
Natural Ethical Facts

Evolution, Connectionism, and Moral Cognition

William D. Casebeer

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

Contents

Preface and Acknowledgements	vii
1 Natural Ethical Facts	1
2 Clearing the Way for Reduction: Addressing the Naturalistic Fallacy and the Open-Question Argument	15
3 The Functional Account of Ethics: Functional Explanation in Biology and a Corresponding Account in Morality	37
4 Moral Judgment, Learning in Neural Networks, and Connectionist Mental Models	73
5 Connectionism and Moral Cognition: Explaining Moral Psychological Phenomena	101
6 Applications and Critique: Moral Theory, Moral Practice, Moral Institutions	127
7 Objections and Conclusions: Nature and Norms	149
Notes	163
References	185
Index	211

Natural Ethical Facts

Why Care about Natural Ethical Facts?

Evolutionary biologists have been at work for more than 100 years telling us about our nature as evolved, embodied creatures. Cognitive scientists have been plumbing the depths of the mind for 50 years, discovering the neural and computational roots of complex behavior and cognition. For more than 2,000 years, moral philosophers have been plugging away at big-picture normative theories regarding how we ought to conduct ourselves and, ultimately, what the point of this blooming and buzzing confusion of life and mind is. Until relatively recently, however, work at the intersection of these three areas of inquiry was difficult to find. Scientific theories of life and mind have had relatively little contact with normative moral theory, and moral philosophers, when they have made contact, have often expressed disappointment with the results. Why is this? What can we do to ensure that fruitful consilience between our best theories in the cognitive sciences, evolutionary biology, and ethics is the norm rather than the exception? Addressing these issues by showing how there can be useful interactions between science and ethics is the critical issue facing the sciences. As we cast about for a post-Enlightenment normative anchor, if we are to prevent backsliding into dogmatic supernatural and non-naturalistic conceptions of the moral life, it is imperative that we demonstrate the possibility of intelligent, useful interactions between the human sciences and human ethics.

This book is an attempt to show that, theoretically speaking, there is no reason to rule out a scientific naturalized ethics *tout court*, and that, practically speaking, by taking into account recent developments in

evolutionary biology and the cognitive sciences, the outlines of one promising form of such an ethics can be sketched. It will be a pragmatic neo-Aristotelian virtue theory, given substantive form by both conceptions of function from evolutionary biology and connectionist conceptions of thought from cognitive science. The rough structure of the book follows from the unfolding of this admittedly synoptic thesis.

Moral Judgments, Connectionism, and the Cognitive and Biological Sciences

The naturalization of ethics has been a problematic enterprise for moral philosophers. Historically, there are several reasons why this is so. For one, theoretical arguments regarding the impossibility of a systematic reductive relationship between the natural realm and the normative realm have stymied attempts to unify the two spheres by those sympathetic to such a union. In addition, the cognitive capacities we use to grasp moral knowledge have been thought by some to be far too subtle for “mere” empirical explanation by a scientifically informed theory of cognition. Finally, some previous attempts to construct a scientifically informed moral theory, and thus remake ethics into a science, have been too simplistic (or have been painted as such by critics) to do justice to the full range of our considered moral intuitions and our reasonably informed moral judgments. As a result, much of the work in the naturalization of morality has taken place in metaethics rather than in normative moral theory, leaving the latter bereft of empirical content. And very little research has attempted to relate the latest findings of the cognitive sciences to moral psychology and moral judgment, let alone normative moral theory, in any systematic fashion.

This isolation has had a debilitating effect on both the empirical plausibility of normative moral theories and the societal impact of the biologically informed cognitive sciences. Our normative moral theories would be greatly enriched if the questions they posed were empirically tractable, and the breadth of our cognitive and biological sciences would be enhanced if they were to offer plausible reconstructions of our cognitive capacity to reason about, grasp, and accede to moral norms. Such an enrichment and enhancement also would pay dividends external to the academic professions, giving us alternate strategies for framing and

resolving moral conflicts and allowing us to improve our methods for cultivating moral knowledge by enhancing the effectiveness of our collective character-development institutions.

My project embodies a synoptic reconciliation of the sciences of cognition with a fully naturalized conception of morality. I argue that we can improve our understanding of the nature of moral theory and its place in moral judgment if we better understand just what morality consists in. Such an understanding will best be informed by treating morality as a natural phenomenon subject to constraints from, influenced by, and ultimately reduced to the sciences, particularly the cognitive sciences and biology. Treating morality as a matter of proper function, biologically construed (e.g., at least partially fixed by our evolutionary history), with a concomitant emphasis on skillful action in the world, will also shed light on just what kind of creatures we must be (cognitively speaking) if we are to possess knowledge about morality so taken. Connectionist accounts of cognition can best accommodate this style of knowledge and can also account for other gross moral psychological phenomena, giving them ample explanatory power and making them the centerpiece of moral cognition. The nature of morality and the picture of moral cognition I defend are rooted in a pragmatic construal of knowledge and in a modern, biologically informed neo-Aristotelianism. Exploring these roots, particularly as they manifest themselves in John Dewey's theory of moral deliberation, will shed light on the role of moral theory in such a scheme and will help distinguish this approach from less fruitful and more purely sociobiological undertakings. Finally, I discuss objections and draw out some practical implications, regarding the nature and form of our collective character-development institutions and our methods for moral reasoning that arise from taking this approach seriously.

The Way Forward

In chapter 2, I discuss and rebut two popular arguments against a reductive and naturalizable account of morality: the naturalistic fallacy and the open-question argument. I contend that both arguments fail, primarily because they rely on an outmoded analytic/synthetic distinction. Arguing for a continuum of analytic and synthetic judgments, thus demonstrating

that moral knowledge and scientific knowledge are commensurable, will open the way for a reductive naturalistic account of morality. I accomplish this by recapitulating W. V. O. Quine's arguments against the analytic/synthetic distinction. I also present the basics of Dewey's theory of moral deliberation, arguing that his conception of "ends-in-view" effectively demonstrates the continuity of scientific and practical knowledge with moral knowledge. The conception of morality I thus offer will be cognitivist and realist but will nonetheless constraint our ability to systematize moral theory. Moral conclusions, I will argue, follow abductively from properly construed non-normative premises. Our moral judgments are part and parcel of our web of beliefs. If the proper reductive relationship between moral terms and natural terms is captured by a theory that relates the two in a fecund way, then inferences from non-normative premises to normative conclusions will not be excessively licentious.

In chapter 3, I articulate the basics of such an approach, rebutting the "error-theory" arguments against a moral science articulated by John Mackie. Moral claims should be reduced to functional claims technically construed, hence the shared roots with an Aristotelian view of the world. Such functional claims should be treated as they are in biology and the life sciences, with a suitably modified Wright-style teleonomic analysis: a Godfrey-Smith-flavored "modern-history" theory of functions. Such a theory will thus take advantage of the explanatory power of the neo-Darwinian synthesis. Some functional facts about human beings fully fix normative claims; others will only constrain the possible state space of moral options. A small percentage of the decisions we face may have no effect at all on functional concerns, in which case we are (morally speaking) simply free to choose. The basics of this account will thus allow some flexibility in the normative structure of our lives. My account also has the resources necessary to distinguish itself from hedonistic, egoistic, desire-satisfaction, and utilitarian theories of morality, particularly after I make some crucial distinctions (including the difference between proximate and distal functions and the difference between ahistorical and historical functions). On this picture, moral facts are not "queer" and unscientific, nor is morality globally relativistic and dramatically contingent. We can in good conscience be moral realists and yet embrace an acceptable form of humility regarding our ability to

know the good; such humility reflects not only constraints on our cognitive economy but also constraints on the form of norm-fixing evolutionary processes in nature. Ultimately, this approach makes empirical and scientific investigation of moral normativity possible. I also examine contemporary work done in the same vein, including more purely sociobiological and Darwinian approaches to morality. I focus primarily on modern accounts, ranging from Larry Arnhart's theory to E. O. Wilson's, although I briefly discuss wrong-headed evolutionary ethical theories, such as those offered by Herbert Spencer and the Social Darwinists. I discuss similarities and differences between these approaches and my own, concluding that the account on offer has strengths that the other approaches lack.

