The Ethical Project

Copyrighted material

Copyright © 2011 by the President and Fellows of Harvard College
All rights reserved
Printed in the United States of America

First Harvard University Press paperback edition, 2014

Library of Congress Cataloging-in-Publication Data

Kitcher, Philip, 1947-
The ethical project / Philip Kitcher.
p. cm.
Includes bibliographical references (p.) and index.
ISBN 978-0-674-06144-6 (hardcover : alk. paper)
ISBN 978-0-674-28428-9 (pbk.)
1. Ethics, Evolutionary. I. Title.
BJ1311.K53 2011
171'.7—dc22 2010051684

Copyrighted material

Contents

Introduction

§1. The Shape of Things to Come

§2. Methodological Preliminaries

I. An Analytical History

1. The Springs of Sympathy

§3. Psychological Altruism: Basics

§4. The Varieties of Altruistic Reactions

§5. Some Dimensions of Altruism

§6. Maternal Concern

§7. Broader Forms of Altruism?

§8. Possibilities of Evolutionary Explanation

§9. The Coalition Game

2. Normative Guidance

§10. The Limits of Altruism

§11. Following Orders

§12. Punishment

§13. Conscience

§14. Social Embedding

3. Experiments of Living

Copyrighted material

§15. From There to Here
§16. Cultural Competition
§17. The Unseen Enforcer
§18. Some Dots to Be Connected
§19. Divisions of Labor
§20. Roles, Rules, and Institutions
§21. Altruism Expanded
4. One Thing after Another?
§22. Mere Change?
§23. Three Ancient Examples
§24. Second-Sex Citizens
§25. Repudiating Chattel Slavery
§26. The Withering of Vice
§27. The Divine Commander
II. A Metaethical Perspective
5. Troubles with Truth
§28. Taking Stock
§29. Prima Facie Problems
§30. Truth, Realism and Constructivism
§31. The Sources of the Troubles
6. Possibilities of Progress
§32. The Centrality of Ethical Progress
§33. Generalizations from History
§34. Problems, Functions and Progress
§35. Modes of Refinement
§36. Functional Generation
§37. Local and Global Progress
§38. Ethical Truth Revisited
§39. Residual Concerns
7. Naturalistic Fallacies?
§40. Hume's Challenge
§41. Authority Undermined?
§42. Troublesome Characters
§43. Settling Disputes

Copyrighted material

III. A Normative Stance

8. Progress, Equality, and the Good
§44. Two Visions of Normative Ethics
§45. Dynamic Consequentialism
§46. Failures and Successes
§47. From the Local Community to the Human Population
§48. Equality and the Good Life
§49. Population Size
§50. Aspects of the Good Life

9. Method in Ethics
§51. Varieties of Ethical Change
§52. Method and the Good
§53. Mutual Engagement
§54. Ethical Debate
§55. Dissent and the Limits of Tolerance
§56. The Challenger Revisited

10. Renewing the Project
§57. Philosophical Midwifery
§58. Scarce Resources
§59. Habits and Their Limits
§60. Conflicting Roles
§61. Ethically Insulated Spheres
§62. Maintaining Equality
§63. The Challenges of Technology

Conclusion
§64. Summing Up

Acknowledgments

Index

Copyrighted material

Moral conceptions and processes grow naturally
out of the very conditions of human life.
—John Dewey

Copyrighted material

Introduction

§1. The Shape of Things to Come

Ethics pervades every human society and almost every human life. People deliberate about what they should do on specific occasions, about what is worthwhile, about the kinds of lives they should aspire to lead. In subtle ways, their everyday actions presuppose habits of conduct, roles and institutions current in their societies, endorsed sometimes after serious reflection, often accepted without much thought. With the exception of those afflicted with psychological disruptions that profoundly limit their cognitive capacities or that cut them off from their fellows, we are all embedded in the ethical project.

Yet for ordinary people, and for philosophers too, the status of our ethical judgments and practices is hard to fathom. What exactly do we mean when we praise someone for a correct decision? How could an evaluation like that be grounded? Bertrand Russell famously described mathematics as a subject in which ". . . we never know what we are talking about, nor whether what we are saying is true."[1] The impressive ability of mathematicians to reach agreement on conclusions that endure fosters confidence in their power to acquire knowledge, despite the mysteries swirling around the content and grounds of their judgments. By contrast, the persistence of ethical debate reinforces a sense of unease about the status of ethical practice, often leading those engaged in public controversies to shy away from "value judgments," as if any hope of reaching consensus were absurd.

One popular view of the ethical life persists. Many people believe, with

Copyrighted material

Dostoyevsky's Ivan Karamazov, that if ethical precepts were not grounded in God's commands, anything would be permitted. From Plato on, however, the philosophical tradition has frequently—and cogently—questioned the idea of a religious foundation for ethics. Supposing ethics to be grounded in the divine will remains popular because alternatives, including the philosophical alternatives, appear so elusive and unconvincing. Could ethical correctness really consist in representing some independent realm of values? Could ethical judgments really express particular privileged emotions? Could they really be arrived at by fathoming the "moral law within" or by apprehending the deliverances of practical reason?

More than a century ago, Darwin outlined a novel way of thinking about the living world: his fundamental insight was to regard the organisms around us as products of history. We can liberate ourselves from mysteries about many of our current practices by emulating Darwin: think of them, too, as historical products.[2] The aim of this book is to pursue this program in the case of ethics. Ethics emerges as a human phenomenon, permanently unfinished. We, collectively, made it up, and have developed, refined, and distorted it, generation by generation. Ethics should be understood as a *project*—the ethical project—in which we have been engaged for most of our history as a species.

