophy* Supp. 24: 57–68.

Haugeland, John. 1982. Weak Supervenience. *American Philosophical Quarterly* 19: 93–101.

Haugeland, John. 1998. Toward a New Existentialism. In *Having Thought: Essays in the Metaphysics of Mind*, 1–6. Cambridge, MA: Harvard University Press.

Hume, David. 1740/1978. *A Treatise of Human Nature.* Oxford: Clarendon Press.

Kelly, Erin. 2004. Against Naturalism in Ethics. In *Naturalism in Question*, ed. Mario DeCaro and David Macarthur, 259–274. Cambridge, MA: Harvard University Press.

Kim, Jaegwon. 1988. What Is Naturalized Epistemology? *Philosophical Perspectives* 2: 381–405.

Kornblith, Hilary. 1999. In Defense of Naturalized Epistemology. In *The Blackwell Guide to Epistemology*, ed. John Greco and Ernest Sosa, 158–169. Malden, MA: Blackwell Publishers.

Kuhn, Thomas S. 1970/2000. Reflections on My Critics. In *The Road Since Structure*, 123–175.

Lance, Mark. 1996. Quantification, Substitution, and Conceptual Content. *Noûs* 30(4): 481–507.

Loewer, Barry. 1997. A Guide to Naturalizing Semantics. In *A Companion to the Philosophy of Language*, ed. B. Hale and C. Wright. Oxford: Blackwell Publishers.

McDowell, John. 1978. Are Moral Requirements Hypothetical Imperatives? *Proceedings of the Aristotelian Society, Supplementary Volumes* 52: 13–29.

McDowell, John. 1994. *Mind and World.* Cambridge, MA: Harvard University Press.

McNaughton, David, and Piers Rawling. 2003. Naturalism and Normativity. *Aristotelian Society Supplementary Volume* 77(1): 23–45.

Peirce, Charles Sanders. 1992. *The Essential Peirce: Selected Philosophical Writings, Vol. I: 1867–1893*, ed. N. Houser, and C. Kloesel. Bloomington, IN: Indiana University Press.

Price, Huw. 2004/2011. Naturalism Without Representationalism. In *Naturalism Without Mirrors*, 184–199. Oxford: Oxford University Press.

Putnam, Hilary. 1981. *Reason, Truth and History*. Cambridge: Cambridge University Press.

Putnam, Hilary. 2004. *Ethics Without Ontology*. Cambridge, MA: Harvard University Press.

Quine, W.V.O. 1981. *Theories and Things*. Cambridge, MA: Harvard University Press.

Rorty, Richard. 1991. Solidarity or Objectivity? In *Objectivity, Relativism and Truth: Philosophical Papers*, ed. Richard Rorty, 21–34. Cambridge: Cambridge University Press.

Sellars, Wilfrid. 1949. Language, Rules and Behavior. In *John Dewey: Philosopher of Science and Freedom*, ed. Sidney Hook, 289–315. New York, NY: The Dial Press.

Sellars, Wilfrid. 1959/1991. Phenomenalism. In *Science, Perception, and Reality*, 60–105. Atascadero, CA: Ridgeview.

Sellars, Wilfrid. 1963a. Abstract Entities. *Review of Metaphysics* 16(4): 627–671.

Sellars, Wilfrid. 1963b. Some Reflections on Language Games. In *Science, Perception and Reality*, 321–358. Atascadero, CA: Ridgeview.

Sellars, Wilfrid. 1969. Language as Thought and as Communication. *Philosophy and Phenomenological Research* 29(4): 506–527.

Sellars, Wilfrid. 1974. Meaning as Functional Classification. *Synthese* 27: 417–437.

Sen, A. 1984. *Resources, Values, and Development*. Cambridge, MA: Harvard University Press.

Simon, Herbert. 1969. *The Sciences of the Artificial*. Cambridge, MA: The MIT Press.

Sturgeon, Nicholas. 1988. Moral Explanations. In *Essays on Moral Realism*, ed. Geoffrey Sayre-McCord, 229–255. Ithaca, NY: Cornell University Press.

Vigen, Tyler. n.d. Divorce rate in Maine Correlates with Per Capita Consumption of Margarine (US). *Spurious Correlations.* http://www.tylervigen.com/view_correlation?id=1703. Accessed 25 June 2015.

Wilson, Edward O. 1998. The Biological Basis of Morality. *The Atlantic Monthly*, April: 53–70.

Wittgenstein, Ludwig. 1969. In *On Certainty*, ed. G.E.M. Anscombe and G.H. von Wright. New York: Harper and Row.

Index

Note: Page number followed by 'n' refers to notes.