In chapter 4, I draw on resources from connectionist accounts of cognition and from the embodied cognition movement to articulate a purely biological notion of moral judgment that bridges the "normativity gap." Using resources from these two approaches, it becomes possible to specify a conception of judgment that harmonizes with the account of moral knowledge discussed in chapter 3. A purely biological notion of judgment is possible, and such a notion comports well with the idea of judgment as the cognitive capacity to skillfully cope with the demands of the environment. Thus, moral judgment is possible only in systems that learn in a natural computational manner, whose nature is at least momentarily fixed,¹ and that exist in an environment where demands are placed upon the organism. Having good moral judgment amounts to being able to accomplish cognitive tasks that enable one to meet the demands of one's functional nature. Morality is therefore a matter of "knowing how" more than a matter of "knowing that." Some of these cognitive capacities can be captured in "representation-free" neural nets that are best described in the language of dynamical systems theory; others require traditional connectionist distributed representations. Some advanced forms of moral reasoning may require a model-theoretic account of reasoning. I discuss what mental models look like in connectionism, postulate how they can accommodate more advanced aspects of moral cognition, and point out their essential connection to action in the world and embodiment in an organism. Certain high-level aspects of connectionist mental models may lend themselves to a truth-functional analysis rooted in a symbolic redescription of network activity, but such

a redescription will be possible only in certain instances and should not be reified into a categorical demand placed upon normative action and its associated psychology. I draw connections between this discussion and Dewey's account of moral deliberation, which I sketched in chapter 2. I also offer a useful typology of moral characteristics that follows from this account, distinguishing between those objects of science that are the proper subjects of moral *cum* functional concerns, and between creatures that are able to effectively model their environment and their relationship to it (and that can hence formulate their own moral science). This generates a continuum among living things that have functions, ranging from simple moral agents (for example, most insects) to maximally robust moral reasoners (most social creatures with a significant range of behavioral repertoires, especially—but not only—human beings).

In chapter 5, I use the explanatory power of a connectionist approach to account for other gross features of moral reasoning. The interaction of advances in connectionist accounts of thought and traditional issues in moral cognition and psychology is an interesting one, as heretofore disparate phenomena in the latter can be unified by an account from the former. Connectionism can serve as a platform on which to reconstruct several high-order moral cognitive phenomena, including moral knowledge, moral learning and conceptual development, moral perception and the role of metaphor and analogy in moral argument, the appearance of staged moral development, the possibility of *akrasia* (acting against one's best considered judgment), the presence of moral systematicity, moral dramatic rehearsal and moral motivation, and moral sociability. A connectionist account of moral cognition best unifies the neurobiology and cognitive psychology of morality and sheds new light on traditional issues in moral psychology, including questions about the motivational efficacy of moral claims, the affective aspect of moral reasoning, and the importance of moral exemplars. I support these contentions with reference to the exponentially increasing body of modeling work in artificial neural networks. Finally, I briefly examine the literature relating brain structure and function to these models, identifying key components of the several cognitive systems that jointly constitute our capacity to be maximally robust moral reasoners.

In chapter 6, I draw together themes from the preceding five chapters, examining how naturalizing morality by way of evolution and

connectionism may affect our moral theories, our moral practices, and our moral institutions. Where does this attempt at reduction leave traditional moral theory? On the one hand, some aspects of moral theory—particularly an appropriately naturalized Aristotle and large parts of Dewey’s attempt to develop a pragmatic ethic—remain components of the moral life; on the other hand, certain traditional moral theories do not fare as well, at least if they are taken to be universally applicable. A Kantian approach, for example, has at best heuristic value but at root makes demands that are psychologically unrealistic. I conclude that it functions well as a device for drawing attention to the strong conditions necessary to enable social reasoning to occur, but that it fails to appropriately accommodate primary functional concerns. This pragmatic approach recognizes a healthy limit to the usefulness of grand moral theory: its existence can be explained, but its limits are outlined. Ethical reasoning becomes a species of pure practical knowledge and as such is responsive to the demands of the present. Just as pragmatic epistemology is a process-oriented philosophy, so too is a pragmatic ethics that draws on the useful portions of previous moral theorizing, insofar as they are informed by and illuminate the issues raised by functional *cum* biological concerns. This emphasis on proper function is rooted in an Aristotelian account of the nature of humanity and requires the defense of at least a “soft essentialism,” which I offer here by adverting to the findings of the neo-Darwinian synthesis. Though we might think that one of the primary lessons of Darwinism is that there is no such thing as a species essence, I argue that population thinking serves as a healthy corrective to the idea that our functions are immutable and that all of us must possess exactly the same functional natures. I discuss the similarities between this explicitly pragmatic approach and an Aristotelian virtue ethic, arguing that the two are successfully unified with very little remainder and that the neo-Darwinian synthesis can give biological bite to Aristotle’s contentions about the limits of moral theorizing. I conclude chapter 6 by using the aforementioned approach as a tool to critique character-development institutions and to illuminate cases of moral conflict. I address real-world case studies in ethics that demonstrate how this conception has the ability to contend with these objections directly and not just abstractly. I focus first on whether an individual should develop deep or wide friendships (modern-history

functions call for deep friendships) and second on how we should structure our societies (modern-history considerations lead to liberal democratic forms of organization). In more abstract and general terms, my account restores an emphasis on habituation and mindfulness that our social institutions would do well to attend to. I examine the implications of this view for character development and moral education, arguing that it propels to the forefront a narrative-driven case-study approach to moral education, a solid grounding in the biological and sociological dimensions of the human situation, a careful tending of the institutional environment in which moral action is situated, a demand for consistency between articulated principles and practical actions, and a healthy flexibility in the practical application of rules and regulations. Nothing teaches like experience, and so the proper environment for moral experience must be carefully cultivated and maintained by institutions of moral education and character development. Such a process is demanding and requires those engaging in it to stay informed of the results from a large number of fields of empirical inquiry.

In chapter 7, I address the remaining objections to the aforementioned approach and outline its additional strengths. It must answer some hard questions usually put to more traditional sociobiological undertakings that any naturalistic account of morality must deal with. Among the grounds for concern are the perceived lack of robust and genuine normativity in the approach, some purportedly morally repugnant “entailments” of the position, an argument that the position demands its own rebuttal for heuristic “Platonic noble lie”-style reasons, and an argument that the position is empty of useful moral content. In the conclusion of this chapter, I outline several areas where there is a notable absence of empirical work or where more empirical work is needed; these areas include the connectionist modeling of moral cognition, applied moral cognitive psychology, moral anthropology, the neurobiology of moral cognition, and biologically informed game-theoretic approaches to skillful coping. I also discuss the need for further exploration of more traditionally philosophic topics, such as alternatives to a simple-correspondence account of cognition. A biological and neurobiologically informed pragmatic ethic holds the most hope for being the unifying procedural glue that can successfully hold together otherwise disparate and possibly mutually antagonistic approaches to the moral

life. Although moral progress using this approach is not a given, I highlight its essentially optimistic character and hold out hope for reconciliation between the humanities and the sciences.

‘Naturalism’ and ‘Ethics’: Problematic Terms?

Before I begin my discussion of the naturalistic fallacy, there are several terms whose use demands clarification so that the nature of this approach is clear. These include ‘naturalism’ and ‘ethics’. (Entire books have been written about the definition of these terms, so my discussion will be concise.)

‘Naturalism’

The principal approach that I will use in the book is best typified as a form of methodological naturalism, by which I mean that the methodological and epistemological assumptions of the natural sciences should serve as standards for this inquiry. If at the end of the inquiry we feel compelled to postulate the existence of a non-naturalistic entity or process, so as to best explain the results of our study, then our methodological naturalism will have led us to a denial of ontological naturalism. However, I don’t think this will be the case, and for the moment we should hold our methodological naturalism close so as to see if normativity can be derived without postulating “spooky” non-natural entities (gods, a noumenal realm, and so on). Of course I will avail myself of the ontologies postulated by the natural sciences during the course of this inquiry, but this will be done with requisite sensitivity to moral experience, and with the fallibilistic view that the ontologies of our current sciences might be wrong, so, although the project will presuppose ontological naturalism to a certain extent, naturalist methodologies are still the primary constraint.