The position to be elaborated—*pragmatic naturalism*, to give it a name—envisages the ethical project as begun by our remote ancestors, in response to the difficulties of their social life. They *invented* ethics. Successive generations have amended the ethical legacy transmitted to them, sometimes, but by no means always, improving it. Doubtless, many traditions have died out, but some have continued into the present, forming the bases of the ways in which people today regulate their conduct. In principle, but not in practice, it would be possible to construct an evolutionary tree, drawing a diagram like the single figure Darwin inserted into the *Origin*, with the important difference that the connecting lines would represent cultural rather than biological descent (so the picture might show fusion, as well as separation of lineages).[3]

As the name suggests, pragmatic naturalism has affinities with both pragmatism and naturalism. In focusing on ethical practice and its history, it attempts to honor John Dewey's call for philosophy to be reconnected with human life.[4] Further, it articulates a Deweyan picture of ethics as growing out of the human social situation; its conception of ethical correctness is guided by William James's approach to truth.[5] The naturalism consists in refusing to introduce mysterious entities—"spooks"—to explain the origin, evolution, and progress of ethical practice. Naturalists intend that no more things be dreamt of in their philosophies than there are in heaven and earth.[6] They start from the inventory of the world allowed by the totality of bodies of well-grounded knowledge (the gamut of scholarly endeavors running from anthropology and art

Copyrighted material

history to zoology), and, aware of the certain incompleteness of the list, allow only such novel entities as can be justified through accepted methods of rigorous inquiry. Appeals to divine will, to a realm of values, to faculties of ethical perception and "pure practical reason," have to go.

Pragmatic naturalism engages with the religious entanglement of ethics more extensively than is usual in secular philosophical discussion—for the pragmatist reason that the entanglement pervades almost all versions of ethical life. Yet, in accordance with its naturalist scruples, it cannot maintain the image favored by those who would ground ethics in the divine will. As we shall see (§27), there are powerful reasons to suppose, even if there were any deity, *ethics* could not be fixed by its (his? her?) tastes. More fundamentally, pragmatic naturalism maintains that, when *religion* is understood as a historically evolving practice, it is overwhelmingly probable that all the conceptions of a transcendent being ever proposed in any of the world's religions are false. For the conceptions introduced in the various religions are massively inconsistent with one another. Each supernaturalist view rests on epistemically similar grounds—typically there was some revelation, long ago, that has been carefully transmitted across the generations to the devout of today—yielding a condition of complete symmetry. Under these circumstances, no believer has any basis for thinking only he and his group are privileged to know the truth about the transcendent realm, while others live in primitive delusions. Further, serious inquiries into the ways in which canonical scriptures are constructed, into the evolution of religions, into the recruitment of converts, into the phenomena of religious experience, demonstrate how radically unreliable are the processes that have yielded the current corpora of belief. Nor can one isolate some core doctrine, shared by all religions, something capable of being viewed as a shared insight. If there are beings of a hitherto unrecognized sort, approximating some idea of the "transcendent," we have every reason to think we have absolutely no clues, or categories, for describing them.[7]

Religious entanglement in ethical practice is no accident. As we shall see (§17), appealing to gods as "guardians of morality" can bring social benefits. Nevertheless, that appeal has distorted the ethical project. Undoing the distortions is not simply a matter of eradicating religion, hacking out the places where false belief has intruded. A secular renewal of the ethical project requires constructive work, positive steps going beyond brusque denial.[8]

Given these clarifications, I can now explain the structure of the following chapters. Part I, Chapters 1–4, elaborates an "analytical history," aimed at providing insight into the evolution of our ethical practice. It provides a basis on which Part II (Chapters 5–7) can explore questions *about* ethics: given this account of the origins and unfolding of ethics, can we make sense of ethical truth or ethical knowledge? The history of Part I and the metaethical account of Part II are

Copyrighted material

then extended, in Part III (Chapters 8–10), into a normative stance, an attempt to suggest how we might best go on from where we are.

It is worth supplementing this bald characterization with a little more detail. A "history of ethical practice" might take many forms, and the one I offer may initially appear strange. Since I suppose our species to have been engaged in the ethical project for tens of thousands of years, it would be hopeless to offer a narrative showing how particular aspects of ethical life have gradually emerged. Until the invention of writing (five thousand years ago), the clues are fragmentary, far scantier than the fossil record, whose poverty provoked Darwin's lament.[9] Primatology, anthropology, and archeology enable us to offer a plausible account of the conditions under which our preethical ancestors lived, but many subsequent steps are beyond our evidential grasp.

The analytical history starts by attempting to understand relevant psychological capacities of the preethical ancestors, and, on that basis, to portray the initial stages of the ethical project. Hominid social life was akin to the contemporary lives of our closest evolutionary relatives: our precursors lived in small groups, mixed by age and sex. For that, they needed a capacity for psychological altruism. Yet the limitations of their altruistic dispositions made living together tense and difficult. The first ethicists overcame some of the problems by agreeing on rules for conduct, rules remedying a few of the recurrent altruism failures that had plagued their group life. Very probably, they began with precepts about sharing scant resources and not initiating violence.

Because the character of early ethical practices is so much simpler than the forms of ethical life visible once written documents are available, it is important to show (Chapter 3), how a series of gradual steps *might* have taken the ethical project from its relatively crude initial phase to the complex articulation of rules and stories found in the first written documents. Thereafter, it is possible to trace, although not with the completeness one might hope for, how *actual* changes in ethical practice have occurred. Chapter 4 considers a few examples from history (rather than prehistory) with the aim of supporting two main theses. First, it is hard to resist the recognition of occasional progress in the evolution of ethics: perhaps ethical progress is rare, but there are transitions (like the repudiation of slavery) in which it seems to occur. Second, even when the records kept by people who participated in apparently progressive ethical change are most extensive, moments of ethical discovery are elusive: there are no analogs of episodes of scientific insight.

The history of Part I offers hypotheses about how the ethical project *actually* began, and how, in recent history, it has *actually* gone. It also addresses concern about the vast difference between the early stages of the project and the rich practices found at the dawn of history by showing how it would have been *possible* for the bare beginnings to evolve, by gradual steps, into the complex systems

Copyrighted material

discernible in the earliest texts. Because the differences in these two modes of explanation need to be clearly appreciated, the next section will address some methodological preliminaries.

How can any history, however carefully focused and articulated, bear on philosophical questions about ethics? One possibility, already illustrated by the example of religion, is that a historical account might undermine current practice. Seeing where our approaches have come from could breed skepticism and disillusionment. In those episodes of ethical change most susceptible to analysis, the participants do not appear to apprehend some previously unrecognized value, or to reason their ways to some novel moral principle. Historical detail, to the extent it can be provided, is inhospitable to philosophical theories about ethical truth (Chapter 5). Yet the history of Part I also reveals examples of ethical progress. The metaethical perspective of Part II centers on trying to reconcile these points.

A “mere change” view of ethical evolution, in which the history is simply one damned thing after another, conflicts with the pull to characterize some transitions as advances. Chapter 6 resolves the conflict by seeking an account of ethical progress, one that abandons the idea of progress as accumulation of (prior, independent) truth. If this appears a strange idea, we should recall that, in some areas of human practice, progress does not consist in the increase of truth. Technological progress is often a matter of discharging certain functions more efficiently or more fully. Moreover, in line with the history of Part I, the initial ventures in the ethical project are readily conceived as introducing a new—social—technology, aimed at remedying disruptive altruism failures.

Amelioration of altruism failure was the initial function of ethical practice. Yet the obvious differences between the pioneering ventures and the complex codes present at the dawn of recorded history show clearly that other functions have emerged. That is the way with technology in general. People begin with a problem and achieve partial successes in solving it. The successes generate new problems to be solved. Chapter 6 attempts to anchor the concept of ethical progress in the discharging of functions, originating with the problem of remedying failures of altruism, and understanding later functions as generated from the solutions previously obtained.

It thereby paves the way for a concept of ethical truth. Ethical truths are those acquired in progressive transitions and retained through an indefinite sequence of progressive transitions. Pragmatic naturalism proposes that some ethical statements—typically, vague generalizations, commending honesty and disavowing violence, for example—are true. They owe their truth to the role they play in ethical progress: “truth happens to an idea.”[10]

To declare that our ancestors *invented* ethics is to deny that they *discovered* it or that it was *revealed* to them. Pragmatic naturalism rejects the idea of a special

Copyrighted material

moment (long ago on Mount Sinai, perhaps) when people received authoritative information about how they should live, and also abandons surrogate philosophical theories about external constraints discovered by special faculties. Yet to declare that ethics is a human invention is not to imply it was fashioned *arbitrarily.* The ethical project began in response to central human desires and needs, arising from our special type of social existence.

There is an obvious concern, one probably already exercising any reader suspicious about misadventures in naturalism. Why should the ethical project, even at its most "progressive," have force on those people who appear late in its evolution? Critics have often charged that naturalism commits a fallacy, and those criticisms need to be addressed. Chapter 7 considers a number of versions of the accusation, attempting to show that pragmatic naturalism has resources equal to those of any non-naturalistic rival. Yet one important version persists. Because the ethical project generates new functions, not necessarily in harmony with one another, it appears to leave open the possibility of different ways of continuing. To settle worries about radical disagreement, to finish articulating the metaethical perspective, a normative stance is needed. That is the work of Part III.

Parts I and II portray the ethical project as an enterprise in which people work out how to live together. It began without presupposing any sources of truth ethical deliberators sought to fathom—whether those sources lay in a divine will or in any of the philosophical substitutes. Our ancestors needed to fashion their shared life by conversing on equal terms. Pragmatic naturalism denies ethical expertise. The role philosophy plays in ethics can be one only of midwifery: to suggest a direction for renewed conversation and some rules for mutual exchange. Chapters 8 and 9 offer a package of proposals, an egalitarian conception of the good at which we should now aim, together with a method of deliberating under conditions of mutual engagement.

Convincing proposals come with some form of support. Pragmatic naturalism's proposals are motivated by conceiving the current human situation as analogous to that initially prompting the ethical project. As it was in the beginning, so too now—for the conflicts to which our ancestors' lives were subject are mirrored in contemporary hostilities across the human population. According to this vision, the original function of ethics—to remedy altruism failures—remains primary. Challenging the enduring importance of this function fails to achieve an important form of coherence, one pragmatic naturalism attains. So (§56) the deepest challenge—the most important accusation against naturalistic ethics—is turned back.

Finally, Chapter 10 attempts further philosophical midwifery, by suggesting some specific places at which current ethical practices might be amended. It offers ideas for continuing the essential conversation.

Copyrighted material

Readers familiar with contemporary philosophical discussions about ethics will recognize the large differences between standard approaches and the material I plan to cover. Although my questions connect with those posed constantly in the history of ethics, they are often framed orthogonally to the preferred philosophical formulations of recent decades (at least in the Anglo-Saxon world). At many points, my treatment has been influenced by the writings of my peers (and betters), but any serious attempt to expose connections (especially one attending to the nuances of intricate positions) would require expanding my discussions by hundreds of pages. Pragmatic naturalism aims to steer ethical practice and ethical theory in new directions, and I apologize to those who would like to see it anchored in the "existing literature." I must also ask for patience, if a response to an important worry or objection is postponed. Not everything can be done at once, and the needed resources have to be assembled before they can be applied.

§2. Methodological Preliminaries

Pragmatic naturalism differs from previous attempts to link ethics to our evolutionary past. It does *not* propose to identify ethical properties in evolutionary terms, say, by equating what is good with what is adaptive.[11] Nor does it suppose ethical practice is already present, at least in embryo, in our evolutionary cousins or our hominid progenitors.[12] The ethical project is not simply the unfolding of previously existent altruistic tendencies—it is more than just a population acquiring capacities for "nice behavior." Ethical practice involves conversation, with others and with yourself, juxtaposing desires you recognize as part of you and other desires you would prefer to move you to action. Neither does it posit a special evolutionary advance, in which our ancestors acquired a "moral instinct," conceived along the lines of our innate capacity for language.[13] Views of this type are in danger of confusing ethics with "nice behavior," and, as we shall see (§14), they underplay the influence of the social environments in which ethical practice occurs.[14]

Given these differences, some familiar methodological concerns about naturalistic ethics do not bedevil pragmatic naturalism. One major worry does arise, however. Because it is important, and because appreciating its significance could easily provoke objections to discussions in Part I, this section endeavors to forestall it.

Darwin's success in applying his historical method to the living world rested on the immense body of evidence amassed in the *Origin*—even hostile reviewers praised him for "his facts." A familiar criticism of later attempts to apply evolutionary ideas to human behavior and to human social life charges Darwin's imitators with failing to live up to the standards set by the master. Evolutionary

Copyrighted material

explanations, it is suggested, run the risk of becoming exercises in storytelling—just-so stories without Kipling's wit.[15] A possible hypothesis for the evolution of some trait is proposed, and, without seriously considering alternatives, discriminating them in light of the evidence, it is adopted as describing the actual course of past events.