Dewey (1902, p. 142) provides a nicely succinct definition of naturalism: “The theory that the whole of the universe or of experience may be accounted for by a method like that of the physical sciences, and with recourse only to the current conceptions of physical and natural science; more specifically, that mental and moral processes may be reduced to the terms and categories of the natural sciences. It is best defined

negatively as that which excludes everything distinctly spiritual or transcendental. . . .”

Some of the traditional methodological and ontological theses of naturalism will be actively defended in this paper; others will be assumed. For example, I will actively defend a realist conception of morality, whereas I will simply assume that there are no miracles and there is no extrasensory perception (at least until evidence demands that we change these assumptions). In other words, my defense of certain traditional tenets of naturalism will take place against the background of (a) uncontroversial findings from the sciences (e.g., no ESP), (b) controversial but eminently defensible findings from the sciences (e.g., the explanatory power of connectionist approaches to cognition), and (c) the interesting points of conflict between fields of inquiry not generally considered to be part of the sciences (e.g., certain assumptions about the nature of ethical claims) and the sciences of cognition and life.

Gerhard Vollmer’s list of the traditional ontological and methodological theses of naturalism (taken from his “Naturalism, Function, Teleonomy,” as published in Wolters 1995) is worth quoting in full:

- A) Only as much metaphysics as necessary!
- B) As much realism as possible!
- C) For the investigation of nature, the method of empirical science is superior to any other.
- D) Nature (the world, the universe, the real) is, at bottom, constituted of matter and energy, both temporally and causally.
- E) All real systems—the universe as a whole included—are subject to development, to evolution, to assembly, and disassembly. That’s why any modern naturalism is an evolutionary naturalism.
- F) Complex systems consist of and originate from less complex parts.
- G) The real world is interconnected and quasi-continuous.
- H) Instances transcending all human experience are conceivable, but dispensable for the consideration, description, explanation and interpretation of the world.
- I) There are no miracles.
- J) There is no extrasensory perception.
- K) Understanding nature doesn’t transcend nature itself.
- L) There is a unity of nature which might be mirrored in a unity of science.

The naturalization of ethics would thus entail making ethics consistent with this list of statements and thereby showing how knowledge of the normative can be derived and justified using this methodology and ontology. As Vollmer notes, every thesis on this list deserves explication

and refinement, but I hope they are intelligible without this and that they serve as useful guideposts for present purposes.

Jay Garfield (2000, p. 423) distinguishes between strong naturalism and moderate naturalism. Strong naturalism requires more than mere consistency (which is demanded by even the weakest forms of naturalism); it also requires entailment or some form of reduction to more fundamental and already unproblematically naturalized theories. Moderate naturalism would require (1) consistency, (2) that the research be guided by the methodological canons of the sciences, and (3) that there be (in Garfield's words) "plausible explanatory strategies for linking the theories, explanations and theoretical perspectives" of the body of knowledge being naturalized to the remainder of science. In my case, I will be happy if I achieve a moderate naturalization, but I keep in mind the goal of strong naturalization as a regulative ideal. This reflects my suspicion that mere supervenience relations, though acceptable in a developing science, often are used as an excuse not to explore the phenomena in question in more depth, or, in the worst of cases, merely restate a problematic relation rather than "solving" it.²

In sum, we should expect that a plausible naturalization of ethics would explain the essential nature of moral judgments, their subject matter, and how we come to make them. Such a naturalization would make full use of background knowledge from the sciences, especially (at least in the case of this book) from the cognitive sciences and evolutionary biology.

The Natural Method

Keeping the background knowledge of the pertinent sciences in mind while constructing a theory has been given a name by Owen Flanagan: the Natural Method.³ Though Flanagan uses it to triangulate on a theory of the nature of the mind (paying attention to results from the associated departments of the cognitive sciences, as well as to first-person phenomenology), there is no principled reason why the process couldn't be applied to any phenomenon of interest. Flanagan (2000, p. 14) characterizes the Natural Method as follows: "The idea is to keep one's eye, as much as is humanly possible, on all the relevant hypotheses and data sources at once in the attempt to construct a credible theory. The

natural method involves seeking consistency and equilibrium among different modes of analysis applied to the study of some . . . phenomenon.” Flanagan’s prescription derives in part from Quinean considerations about confirmatory holism. Insofar as these considerations also drive my inquiry (as will become evident at the end of chapter 2), it is no surprise that the method I advocate for framing theories of morality is, in essence, the Natural Method.

Two Desiderata for Naturalization

To summarize the desiderata for naturalism (for comparison to the conclusions of chapter 7), naturalizing ethics would therefore consist in producing (1) an account of moral normativity that roots normativity in nature, where the content of nature’s ontology is (provisionally⁴) provided by the methodological canons of the natural sciences, and (2) an account of our capacity to grasp and accede to these norms that is rooted in the best theoretical frameworks that the mind sciences have to offer.

‘Ethics’

What does the subject matter of the study of morality consist in? Broadly speaking, it is the study of what we ought to do, what we ought to intend, or what kind of people we ought to be, all in the largest sense—how ought we live our lives? The three traditional theoretical approaches to ethics have been thought to answer these questions in turn: utilitarianism⁵ focuses primarily on the consequences of actions (as they relate to the production of pleasure and the reduction of pain), deontology⁶ concentrates on what duties we owe to one another (and, in its most famous Kantian version, on what duty-filtered maxims or intentions we ought to form in our minds), and virtue theory⁷ considers what states of character we ought to cultivate in ourselves. In the course of this book, I discuss all three of these theories as they relate to naturalization, particularly virtue theory.

There are many more fine-grained distinctions to be made here, beginning with the difference between instrumental reasoning and reasoning

about final ends. On the one hand, we can ask what we ought to do given some desire or project; such a question is one of means and involves instrumental reasoning. What is the best means or instrument I can use to accomplish my goal? On the other hand, we can ask what we ought to desire or what projects we ought to have; such a question is one of ends and involves practical reasoning about *final* ends. Naturalized systems of ethics, particularly modern approaches, are often accused of dealing only with the former, and hence of not dealing with ethics proper at all. In this project, I intend to deal with both instrumental and final norms, although the distinction often obscures the true nature of moral reasoning and can cloud inquiry. Rather than construing “grand theory” ethics as the search for final ends, we should seek explanatory unification of reasoning about both instrumental and final ends.

Some authors draw a distinction between morality and ethics. For example, Bernard Williams argues that morality is a subset of ethics, and that the former concentrates on obligation whereas the latter deals with larger questions.⁸ Others argue that ethics is a specialized body of knowledge applicable only to certain roles, and that morality is actually the larger term; there can be “military ethics” or “medical ethics,” both of which derive their content from more general *moral* considerations.⁹ I am dubious about the work done by drawing these distinctions, at least for this project (although in other contexts, such a distinction might be eminently useful). For present purposes, then, the terms ‘ethics’ and ‘morality’ will be used interchangeably, and no particular substantive inferences about the project should be drawn from my use of one term instead of the other.

Final Context

Philip Kitcher offered an enlightening list of potential alternative goals for those who would “biologize” ethics. Kitcher formulated the list while attempting to discern the exact nature of the project encompassed by E. O. Wilson’s sociobiology, which Kitcher criticized in his 1985 book *Vaulting Ambition*. Kitcher’s piercing critique of Wilson is a healthy corrective to both excessive ambition and vagueness, though

Wilson's program has much about it that is worth admiring.¹⁰ Kitcher (1985, pp. 417–418) postulates four possibilities for “biologicizing” (E. O. Wilson's neologism) morality:

A. Evolutionary biology has the task of explaining how people come to acquire ethical concepts, to make ethical judgments about themselves and others, and to formulate systems of ethical principles.