Reconstructing the *actual* history of the ethical project, from its beginnings to the present, is plainly beyond the evidence available—and probably beyond the evidence anyone could ever hope to obtain. On the account sketched in §1, and developed in Part I, many important changes occurred in the Paleolithic and early Neolithic, in a period lacking any written documents. Visions of human life at that time can be based on just a few tantalizing clues: deposits showing increasing group size, tools found at a far remove from the closest source of material, burial sites, figurines, and cave art. These data are too sparse to screen out rival hypotheses about the sequence of events leading from the beginnings of the ethical project to the complex form in which it appears in the first written texts we have.

Yet some hypotheses about the actual history can be defended. Archeological evidence provides grounds for thinking that, until about fifteen thousand years ago, human beings lived in groups of roughly the size of contemporary bands of chimpanzees, and, like the societies of our evolutionary cousins, these groups were mixed by age and sex. Combining this conclusion with anthropological studies of living people who are closest to the circumstances of our ancestors, researchers have provided a picture of the social context of our forebears, the context in which they began the ethical project. We know what the problems faced by chimpanzees are, and how contemporary bands of hunter-gatherers overcome them. The first ventures in ethical practice probably involved group discussions, on terms of rough equality, directed toward issues of sharing and intragroup aggression.

Those initial efforts likely occurred tens of thousands of years ago, after human beings acquired full linguistic abilities. At a conservative estimate, we can set the age of the ethical project as fifty thousand years. I conclude that roughly the first forty thousand years of the project were directed toward the needs of small groups, whose members worked out their social lives on terms of approximate equality. Their original rules were crude and simple. Out of the social life they permitted came a sequence of dramatic changes, generating societies of greater size, and eventually the hierarchical cities of Mesopotamia and Egypt. That part of the story, I shall argue, can be defended as *actual* history.

How did our ethical ancestors move from those simple beginnings to the complex forms of ethical life recognizable in the world of ancient history? No answer can claim to tell the actual story. Here pragmatic naturalism must face the difficulty of discriminating a preferred story from potential rivals. It is,

Copyrighted material

however, committed to supposing some sequence of transitions led—without revelations, without discoveries of the structure of values or the moral law, without any "spooks"—to an endpoint enormously richer and more complex than the original practices. To answer skeptics claiming that "real ethics" requires resources naturalists cannot allow, some narrative needs to be given. It cannot be advertised as a "how actually" explanation; instead it is a "how possibly" explanation.

Explanations come in many varieties, but for the purposes of this book, these two types will suffice. A historical "how actually" explanation aims to tell the truth about a sequence of events: if it is properly supported, rival options have to be eliminated by the evidence. A historical "how possibly" explanation, by contrast, aims only to tell a story, consistent with the evidence and with background constraints: its status is not impugned by pointing out that there are other options (the more, the merrier). A "how possibly" explanation is important because we sometimes wonder whether a chain of occurrences *could* have occurred, or whether the occurrence of the sequence is permitted by a particular theory. Opponents wonder, for example, if the processes countenanced by Darwin and his successors allow for the evolution of the cell. Answering their doubts requires showing how Darwinian processes *might* have produced the cell. It would be marvelous, of course, to be able to say how the history actually went, but, given the temporal remoteness of the events and the limitations of our evidence, modesty is required. In the context of rebutting the skeptical challenge, modesty —settling for "how possibly"—is enough.

Pragmatic naturalism can advance probable hypotheses about the original state in which the ethical project began, and about the character of the evolution of the project during recorded history. With respect to the transformations that occurred between the early phases and the practices of the ancient world, all that can be claimed is that these *could* have happened without supposing processes or causes of kinds pragmatic naturalism rejects. The history of Part I is self-conscious about the distinction. I have attempted to be clear about the kinds of explanations at which particular discussions aim—and hence the standards they are expected to satisfy. In consequence, the metaethical perspective of Part II can sometimes build on claims about *how* a process unfolded (delivered by "how actually" explanations), and sometimes only on theses *that* particular initial states evolved into particular final states (defended against skeptical challenge by "how possibly" explanations). Explicitness may be pedantic—but pedantry is probably better than arousing suspicion that proper standards have been violated.

The aim is to use history—in the ways, and to the extent, we can reconstruct it —to liberate discussions of ethics from the confining pictures that prompt a sense of mystery. Let us now turn to the work of reconstruction.

Copyrighted material

1. “Recent Work on the Principles of Mathematics,” *International Monthly,* 4, 1901, 84. I shall not venture into the controversy about how seriously Russell intended his characterization.

2. Darwin himself made some first efforts to inaugurate this program in the early chapters of *The Descent of Man* (John Murray, 1872). Thomas Kuhn proposed a similar understanding of the natural sciences as historical products (The *Structure of Scientific Revolutions* [Chicago: University of Chicago Press, 1962 and 1970]). I have followed their lead, both for the sciences and for mathematics (in *The Advancement of Science* [New York: Oxford University Press, 1993] and in *The Nature of Mathematical Knowledge* [New York: Oxford University Press, 1983]).

3. Charles Darwin, *On the Origin of Species* (facsimile of the first edition), edited by Ernst Mayr (Cambridge, MA: Harvard University Press, 1964), between 116 and 117.

4. See *The Quest for Certainty,* vol. 4 in *John Dewey: The Later Works* (Carbondale: Southern Illinois University Press, 1988), 204; and *Reconstruction in Philosophy* (Boston: Beacon Press, 1957), 147.

5. See Dewey and James Tufts, *Ethics,* 2nd ed. (New York: Henry Holt, 1932), 307–9; William James, *Pragmatism* (Cambridge, MA: Harvard University Press, 1978), Lecture VI.

6. The inversion of Hamlet stems from Nelson Goodman: *Fact, Fiction, and Forecast* (Indianapolis, IN: Bobbs-Merrill, 1956).

7. The considerations bluntly advanced here are elaborated at much greater length in the last chapter of my *Living with Darwin* (New York: Oxford University Press, 2007).