B. Evolutionary biology can teach us facts about human beings that, in conjunction with moral principles that we already accept, can be used to derive normative principles that we had not yet appreciated.

C. Evolutionary biology can explain what ethics is all about and can settle traditional questions about the objectivity of ethics. In short, evolutionary theory is the key to metaethics.

D. Evolutionary theory can lead us to revise our system of ethical principles, not simply by leading us to accept new derivative statements—as in (B)—but by teaching us new fundamental normative principles. In short, evolutionary biology is not just a source of facts but a source of norms.

Though it is a stretch to say that any single science (let alone evolutionary biology) can do all these things, I will claim that collectively the sciences can accomplish A–D.¹¹ The methodologies and the ontologies of the science are up to the task, particularly if our approach is subtle. In particular: I think the cognitive sciences have the leading role in A; both cognitive science and biology can contribute to B; the evolutionary sciences—evolutionary biology, ecology, systematics, etc.—can answer C (I will defend a version of realism using those resources); and both cognitive science and evolutionary biology can answer D (they reaffirm an appropriately naturalized virtue ethic, such as that developed by Aristotle and Dewey, and they can inform normative principles in interesting and enlightening ways). Minimally, and relatively uncontroversially, this book will make a contribution to A and B. Maximally, and controversially, it will also make a contribution to C and D.

So, on to certain pieces of philosophical undergrowth that must be cleared out before the project can begin in earnest, beginning with the naturalistic fallacy. Is ethics explanatorily autonomous from the sciences? Can a valid argument be given that has only factual premises and a normative conclusion? Doesn't the nature of the concepts of “normative” and “empirical” preclude any meaningful interplay between the two, and if it does, what kinds of interaction are prohibited? Depending on our answers to these questions, we may be able to rule out naturalization from the start.

Clearing the Way for Reduction: Addressing the Naturalistic Fallacy and the Open-Question Argument

Metaethics: Cognitivism and Non-Cognitivism

The status and the nature of moral claims have been topics of controversy in metaethics for as long as the field has existed as an independent arena of inquiry; settling arguments about these issues is in fact the metaethical *raison d'être*. One way of resolving disputes regarding just what it *is* that moral judgments make claims *about* is to ask whether such judgments are truth evaluable.¹

The non-cognitivist argues that moral judgments are not truth evaluable because (for example) they are merely expressions of attitudes or emotions—in much the same way that “jealousy” is not a truth evaluable claim (as jealousy does not refer to anything independent of the emotional state of the person experiencing jealousy), neither are moral claims. This “boo-hurrah”² metaethical view stands in opposition to cognitivism, the school of thought according to which moral claims are indeed truth evaluable. The cognitivist claims that, just as the statement “This dog’s mass is 20 kilograms” can be true or false, so too can the statement “This act is immoral.” Though most ethicists today adopt cognitivism as a default position,³ there is still heated debate within the cognitivist camp regarding just what should happen next.⁴ Though many cognitivists want to be good reductive naturalists too, the seeming irreducibility of moral claims to perfectly ordinary and empirically tractable ones has presented an “anti-reductionist roadblock” past which many have been afraid to travel.

The arguments for irreducibility have driven some philosophers, such as George E. Moore, to abandon naturalism about ethical claims; others, such as John McDowell, have become non-reductive naturalists.

Some non-cognitivists even offer these irreducibility arguments as a strong motivation for abandoning cognitivism. By my lights, however, the two main historical arguments against reduction, Hume’s “naturalistic fallacy” and Moore’s “open question argument,” fail to establish such a roadblock. Supporting this claim will pave the way for an explanation of my particular brand of reductive cognitivism—there is such a thing as a moral fact, and such facts are complexes of functional claims, where functionality is given a thoroughly naturalistic interpretation.

The Naturalistic Fallacy and the Open-Question Argument: Barriers to Naturalization?

In this chapter I will argue that both the naturalistic fallacy and the open-question argument fail. Each, either implicitly or explicitly, relies on the distinction between analytic and synthetic statements for its force. Insofar as we have good reasons (thanks to Quine and Dewey⁵) to doubt that such a distinction exists, anti-reductionism has lost much of its force.

I will end the chapter with a survey of the nature of the relationship between empirical statements and moral theories. Although the use of normative language does capture a unique and important aspect of the world (namely, planning by organisms to achieve ends), it does not point to an ontological barrier that somehow separates the natural world from non-natural normativity. The leap from ‘is’ to ‘ought’ becomes an ever-so-tiny web-of-belief-driven inference when the objective correlates of normative terms are appropriately scientifically explicated, and when we view “ought” statements as recommendations about the habits humans and other organisms need have if they are to relate in fruitful ways to those objective correlates.

Terminology

Before I offer a brief exposition of the naturalistic fallacy and the open-question argument, I should clear up some terminology. Although Hume was the first to note the seeming invalidity of inferring an ‘ought’ statement from a list of ‘is’ statements, he did not actually use the phrase “the naturalistic fallacy.” Rather, G. E. Moore (1902) popularized these words in his discussion of his own “open question argument.” Moore’s argument was directed specifically against attempts to naturalize the term

‘good’, whereas Hume’s argument applied more generally to all normative terms. Following most other philosophers, I will thus treat the open-question argument as a species of a naturalistic fallacy, giving Hume credit for the general argument and Moore credit for the specific one.⁶

What Is Not at Stake

Before examining Hume and Moore’s arguments, let me briefly detail what exactly is *not* at stake in the debate. This is crucial, as wrong-headed refutations of the naturalistic fallacy can do more harm than good for naturalism in ethics. First, no reasonable naturalist in ethics would deny that certain states of affairs in the world are good and others are bad. The point of a naturalistic ethics is just to give a natural yardstick against which to measure such affairs. Thus, it won’t do to say in response to the naturalist “You can’t infer from the fact that x exists that x is good,” as any plausible naturalistic ethical theory will be in agreement. For example, we can’t infer from the fact that there is inequality that inequality is good. The question is: Will the norm that we use to criticize inequality originate in nature, or will it originate and be justified supernaturally? Second, no reasonable naturalist in ethics would argue that naturalism in ethics entails the elimination of normative language from our vocabulary. It might very well be that normative terms (such as ‘ought’ and ‘should’), when given the appropriate theoretical explication, are proxies for sets of empirical statements (or, more richly, as statements about what would happen if we behaved in certain ways—that is, as scientific statements), but that is not to say that we should then use these statements rather than the normative terms in everyday discourse. When embedded in the appropriate theory, such normative terms will have explanatory power and pragmatic use. We might have to reform or modify some of our moral concepts, true, but there is no need to dispense with moral language as a result.

What Is at Stake

What is at stake is the nature of the relationship between normative moral theories and traditional empirical scientific theories. Both of the arguments I discuss in this chapter contend that we have *a priori* reason to think that there can be no legitimate form of strong intercourse

between normative theories and empirical theories. Can normative theories be justified with the appropriate sets of empirical statements? Hume says no, as any inference from a list of 'is' statements to an 'ought' statement will be invalid—we cannot expect a normative theory to be supported only by scientific findings. Moore also says no, as we will never be able to reduce the primitive unanalyzable term 'good' to any natural predicate or term. Thus, the arguments turn on the question of legitimate possible relationships between empirical findings and normative theories.