8. See the closing pages of *Living with Darwin,* as well as “Challenges for Secularism,” in *The Joy of Secularism,* ed. George Levine (Princeton, NJ: Princeton University Press, 2011).

9. Darwin, *Origin of Species,* chap. 9, esp. 310–11.

10. James, *Pragmatism,* 97.

11. This idea, defended by some sociobiologists, is criticized in the final chapter of my *Vaulting Ambition: Sociobiology and the Quest for Human Nature* (Cambridge, MA: The MIT Press, 1985).

12. For a forthright attempt to link human ethical practice to the altruistic tendencies of other primate species, see Frans de Waal, *Primates and Philosophers* (Princeton, NJ: Princeton University Press, 2006).

13. A view well articulated and defended by Marc Hauser: *Moral Minds* (New York: Ecco, 2006).

14. Finally, pragmatic naturalism also rejects the thought that significant advances can be made in understanding ethical issues by undertaking psychological, or neurological, experiments in which subjects are asked to respond to abstract philosophical scenarios. It is unclear what capacities are being fathomed in posing the questions: for questions and concerns that would arise in everyday life are artificially excluded. Moreover, pragmatic naturalism looks for an alternative to current ethical theorizing, rather than for an experimental extension building on available options.

15. The locus classicus for this accusation is Stephen Jay Gould and Richard Lewontin, “The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme,” *Proceedings of the Royal Society,* B, 205, 1979, 581–98. In *Vaulting Ambition,* I argued that storytelling vitiates much work in human sociobiology.

Copyrighted material

PART ONE

An Analytical History

Copyrighted material

CHAPTER 1

The Springs of Sympathy

§3. Psychological Altruism: Basics

At some point in our evolutionary past, before the hominid line split off from the branch that leads to contemporary chimpanzees and bonobos (possibly quite a long time before), our ancestors acquired an ability to live together in small groups mixed in terms of sex and age. That achievement required a capacity for altruism. It also prepared the way for unprecedented forms of cooperation, and ultimately for the enunciation of socially shared norms and the beginnings of ethical practice. Altruism is not the whole story about ethics, but it is an important part of it.[1]

My analytical history of the ethical project thus begins with a hypothesis about the social groups in which the project originated and about the psychological capacities of the members of those groups. Fossil evidence, together with the remains found at hominid and early human sites, reveals that our ancestors lived in bands akin to those in which chimpanzees and bonobos live today: the members were young and old, male and female; the band size was (roughly) 30–70.[2] This chapter argues that, to live in this way, hominids and human beings had to have a capacity for altruism, one contemporary people almost certainly retain. To understand the historical unfolding of ethics we shall need to recognize the intricacies of the notion, as well as the varieties and limitations of hominid/human altruism. The next sections supply the necessary preliminaries.

It is important to distinguish three types of altruism. An organism *A* is *biologically altruistic* toward a beneficiary *B* just in case *A* acts in ways that decrease its own reproductive success and increase the reproductive success of *B.*

Copyrighted material

For a century after Darwin, there was a deep puzzle about how biological altruism is possible. During the past fifty years, however, that puzzle has been solved. Biologically altruistic actions directed toward kin can promote the spread of the underlying genes. Moreover, when organisms interact with one another repeatedly, biological altruism exhibited on some occasions can gain dividends from future reciprocation.[3]

Biological altruism requires no perceptive or cognitive abilities. Even plants can have traits that make them biologically altruistic, for their propensities to form roots or to set seeds can limit individual reproductive success and facilitate the reproduction of neighbors. For animals capable of recognizing the wishes of those around them, however, we can develop a useful behavioral analog of the notion of biological altruism.[4] An animal *A* is *behaviorally altruistic* toward a beneficiary *B* just in case *A* acts in ways that detract from its fulfillment of its own current desires and that promote the perceived wishes of *B*.[5] Behavioral altruists do what they take the animals around them to want. They may act in this way not out of any particular concern for those other animals, but because they think that some of their own wishes will ultimately be well served by doing as they do. Behavioral altruism may be practiced by Machiavellian egoists (and, as we shall eventually see—§11—it can also be practiced by individuals who fall into a category intermediate between egoism and psychological altruism).

Neither biological altruism nor behavioral altruism is of much help in understanding the origins of the ethical project. For our purposes, the significant notion is that of *psychological altruism.* Psychological altruism has everything to do with the intentions of the agent and nothing to do with the spread of genes, or even the successful satisfaction of the wishes of others. Assuming for the moment that there have been human beings who are psychological altruists, the vast majority of them have not known much about heredity, and even those who have were rarely concerned with spreading genes. They acted to promote what they took to be the wishes, or the interests, of other people.[6] Sometimes they succeeded. Yet, even when they did not, their serious efforts to do so qualified them as psychological altruists.

Many people believe psychological altruism does not exist, even that it is impossible. Often they are moved by a very simple line of reasoning: when a person acts in a way that could be appraised as altruistic, he or she acts intentionally; to act intentionally is to identify an outcome one wants and to attempt to realize that outcome; hence, any potential altruist is trying to get what he or she wants; but to strive for what you want is egoistic; consequently, the potential altruist turns out to be an egoist after all. The key to rebutting this argument is to distinguish different kinds of wants and goals. Some of our desires are directed toward ourselves and our own well-being; other desires may be directed toward the welfare of other people. Desires of the former type are the

Copyrighted material

hallmark of egoism, but those of the latter sort are altruistic. So altruists are intentional agents whose effective desires are other-directed.[7]

I shall develop this approach to psychological altruism further, by giving a more detailed account of the character of other-directed desires, and thereby bringing into the open some of the complexities of the concept of altruism. In focusing on desires, I ignore for the moment the fact that there are other psychological attitudes—hopes, aspirations, and particularly emotions—that can be properly characterized as altruistic. Attention to these other types of states will occupy us in the next section. Because of the connection of desires with intentions and actions, altruistic desires have a certain priority. They are thus the topic of the basic account.