Hume and the Naturalistic Fallacy

Hume first offered a general argument for the existence of the naturalistic fallacy in *A Treatise of Human Nature* (1739),⁷ where he discusses the transition from 'ought' to 'is', reminding us that it "is of the last consequence. For as this ought, or ought not, expresses some new relation or affirmation, 'tis necessary that it shou'd be observ'd and explain'd; and at the same time, that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it." Hume is "surprised" when authors writing about morality who were previously reasoning in the 'usual way' suddenly begin to substitute 'oughts' in places where before only 'is' had been present. Since Hume is often cited as a pre-eminent advocate of a naturalized ethics, one might be surprised to hear him offering this argument. However, in the context of the work, Hume is arguing that moral judgments (as it were) arise not from reason but from our passions. We should not look to reason for the wellspring of morality, for reason is the faculty we use to judge things true or false—it does not motivate us; rather, our *passions*, which are not ratiocinative, move us to act, and therefore only they can adequately ground morality. Thus, Hume is a non-cognitivist about moral claims, and hence the apparent tension between his naturalization of ethics and his formulation of the naturalistic fallacy is only apparent.⁸ For the naturalist who would also be a cognitivist, however, Hume's remarks do pose a problem, so much so that the Humean version of the naturalistic fallacy has its own name: "Hume's Law." It would appear that Hume has pointed out a serious flaw in any attempt to reason from the empirical to the

normative: that in your conclusion you will make reference to an unexplained term (the ‘ought’ term) that was nowhere present in the (empirical) premises of the argument. Such an argumentative structure is invalid, as the truth of the premises does not guarantee the truth of the conclusion.⁹

Moore and the Open-Question Argument

The open-question argument takes a similar approach. In his *Principia Ethica*, Moore argues that all naturalists about ethics are guilty of a common fallacy. They confuse the property of goodness with the things that possess it or with another property that the good things have. To commit the naturalistic fallacy is just to confuse the good with one or both of these other things. Moore offers two arguments to support his claim. One is the open-question argument; the other is an argument from the addition of meaning (the import of this phrase will become clear later). First, I will examine the open-question argument.

If goodness were identical with another property, then every competent speaker of a language would consider it an ill-formed question to ask if the property in question is itself good; this would be akin to asking a fluent English speaker “Are birds birds?” But in fact we do not consider questions of the type “Is x good?” (where x represents your favorite contender for the reduction of the moral property “good”) to be nonsensical. Thus, if your brand of reductive naturalism is utilitarian, then others can, Moore argues, legitimately and sensibly confront you with the question “But is it good to maximize aggregate pleasure?” This indicates that the property in question and the property of being good are not actually identical. It is an open question for *any* natural property as to whether it is good. Moore’s conclusion is thus that goodness is and must be a simple, non-natural property.

The second argument Moore offers is an argument from the addition of meaning. If, for example, ‘good’ meant *pleasant*, then to say “What is pleasant is good” would provide us with neither additional information nor any extra reason to promote pleasurable states of affairs. But since saying “What is pleasant is good” *does* provide us with additional information and *does* give us extra reason to promote pleasure, then we cannot reduce the good to the pleasurable. Such an argument, Moore

says, generalizes to prevent any reduction of the term ‘good’.¹⁰ Again, goodness, on Moore’s view, is a simple, non-natural property.

Several moral philosophers (including Mark Johnson and Geoffrey Warnock) think that Moore did great damage to ethics by advancing these claims. He set the stage for the emotivism that predominated in early-to-mid-twentieth-century ethics. Johnson (1993, p. 140) summarizes:

By claiming that empirical evidence about who we are and how we function is simply irrelevant to the fundamental questions of moral philosophy, Moore initiated a serious decline in ethics (and in value theory generally) in this century, from which we are only beginning to recover. Quite simply, he so impoverished and marginalized reason that its only role in ethics was the determination of efficient means to ends and of probable causal connections. As Warnock has summed up, Moore leaves us with a realm of *sui generis* indefinable moral qualities about which reason can say nothing. We are confronted with a “vast corpus of moral facts about the world—known, but we cannot say how; related to other features of the world, but we cannot explain in what way; overwhelmingly important for our conduct, but we cannot say why.” [Warnock 1967, p. 16]

Moore and Hume Rely on an Implicit Analytic/Synthetic Distinction

One very important feature of Moore’s argument that may be transparent at this point is worth discussing in more detail. Moore is essentially arguing that the good itself is a simple, unanalyzable concept. In *Principia Ethica* (1902, p. 9) he writes: “‘Good’, . . . , if we mean by it that quality which we assert to belong to a thing, when we say that the thing is good, is incapable of any definition, in the most important sense of that word. . . . It is simple and has no parts.” Arguments from open questions and the addition of meaning all imply that the good *qua* good is non-synthetic, a simple property not amenable to reductive theoretical analysis. That is, if I say “The good is the pleasant,” the *reason* it makes sense to ask of the pleasant “But *is* it good?”—and the *reason* I acquire additional information and may obtain motivation to promote pleasant states of affairs when someone informs me that the pleasant *is* good—is just that we purportedly learn something new when we append the concept “good” to the concept “pleasant” (or whatever our contender for naturalization is). The good is *not analytically given* by any natural definition. If we think that there is no clear distinction between analytic and synthetic statements, and if we think that even simple statements

about the good are revisable in light of experience, then we will have gone a long way toward defusing Moore's in-principle objections to a naturalized ethic.

Interestingly, among Moore's belongings when he passed away was a new preface for a never-written second edition of *Principia Ethica*. This preface was published posthumously. In it, Moore spends a considerable time backing away from some of the claims he seems to be making in the text, concluding with this startling statement: "Some such proposition as this, namely, that G [the Good] is not identical with any natural or metaphysical property (as now defined), was more or less vaguely in my mind, I think, there is no doubt. . . . I was, I think, certainly confusing this proposition to the effect that G is not analyzable in one particular way, with the proposition that it is not analyzable at all." This is an incredible admission—we learn that Moore did *not* intend for the open-question argument to establish *a priori* that G could not be a natural property. Thus, Moore's argument boils down to this: We haven't been given a perfect naturalistic ethic *yet*, to which all but the most partisan naturalists about ethics would agree, myself included (although, with others, I think an appropriately scientifically updated Aristotle comes very close). Strangely, I have not been able to find a single work about the open question as it relates to evolutionary accounts of morality that discusses these interesting admissions. Since Moore examines only two naturalistic accounts of the meaning of 'Good' in his book (namely, hedonism and Herbert Spencer's evolutionary ethics), his conclusions suddenly seem much less grand. More realistically, his point becomes that Hedonism and Spencerian ethics are not good candidates for a reduction of moral properties to naturalistic properties. I agree, as do many other naturalists. Nonetheless, despite these clarifications, Moore *still* insists that "ethical propositions do involve some unanalysable notion, which is not identical with any natural or metaphysical property." I assume that the reason there hasn't been more discussion of these remarks is that they are taken from a posthumous manuscript.

In any case, Hume similarly relies on an implicit analytic/synthetic distinction. We find the new copula 'ought' strange and confusing, apparently, because it references concepts that are not analytically identical to those referenced by the copula 'is'. If it were, on popular accounts of what analyticity consists in, we could, by the law of substitution,

merely replace ‘ought’ with ‘is’ in the conclusion of the fallacious naturalistic argument and go on our merry way. But such a story about why we *don’t* substitute ‘is’ for ‘ought’ relies on our ability to clearly distinguish analytic from synthetic statements—that is, on our capacity to delineate meaning independent of factual content. If there is no clear distinction to be drawn between these two types of statements, then there must be another reason why we find the inference a strange one. It could be that only empirical statements of the *proper* kind, namely those informed and organized by an appropriate naturalized ethical theory, can productively inform a normative statement. But an admission that our logic can be informed by experience—that the laws of logic are open to revision in light of recalcitrant experience—amounts to an admission that the laws of logic are not analytic. Thus, our intuitions that Hume is on to something with the naturalistic fallacy are driven by either (a) implicit analytic/synthetic distinctions or (b) an inappropriate theory of naturalized ethics. Quine effectively undercuts (a), and the purpose of this book is to provide more support for a theoretically fecund notion of naturalized ethics, so (b) is not a threat to the project.

There is another sense in which Hume’s argument reduces to Moore’s argument. One could grant that it is illegitimate to make an inference from an ‘is’ to an ‘ought’, but only if, as Hume implicitly assumes, you do not *define* ‘oughts’ in terms of ‘is’ statements (e.g., “One *ought* to do what *is* pleasurable”). Hume’s argument then relies on Moore’s argument for its force: you can’t give a naturalistic definition of the good, and so the naturalistic fallacy will forever remain a fallacy.