The other-directed desires central to the defense of the possibility of altruism are desires that respond to the altruistic agent's recognition of the impact of his or her actions on the situations of others. To be an altruist is to have a particular kind of relational structure in your psychological life—when you come to see that what you do will affect other people, the wants you have, the emotions you feel, the intentions you form, change from what they would have been in the absence of that recognition. Because you see the consequences for others of what you envisage doing, the psychological attitudes you adopt are different. You are moved by the perceived impact on someone else. If your response leads you to *act* altruistically, that is because your *desires* have been affected.[8]

So far, that is still abstract and vague. I shall motivate the underlying idea with a simple and stylized example and then offer a more precise definition.

Imagine that you are hungry and that you enter a room in which some food is spread out on a table. Suppose further that there is nobody in the vicinity who might also be hungry and want all or part of the food. Under these circumstances, you want to eat the food; indeed, you want all of it. If the circumstances were slightly different, however, if there were another hungry person in the room or believed to be in the neighborhood, your desire would be different: now you would prefer the outcome where you share the food with the other person. Here your desire responds to your perception of the needs and wants of someone else, so that you adjust what you might otherwise have wanted to align your desire with the wants you take the other person to have.

This is a start, but it is not sufficient to make you an altruist. For you might have formed the new want when you see that someone else will be affected by what you do, because you saw profitable future opportunities for accommodating this other person. Maybe you envisage a series of occasions on which you and your fellow will find yourselves hungry in food-containing rooms. You see the advantages of not fighting and of not simply having all the food go to the first person who enters. You resolve to share, then, because a future of cooperation will be better from your point of view. For real altruism, the adjustment of desires

Copyrighted material

must not be produced by this kind of self-interested calculation.

I offer a definition of "A acts psychologically altruistically towards *B* in C"—where *A* is the agent, *B* is the beneficiary, and *C* is the context (or set of circumstances). The first notion we need is that of two situations differing from each other in the recognizable consequences for others (people or nonhuman animals). Let us say, then, that two contexts *C* and *C** are *counterparts,* just in case they differ only in that, in one (C*, say) the actions available to *A* have no perceived consequences for *B,* whereas in the other (C) those actions do have perceived consequences for *B. C** will then be the *solitary* counterpart of *C,* and *C* will be the *social* counterpart of *C*.* If *A* forms different desires in *C** from those *A* forms in *C,* the set of desires present in *C** will be A's *solitary* desires (relative to the counterparts *C* and C*). Given these preliminary specifications:

> *A* acts psychologically altruistically with respect to *B* in *C* just in case
>
> (1) *A* acts on the basis of a desire that is different from the desire that would have moved *A* to action in *C**, the solitary counterpart of *C.*
> (2) The desire that moves *A* to action in *C* is more closely aligned with the wants *A* attributes to *B* in *C* than the desire that would have moved *A* to action in *C*.*
> (3) The desire that moves *A* to action in *C* results from A's perception of B's wants in *C.*
> (4) The desire that moves *A* to action in *C* is not caused by A's expectation that the action resulting from it would promote A's solitary desires (with respect to *C* and *C**).

Condition 1 tells us that *A* modifies his or her desires from the way they would otherwise have been, when there is an impact—more accurately, when there is a perceived impact[9]—on the wants of *B.* Condition 2 adds the idea that the desire, and the behavior it directs, is more in harmony with the wants attributed to *B* than it would have been if *B* were unaffected by what was done. (It is possible to modify your desires in response to the perceived wishes of another, but to do so in a way that *diverges* from their perceived wants—that is spite.) Condition 3 explains that the increased harmony comes about because of the perception of *B*'s wants; it is not, say, some caprice on A's part that a different desire comes into play here. Finally, condition 4 denies that the modification is to be understood in terms of A's attempt to promote some desire that would have been present in situations where there was no thought of helping or hurting *B;* this distinguishes *A* from the food sharer who hopes for returns on future occasions when *B* is in the position of disposing of the goods. Condition 4 requires that genuine psychological altruists be different from Machiavellian calculators who aim to satisfy the wants they would have in solitary situations (I shall sometimes refer to condition 4 as the "anti-Machiavelli" condition).

Copyrighted material

Given this account of psychological altruism, it is now possible to characterize *behavioral* altruism more carefully. Behavioral altruists are people who look like psychological altruists. That is, they perform the actions people with psychologically altruistic desires would have been led to perform. In ascribing behavioral altruism, however, we do not suppose any particular psychological explanation of the actions. Perhaps they are indeed the products of psychologically altruistic desires, or perhaps the actions are produced by quite different desires having nothing to do with the satisfaction of the beneficiary—a desire for status, or for feeling oneself in accordance with some socially approved pattern of conduct, or even a self-interested calculation. (We shall explore some possibilities of behavioral altruism later; see §§7, 11.)

The stylized food example allows the introduction of an obvious concept, one that will be important in future discussions, and that further articulates the account of altruistic desires. The altruistic modification of solitary desires can be more or less *intense.* I have spoken—somewhat vaguely—of the altruist as aligning his or her wants with those attributed to the beneficiary.[10] That alignment is often a matter of degree, for example, when there is a continuum of possibilities intermediate between complete egoism (retaining one's solitary desires in the social counterpart) and complete subordination of one's solitary wishes to those one perceives the other to have (where one comes to want exactly what one perceives the other as desiring). In sharing food, this is easily expressed in terms of the mode of division: egoists give nothing, self-abnegating altruists give everything, and in between lie a host of intermediate altruists. One obvious style of altruism is *golden- rule altruism,* distinguished by its equal weighing of the solitary desires and those attributed to the beneficiary.

Inspired by the food example, we can undertake a simple way of representing the intensity of psychological altruism, one that will be useful in some (but by no means in all) instances. Suppose that people's desires can be represented by (real) numbers that correspond to how much they value a given outcome. If one result, eating all the food, say, is worth 10 to me, and another, eating half the food, is worth 7, then I prefer eating everything to eating half, but I also prefer an assured outcome in which I receive half to the state of being awarded all or nothing dependent on the flip of a fair coin. (For, in the latter case, my expected return is measured by 5—half of 10 plus half of 0—which is less than 7.)