The secondary literature on the naturalistic fallacy is large. However, it would be a fair summary to say that contemporary philosophers of a non-naturalistic stripe accept one version or another of either the Humean or the Moorean naturalistic fallacy. I will spend a good part of the remainder of this chapter outlining two possible responses to Hume and Moore. One draws on the explanatory resources of Quine, the other on a little-discussed account of moral reasoning proffered by Dewey. By my lights, Quine and like-minded philosophers such as Nelson Goodman and Morton White¹¹ make short work of the analytic/synthetic distinction. In doing so, they remove a crucial premise necessary for Hume and Moore to cleanly separate the empirical and the normative. Similarly, Dewey’s philosophical method tends to dissolve dualisms of

all kinds, including the analytic/synthetic distinction; nowhere is this clearer than in his discussion of means-ends reasoning. Though which means is most effective to a given end may be “merely” a matter for empirical demonstration, it may also be, if Dewey’s picture of moral judgment is at all correct, an empirical matter as to which ends we ought to have *simpliciter*.¹² Dewey and Quine are thus cozy bedfellows, which should come as no surprise since both fall under the pragmatist umbrella.

The upshot of Quine’s and Dewey’s responses to Hume and Moore will be that all of our beliefs, including seemingly analytic ones, are open to revision based on recalcitrant experience. If our beliefs are appropriately (that is, pragmatically) formed, so-called analytic statements are nothing more than extremely well confirmed scientific facts. Any attempt to argue that “come what may, we can never infer norms from empirical judgments,” as both Hume and Moore do, would entrench an indefensible assumption. We should therefore be open to the possibility of a reduction of normative properties to natural, functional properties.

Quine: Rejecting the Analytic/Synthetic Distinction

In “Two Dogmas of Empiricism,” Quine attacks two ill-founded beliefs that have conditioned the modern empiricist epistemological project. The first dogma is, of course, the analytic/synthetic distinction. The second is reductionism. The reductionism Quine attacks is not the kind of intertheoretic reduction that I am pressing. Rather, he attacks the reductionism of the logical empiricists, who thought that all meaningful statements were equivalent to logical constructs built out of terms referring to immediate experience. Quine would guardedly approve of the unity-of-science considerations that often drive both the articulation of traditional theories of reduction and more broadly ecumenical theories such as domain integration.¹³ I focus primarily on the first dogma, although, as Quine notes, the two are, at root, identical.

Quine first distinguishes between logically true analytic statements and other statements that appear to be analytic but do not obviously share the “logically true” status. An example of a logically true statement is “No unmarried man is married.” If we presuppose a class of

“logical particles” (e.g., truth-functional connectives such as ‘not’ and ‘and’), this statement remains true under any reinterpretation of its components (unless, of course, we reinterpret the logical particles themselves).

Quine later demonstrates that even the first class of logically true statements begs the question against the problem of analyticity. But we can set this concern aside for the moment to at least consider whether we can reduce the second class to the first so as to further constrain the bounds of the problem. Quine thus begins his argument with the second class of “analytic statements.” His example is “No bachelor is married.” At first glance, this statement seems analytic. But how can we demonstrate that it is? One strategy is to reduce this second class of statements to the first class by leveraging definitions. “Bachelor” is *defined* as “unmarried man,” so the second statement is actually equivalent, via substitution, to the first. To this, Quine responds “But who defined it thus, and when?” Appealing to dictionaries written by lexicographers begs the question, as those empirical scientists already had a standard for synonymy in mind—that is exactly why they listed ‘bachelor’ and ‘unmarried man’ next to each other in their dictionary. Thus, adverting to “definitions” does not adequately analyze the notion of synonymy to which friends of analyticity were appealing in the attempt to reduce definitional truths to logical truths.

An alternative explication of synonymy is to equate it with interchangeability. On this view, terms are synonymous if they can be interchanged without loss of truth value. Quine rightly notes that in this case we are concerned only with “cognitive synonymy,” not with psychological synonymy (e.g., terms can be cognitively synonymous with regard to the logical structure of the arguments they will support without necessarily calling to mind similar associations in you and me). According to Quine (1953, p. 158), “to say that ‘bachelor’ and ‘unmarried man’ are cognitively synonymous is to say no more nor less than that the statement: ‘. . . all and only bachelors are unmarried men’ is analytic.” Thus, this move is just question begging yet again. We still have no criteria for distinguishing this purportedly analytic statement from a statement that is true but only contingently so.

The final option that Quine examines for reducing statements of the second class of seemingly analytic truths to statements of the first logically true class relies on semantical rules. By examining and rejecting

this final option, Quine undermines any clean distinction between analytic and synthetic statements of *either* class, as logically true statements also lean heavily on the concept of a semantic rule.

One might think that it is only the sloppiness of ordinary language that prevents us from drawing a bright analytic/synthetic line. In an appropriately constructed artificial language, such as a good logic, can't we just define sets of semantical rules that stipulate what statements are analytic? However, as Quine quickly points out, such a move does not offer an analysis of analytical statements but instead solves the problem by fiat; stipulations and truths by fiat can, of course, be wrong. Perhaps then, we can merely add that such stipulations must be *true* stipulations. But this doesn't help, as that amounts to saying that *any* truth can be an analytic truth. Semantical rules would then be distinguished from the statements of (say) a true science merely because they happen to appear on a page under the heading "Semantical Rules" rather than in the "Well-Confirmed Experimental Results" section.

Quine concludes by noting the obvious fact that "truth in general depends on both language and extralinguistic fact." But, crucially, the belief that we can therefore somehow analyze a statement into a linguistic component and a factual component is, as Quine famously puts it, "an unempirical dogma of empiricists, a metaphysical article of faith" (*ibid.*, p. 163).

What of the second reductionist dogma? Quine argues that Rudolf Carnap's attempt to translate sentences about the physical world into sentences about immediate experience (in the technical sense intended by the logical empiricists—for example, that complexes of simple sentences of the form "Quality *q* is at point-instant *x;y;z;t*" will latch on to immediate experience and serve to ground all other sentences) implicitly relies on a language/fact distinction. The confirmation of a sentence leans heavily on the fact that one can distinguish the linguistic content of the sentence from the factual content supplied by the basic experience. But it was exactly the inability to demonstrate that such a thing is possible that led to Quine's abandonment of the analytic/synthetic distinction. Quine remarks: ". . . as long as it is taken to be significant in general to speak of the confirmation and infirmation of a statement, it seems significant to speak also of a limiting kind of statement which is vacuously confirmed, *ipso facto*, come what may; and such a statement

is analytic. The two dogmas are, indeed, at root identical” (ibid., p. 166).

Of course, Quine remains a good empiricist. He thinks, however, that our empiricism cannot make the simplistic assumptions required to get the project of *logical* empiricism off the ground. Rather, we should view belief formation more pragmatically. Each of us approaches the world armed with our theories (our “scientific heritage”) and an ongoing barrage of sensory stimuli. The considerations that guide us in warping our scientific heritage to fit our “continuing sensory promptings” are “where rational, pragmatic” (ibid., p. 168). All our beliefs exist in a web (including our theories about ethics, logic, and the various sciences),¹⁴ and we should not be so arrogant as to think that any of them, even the purportedly analytic ones (or normative ones), are immune to revision in light of experience.

Quine realized that his approach to philosophy would have tremendous implications for ethical theorizing. Indeed, he discussed his thoughts about the relationship between pragmatism and ethics in “On the Nature of Moral Values.” With Owen Flanagan, however, I think that Quine did not go far enough in allowing normative theories full play in our web of beliefs.

Quine, Hume, and Moore

Quine’s arguments interact with those of Hume and Moore in three significant ways.

First, as was discussed in chapter 1, both Hume and Moore rely in some respects upon a hard and fast analytic/synthetic distinction. If such a distinction cannot be supported, then there is reason to believe that the normative and the natural might be more closely related than they (especially Moore) argued. Recall particularly that Hume’s argument relies on Moore’s argument for its force. With Quine in hand, we can insist that any *a priori* attempt to isolate the good from natural definition dodges tough questions about theory change: rather than insist that the meaning of good precludes natural definition, why not admit that you have a theory of the good (rather than merely a definition of it), and let such a theory be adjudged as theories are: by their relationship to other theories, and by their encounters with experience?