When you are in the picture, I also take into account the values you attribute to various outcomes. My social desire could be represented as a weighted average of the values represented in my solitary desires and those I take to measure your solitary desires. Thus, the numbers assigned in my social desires would be given by the simple equation:

$$v_{Soc} = w_{Ego} v_{Sol} + w_{Alt} v_{Ben}$$

Copyrighted material

where v_{Soc} measures my social desires, v_{Sol} my solitary desires, v_{Ben} the measurements of desire I attribute to the beneficiary (you), and w_{Ego} and w_{Alt} the weights given to my solitary desires and my attributions of desire values to you (so that $w_{Ego} + w_{Alt} = 1$). The intensity of my altruism is represented by the size of w_{Alt}—and hence inversely by the size of w_{Ego}; if $w_{Ego} = 1$ ($w_{Alt} = 0$), then I am, at least with respect to you on this occasion, a psychological egoist; if $w_{Alt} = 1$, then I am a self-abnegating altruist; if $w_{Alt} = 0.5$ ($= w_{Ego}$), then I am a golden-rule altruist.

We should not assume that all types of altruistic alignment with the wishes of others can be conceived in this very simple way. Cases of sharing show that a simple approach sometimes works, and the simple expression of social wants as weighted averages will be useful in explaining and illustrating some of the ideas of later sections.

§4. The Varieties of Altruistic Reactions

As already recognized, altruism is not always about the modification of desire, though we are often reasonably suspicious about alleged examples of altruism that do not change desires in ways leading to action: it is not enough simply to "feel another's pain." We can be moved to share the hopes of others, to modify our own long-term intentions and aspirations to accommodate what we see them as striving for, and, most important, we can feel different emotions because of our awareness either of what they feel or of the situations in which they find themselves. For some kinds of psychological states, hopes and long-term intentions, for example, accounts of altruistic versions of these states can be generated straightforwardly in parallel to the treatment of the previous section. Emotions, however, deserve special consideration, both because they are frequently components of the psychological attitudes with which we shall be concerned, and because they involve types of reactions more broadly shared among animals than the psychological states on which I have so far concentrated.

Altruistic emotional responses to others *might* be—and probably often are—mediated by perception and cognition. We see that another person is suffering—or jubilant, for altruistic emotions are not always dark—and our own emotional state changes to align itself more closely with that attributed to the other. Or, in a different mode of altruistic response, we understand the situation in which another person is placed, and our emotional state changes to take on, to some extent, the feeling(s) we would have if placed in that state.[11] When people, or other animals, have dispositions to modify their emotional states in light of their understanding of the feelings or the predicaments of others, we can treat emotional altruism just as §3 analyzed altruistic desire. The emotional altruist feels one thing in the solitary counterpart and feels differently in the social

Copyrighted material

counterpart; the emotion in the social counterpart is more closely aligned with that attributed to the other (or more closely aligned with the emotion the altruist supposes he or she would feel if placed in the other's shoes), and the alignment comes about because of the recognition of the other's feelings (or of the other's situation); finally, it is not caused by any background solitary emotion or solitary desire. Now, whereas in the understanding of altruistic desire this last condition responds to a genuine worry—for we readily think people can form ostensibly other-directed desires on the basis of selfish calculations (I can want to share with you because I think it will be good for me in the long run)—the anti-Machiavelli condition seems odd and gratuitous in the emotional case. It is natural to think, and it may even be true, that self-directed psychological states simply have no power to generate *emotions* toward others, that our emotional life is not under that sort of control. Emotional responses, one may suppose, are caused by processes more direct and automatic than the perceptions and cognitions figuring in my analyses. Consequently, an account of emotional altruism parallel to the analysis of altruistic desire will be at least incomplete, and perhaps even radically misguided.

This is a serious challenge. To meet it, we shall have to consider, if only briefly, the character of emotions. Without taking sides in unresolved controversies, I shall argue that some kinds of emotional response can be understood along the lines just sketched, while others cannot. An account of more basic altruistic emotional reactions, or "affective states," as I shall call them, provides a valuable supplement to the approach to psychological altruism begun in the previous section.

Emotions involve changes in our physiology, and some students of emotion have identified the emotion with the alteration in physiological state. Others propose that there are important distinctions among emotions that cannot be recognized without supposing those who feel the emotions to have particular beliefs, desires, and intentions: specific forms of awareness are required for guilt and shame, for resentment and indignation, and for certain kinds of contentment and anger. A natural way of responding to the findings of neuroscientists, psychologists, and anthropologists is to suppose that many emotions are complex entities, perhaps processes in which particutar types of physiological conditions are accompanied by special kinds of cognitive and volitional states. When someone resents the insensitive remarks made by another, he or she undergoes a physiological response connected to judgments about what has been said and desires about what will happen next. The causal details of these connections are matters of speculation, but, even in advance of knowing them, we can reject an approach to emotions that would leave out either the physiological or the cognitive/volitional features.[12]

Yet there may be emotional states, felt by nonhuman animals and by human

Copyrighted material

Machiavellian manipulation of our emotional lives is beyond our powers, and, if that is indeed so, this requirement is redundant.

The analysis just given preserves a fundamental feature of my original characterization of psychological altruism (§3): altruists have a particular type of relational structure in their psychological lives—when others are around, the altruist's desires, hopes, intentions, and emotions are different from what they would otherwise have been, closer in some way to those of the others, and the difference is produced by some sort of response to those others, not by something enclosed within the self (calculations of future benefit, for example). What the more complex approach to altruistic emotions adds is the possibility that the generation of the response might involve some precognitive mechanism.

It is easy to overinterpret this last point. One might suppose that affective states are always generated by some mechanism that does not involve cognition—but, not only do I see no basis for holding so sweeping a generalization, but it also seems belied by the fact that affective reactions are often founded in complex and explicit understanding (when I see pictures of Jewish refugee children being greeted at English ports by policemen and willing foster parents, I feel a complex mixture of emotions, surely involving affective states, but these states are clearly dependent on my conscious understanding of what the photographs display). The causal relations among affective and cognitive states may be quite various, and, while we await definitive accounts of them, it is well to suspend judgment and to be open to many possibilities.

Nor should we suppose that noncognitive mechanisms are inevitably involved in whatever altruistic responses occur in nonhuman animals. Although questions about the extent of animal abilities to recognize the wishes and thoughts of their conspecifics are much debated, there is no reason to take an advance stand on these issues.[17] I shall later defend the thesis that some of our evolutionary cousins have altruistic desires (in the sense of §3; see §7) and that similar capacities were shared by our hominid ancestors.