Second, Quine's arguments also had an impact on *a priori* truth, at least insofar as analytic statements captured a large subset of those truths that could purportedly be justified without appeal to experience. If moral truths weren't those that could be known *a priori*, then we must come to have knowledge of them via experience, which opens the door for a robust empirical/normative interaction.

Third, Quine leveled the playing field with regard to an implicit hierarchy of things known—those things that were certain and were often known with certainty (the rules of logic, the truth values of definitional sentences, moral rules) were not categorically different from those things that were contingent and usually known contingently (the deliverances of the natural sciences). On the Quinean picture, theories about all these entities were conjoined and made responsive to experience. As a result, areas of inquiry that were not previously thought to be amenable to empirical interpretation, such as epistemology, were ripe for naturalization as the old hierarchies collapsed.¹⁵ Likewise for ethics.¹⁶

Dewey on the Naturalistic Fallacy and Moral Reasoning

John Dewey, one of the founders of modern pragmatism, anticipated much of Quine's work. Dewey was highly sensitive to dualisms of all sorts and the damage that they could do to our interests, particularly when they prevented us from expending our energies appropriately when dealing with our problems. Like Quine's, Dewey's logic was at root a compendium of empirically successful ways to deal with problematic situations; he did not have patience for those who would reify logic, making it a part of the formal structure of the universe that existed independently of reasoning creatures interacting with the world. His ethical theory, and the framework for moral judgment that constitutes its epistemological machinery, also eschews supernaturalism about the ethical and roots moral concerns in the activity of people coping with an environment. In this section, I will briefly discuss the basics of Dewey's moral theory, highlighting especially his appeal to the means-ends continuum, so as to sketch Dewey's conception of a science of morality. I will also gloss his theory of moral reasoning, which establishes the necessity of several crucial cognitive capacities that are especially amenable to connectionist reconstruction.

Dewey's general position on the naturalistic fallacy was that the "is/ought" gap did capture something about moral reasoning: that to articulate norms consisted in discussing intelligent methods of regulating consummatory experience. But Dewey did not think that this implied that there could be neither a science of ethics nor a naturalistic explanation of the ontology of the good and how we comprehend it and regulate it. Crucially, Dewey distinguishes between the desired and the desirable. The presence of a desire for dessert does not mean I ought to eat the dessert; to do so would be to improperly balance my *desire* for sweet food with something *desirable*, namely maintaining a healthy body. In the short term, regulating my experience by giving in merely to what is desired rather than to what is desirable would be disastrous and would lead to non-consummatory experience in the long run. I should regulate my desires and resolve conflicting wants and needs, or I should triangulate a reasonable course of action when faced with apparently conflicting values. The reasoning process that I use to regulate action in this way is the moral reasoning process.¹⁷ But such a process does not rely on a *supernatural* capacity to identify *pre-existing* "eternal norms." And neither does the fact that I have desires on which I ought not act preclude my using positive moral experience as a fallible basis for generating norms and "oughts." Dewey's general approach to ethics is thus consistent with his naturalistic humanism, and with his appreciation for evolutionary theory.¹⁸ As Dewey notes in his introduction to *Human Nature and Conduct* (1922, p. 12), "a morals based on study of human nature instead of upon disregard for it would find the facts of man continuous with those of the rest of nature and would thereby ally ethics with physics and biology."

There is some disagreement in the small secondary literature on this matter. Marga Vicedo (1999, p. 234) insists, using strong language, that Dewey would approve of an evolutionary ethic,¹⁹ whereas Paul Lawrence Farber (1994, p. 113) argues that Dewey rejects evolutionary approaches to ethics as "fundamentally misguided." Scholars such as Farber often support their contentions with quotations from the first edition of the *Ethics* (1908). But a close reading of the second edition (1932) reveals that much of the controversial language that can be construed as eliminating in principle an evolutionary ethic has been removed. Moreover, examination of the context of the remarks in the first edition reveals

that they are intended as criticisms of *existing* systems of evolutionary ethics, mainly those proposed by Darwin and by Spencer. Finally, Farber draws mostly upon Dewey's early work, which is tainted with a Hegelian residue from Dewey's early philosophic training. Though Dewey learned the theory of evolution in college and believed it to be accurate, it took almost 10 years for the import of it to leach into his philosophy. Dewey makes several precautionary remarks regarding an evolutionary ethic, but in view of his general approach of having Darwinian considerations inform philosophy *en toto* we have *prima facie* reason to believe that Dewey would be amenable to an appropriately formulated evolutionary ethic.

For Dewey, organisms like ourselves engage in inquiry when we are faced with problematic situations. Such organic, "lived" problems are what spark reflection and issue in choice. Thus, in moral inquiry there are three predominant stages: (1) an agent finding herself in a morally problematic situation, which leads to (2) moral deliberation involving experimental, emotional, and imaginative processes, which then issues in (3) a judgment, choice, or an action. Though all three of these phases are crucial, of particular interest for this chapter is moral deliberation as it relates to imagination.

Dewey on Moral Imagination

Dewey thought that, if we applied ourselves, we would come to regulate our activities intelligently so as to provide an optimum amount of consummatory experience. Although the world (e.g., organisms and environments) contains both value and disvalue, and although we cannot hope to alleviate the latter entirely, we can certainly ameliorate our situation, improving it as much as possible.

Language such as "consummatory experience" should not lead one to think that Dewey or other pragmatists were concerned with maximizing *subjective* happiness or pleasure. For Dewey, values are part of the world-organism relationship, and, owing to the facts of our biology and our evolutionary history, we can come to discover them (although this is not to say that they were there before the organism was). David Brink (1989) argues that accounts of value that make values subjective (such as hedonistic or desire-satisfaction theories) fall prey to a fatal *gedanken*

from Robert Nozick. If we had an “experience machine” that we could connect to our brains so as to provide continual satisfaction of our desires, none of us would choose to connect ourselves to this machine. This belies the fact that value is not merely a reflection of our subjective desires but involves interaction with a world that contains value.

One important cognitive method we use to hold an end in view so as to ascertain the consequences of its pursuit (and fix effective means to achieve it) is imagination. The capacity to imagine is crucial for moral reasoning on Dewey’s account. In *Human Nature and Conduct* (1922, pp. 132–133), Dewey explains:

Deliberation is an experiment in finding out what the various lines of possible action are really like. It is an experiment in making various combinations of selected elements of habits and impulses, to see what our resultant action would be like if it were entered upon. But the trial is in imagination, not in overt fact. The experiment is carried on by tentative rehearsals in thought which do not affect physical facts outside the body. Thought runs ahead and foresees outcomes, and thereby avoids having to wait the instruction of actual failure and disaster. An act overly tried out is irrevocable, its consequences cannot be blotted out. An act tried out in imagination is not final or fatal. It is retrievable.

Though at first glance it might appear that Dewey is merely referring to our ability to model events in the world, he is doing more than this, as he has very subtle accounts of what it means to possess a habit. Habits for Dewey are rich cognitive and conative capacities that are influenced by experience and, in turn, influence what we make of experience. Later I will argue that Dewey has in mind a complex of cognitive capacities when he speaks of imagination, only some of which include our ability to engage in mental modeling, and all of which are amenable to connectionist interpretation. In some cases, Dewey’s language anticipates radical connectionist, sub-symbolic, and dynamical systems theory approaches to situated action; in addition, some of the otherwise strange language that he uses when describing moral reasoning and character development can be viewed as an anticipation of developments in the cognitive neuroscience of judgment and decision making. Dewey’s account and these influences and connections will be explored in more depth in chapter 5.²⁰

For the time being, the important thing to note is the existence of a fluid continuum in this picture of moral reasoning between means and ends. A trivial example: I have a quite natural and possibly appropriate

desire for ice cream; ice cream is of value to me. I hold fixed this end in view so as to imagine the consequences associated with the consumption of the ice cream. I discover that there are many possible futures wherein I gain an unhealthy amount of weight, and I discover also that in those circumstances many other things I value as consummatory experience would not be available to me—I could no longer fit into the cockpit of my stunt airplane, say, and there is a good chance that I would suffer a heart attack owing to arterial sclerosis. I choose instead to eat an apple, and as I eat apples rather than ice cream I come to enjoy the experience of apple eating and focus approvingly upon it, making it a habit. I react to apples differently now (“Oh, an apple! How delightful!”) and have different experiences around them as a result of my initial encounter with and cognition about ice cream.

In this case, my moral imagination has caused me to transform an end in view (consumption of ice cream) into a different end in view (consumption of apples), which at first I conceive of as merely a *means* to the end (remembering that ends are something desirable and not merely desired) of health, but which I eventually transform into an end *in and of itself* also. I finally get in the habit of eating apples, and such a habit is not *merely* the repetition of a bodily movement but rather a rich set of cognitive experiences that transforms my daily activity into something quite different than it was before. A better example: think of exercise. It is no accident that this process of habituation (*richly* construed) is essentially a character-development activity. On this view, the sets of capacities we gain by reasoning morally are more accurately characterized as sets of cognitive skills and habits rather than as linguistic knowledge as such.

Ends become means and means become ends. This process of transformation demonstrates that, according to Dewey, we do an injustice to the world if we construe ends as being fixed, permanent, final and out of the reach of a scientific analysis. Most people look upon engineering as an applied science, and would view it as an expertise that focuses on means, yet we have no bitter ontological struggles about engineering (at least, none that make their way into common parlance, quite unlike ethical ontologies). The transformation of one thing formerly valued as an end into something that is merely a means for another end, and the reverse transformation of ends into means (e.g., at first I enjoy going to the library because I like to read, but later reading becomes a means to

enable me to acquire the skill of being able to philosophize) demonstrates that the fact/value distinction is not hard and fast, but rather is one of degree.

In the perfect world, all experience would be continually consummatory. Note that in the analysis of moral function in the next chapter this amounts essentially to being perfectly adapted to the range of environments with which you regularly interact. Note also that if your environment is perfectly stable, being perfectly adapted would abnegate the need for creative abstract thought. On some pictures, if this world were simple enough, cognition would altogether cease to have a function. Peter Godfrey-Smith's 1996 book *Complexity and the Function of Mind in Nature* contains an excellent discussion of these issues as well as illustrative treatments of both Dewey and Herbert Spencer. But, since we do not exist in a perfect world, not all experience is consummatory. However that is not to say that norms can't be grounded in empirical facts about human flourishing, nor is it to say that ends can never be means and vice versa.

In line with these thoughts, Dewey's account is both normative and empirical. It is normative insofar as it represents the way we ought to think about moral matters (that is, in a scientific spirit), and it is empirical insofar as Dewey thought that this was the way we do in fact proceed when engaging in fruitful moral inquiry. It is naturalistic through and through, and the open-question argument and naturalistic fallacy find no purchase on it.

Dewey, Hume, and Moore

The open-question argument merely amounts to a description of one crucial phase of moral experimentation, namely that of testing ends in view to see if they should be adopted as ends proper. However, nothing about this process implies that ends are metaphysically strange or that they are not facts about creatures and environments and their relationships. There is a singular, crucial difference between Dewey's method and Moore's: despite Moore's lament that philosophers too often engage in purely speculative metaphysics, the open-question process at its best is still basically a form of non-empirically informed conceptual analysis. At its worst, it can legitimize armchair metaphysics (as in: not

only must we proliferate moral ontological simples, but perhaps there are open questions about every concept at every turn). Dewey, however, intends for moral reasoning to be empirically informed. On his picture, it has a scientific aspect that is missing from Moore's "open questioning." And, as discussed earlier, open-question arguments implicitly rely on the analytic/synthetic distinction, which Dewey, anticipating Quine, rejects as yet another ill-advised dualism.

As for Hume's naturalistic fallacy, Dewey's process of moral reasoning will, he thinks, help us identify those extant values that are worthy of pursuit. These values, though, are discovered by examination of the biological world of organism-environment interaction: they are facts, empirical matters in any reasonable sense of the phrase. Dewey's ethical theory has many points in common with Hume's,²¹ although once the teleological aspects of Aristotle are canalized and given limits by a biological analysis of function, we will see that Dewey's project is actually much more like a modern-day virtue theory.

A Pessimistic Coda: Why This Project Is Still Important Even If This Chapter Is All Wrong

Even if Quine, Dewey, and I haven't convinced you that the naturalistic fallacy and open-question arguments do not stand in the way of attempts to sketch a naturalistic account of the content of morality and the form of moral judgment, you still have reason to keep reading. Only the most stalwart anti-naturalist would think that facts about human beings and how they reason have absolutely no bearing on normative concerns, and only a small number of contemporary moral philosophers have taken this position. Even if this chapter seems misguided, we can at least maintain that the biological and cognitive sciences can constrain moral theorizing by identifying the realistic limits of our biological and moral capacities.

Usefully, we can sketch out three possible personality types that embody sets of positions regarding the relationships between science and the norms of morality (since the question is ultimately one of governance, I have used political terms): Separatists, Confederates, Unionists.

Separatists advocate abstinence: there shall be no intercourse between the findings of science and the articulation of norms. What is would be

irrelevant to what ought to be; the methods of the sciences would be orthogonal (at best) to the formulation of norms, and there would be no common ground between science and morality. Virginia Held, Kelly Nicholson, and Alvin Plantinga are modern-day separatists.

Confederates are moderately promiscuous: they allow the findings of the sciences to place limits on the demands that norms can legitimately place upon us, or to rule out some moral theories as inconsistent with our best natural knowledge. James Sterba and David Brink are contemporary Confederates.

The fecund Unionists are of two stripes. There are those who think that robust moral norms are part of the fabric of the world and can be constrained by and derived from the sciences. These are the “Conservative Unionists,” who wish to subsume ethics by making it into a science. Mark Johnson and Larry Arnhart are Conservative Unionists, as is Owen Flanagan. Sharing similar views about the relationship between science and morality, but disagreeing about what the sciences will tell us about moral nature, are the “Eliminative Unionists,” who wish to “unify” science and ethics by eliminating the purportedly illusory subject matter of ethics. Michael Ruse is presiding president of this party; J. L. Mackie is past president, and E. O. Wilson is vice-president.

The point of this section was to make a plausible case for Conservative Unionism. (I will deal with the complications presented by Eliminative Unionism in the next chapter.) Though the inertia of the history of moral philosophy is against Conservative Unionism, the party platform has much to offer. But even if you remain a Confederate, the remainder of this book will be very useful, as it will identify constraints placed upon our normative moral theories by the results of the cognitive and biological sciences. If you are still a Separatist, then it will at least be a provocative read. But I would hasten to point out that your party is growing smaller and more disorganized day by day. The future lies with Conservative Unionism and consilience. Nothing about the term ‘Conservative’ is meant to imply that the viewpoint won’t be progressive. It will be; rather, it merely indicates that the party wishes to maintain the general moral stance, identifying parts of the ethical tradition that are especially useful in view of the findings of science. The view will not be radically eliminative, but neither will all moral concepts be maintained.

Conclusion

I have argued in this chapter that cognitive naturalists about morality have often been stymied in their attempts to fruitfully unify ethics and the sciences by the two non-reductive roadblocks of the naturalistic fallacy and the open-question argument. However, both of these positions rely upon the analytic/synthetic distinction for their force, and the arguments of Quine give us good reason to doubt that such a hard and fast distinction exists. In addition, the theory of moral judgment on offer from Dewey belies the fact that facts and values intermingle and correlate in ways subversive to both roadblocks. “Conservative Unionism” about the relationship between science and norms remains a live option. In the next chapter, I give content to the party platform by outlining a neo-Aristotelian conception of function that is biological and is naturalistic through and through.