§5. Some Dimensions of Altruism

One further aspect of psychological altruism needs to be emphasized before we have all the tools required for probing the hominid preethical state. On the account of the last sections, there are many varieties of altruism. Or, to use a suggestive metaphor, altruism is a multidimensional notion. For animals capable of psychological altruism, each individual occupies a particutar place in a multidimensional space where brute (non-Machiavellian) egoism is represented by a single plane, and the various forms of altruism range over the entire rest of the space.[18]

An animal's *altruism profile* (where he or she is located in altruism space) is

Copyrighted material

determined by five factors: the *intensity* of the animal's responses to others, the *range* of those to whom the animal is prepared to make an altruistic response, the *scope* of contexts in which the animal is disposed to respond, the animal's *discernment* in appreciating the consequences for others, and the animal's *empathetic skill* in identifying the desires others have or the predicaments in which they find themselves. Non-Machiavellian egoists never respond to anyone else in any context: for the dimensions of intensity, range, and scope they score 0, 0, and 0; their discernment and empathetic skill can be as you please, for these are never called into play.

Altruists are not like that. They modify their desires and emotions to align them with the perceived desires and (perceived or actual) emotions of at least some others in at least some contexts. As §3 already proposed, their responses may be more or less intense. With respect to altruistic desires, an altruist may give more or less weight to the perceived desire of the beneficiary. My treatment of the stylized example in terms of weighted averaging provides a clear paradigm for intensity—the intensity of altruism is represented by how much of the food you are willing to relinquish. If

$$v_{Soc} = w_{Ego} v_{Sol} + w_{Alt} v_{Ben}$$

egoists set w_{Ego} at 1 and w_{Alt} at 0. People for whom $w_{Ego} = 1 - \varepsilon$, where ε is tiny, are altruists in a very modest sense: they will act to advance the wishes of others only when the perceived benefits to others are enormous compared to the forfeits for themselves—they may suffer the scratching of their finger in order to avoid the destruction of the world, but refuse larger sacrifices. People for whom $w_{Alt} = 1$, by contrast, are completely self-abnegating. They abandon their own solitary desires entirely, taking on the wishes they attribute to the beneficiary. In between, we find golden-rule altruists, for whom $w_{Alt} = {}^{1}/_{2}$, who treat the perceived wishes of the other exactly as they do their own solitary desires.

Even when averaging is not appropriate for representing altruistic desires, there will often be a comparable notion of the degree to which one has accommodated the perceived wishes of the other. Moreover, with respect to altruistic emotions there is surely a similar concept. Notoriously, we can be relatively unsympathetic, even with those who are dearest to us, when we are preoccupied or distracted. At other times, we enter fully into the feelings of friends and loved ones, even of strangers. It is not obvious how to delineate the notion of intensity in the emotional case as precisely as the food-sharing example allows, but the varying intensity of altruism in emotional responses is uncontroversial. Notice, however, that it should not be confused with the intensity of *emotion:* intensity depends on the degree of *alignment* with the other's feelings (or with the feeling one would have had in the other's situation), not with

Copyrighted material

the force of what one feels.

Most altruists, indeed probably all, lack a fixed intensity of response, applying with respect to all potential beneficiaries and all contexts. There are many people to whom we would rarely make an altruistic response: these people effectively fall outside the range of our altruism. Even with respect to those to whom we are disposed to respond, there are many contexts in which we do not take their perceived wishes or their feelings into consideration (or into our own minds). For many, perhaps, we are prepared to offer limited forms of aid and support; for a few, we are willing to sacrifice everything. Often our altruistic responses to some are colored by indifference to others: parents who make sacrifices to help their children obtain things the children passionately want frequently do not take into account the wishes of other children (or the altruistic desires of the parents of the other children).

Someone's altruism profile typically shows a relatively small number of people to whom the focal individual responds, frequently with significant intensity, across a wide set of contexts. The beneficiaries lie at the center of the range of altruism for the focal individual, and the scope for these beneficiaries is wide. As we consider other potential beneficiaries more distant from the center, the scope narrows (there are fewer contexts in which the more peripheral people elicit an altruistic reaction) and the intensity falls off, until we encounter people to whom the focal individual makes no altruistic response at all. Henceforth, I shall conceive of the range of A's altruism in terms of the metaphor of center and periphery: the center is the select set of potential beneficiaries for whom A's response is relatively intense across a relatively wide scope of contexts; at the periphery, the intensity of the response and the scope of contexts narrow and vanish.

Someone's character as an altruist is not fixed simply by the factors so far considered—intensity, range, and scope—because there are also significant cognitive dimensions to altruism. *A* may make no response in a particular context through failure to understand the consequences for *B;* perhaps *A* does not differentiate the social from the solitary counterpart. Often this is an excusable feature of our fallibility, for the impact on the lives of others may be subtle; we may just not see that following some habitual practice—buying at the most attractive price, or investing in promising stocks—has deleterious consequences for people about whose welfare we care. Evidently, however, acuity with respect to consequences comes in grades, and we admire those who appreciate the intricate ways in which others can be affected, while blaming those who "ought to have seen" the damage they cause.

Similarly, there are degrees to which people are good at gauging the desires of others. Almost everyone is familiar with the well-intentioned person who tries to advance the projects of an intended beneficiary but who is hopelessly misguided

Copyrighted material

about what the beneficiary wants: almost everyone has had a friend or relative who persists in giving presents no longer appropriate for the recipient's age or conditions of life. It would be hard, I think, to declare that people who attribute the wrong desires to their beneficiaries, or who overlook consequences for those whom they intend to benefit, are not acting altruistically when they carry out their variously misguided plans—their intentions are, after all, directed toward doing good for others—but their altruism needs to be differentiated from that of their more acute fellows. Hence I add two cognitive dimensions, one representing *A*'s skill in understanding the nature of a social counterpart to a solitary context, and one assessing *A*'s ability to empathize with *B,* to ascribe desires *B* actually possesses.

A simple reaction to the prospect of human egoism is to propose that people living in community with one another—or even all people—should be altruistic; some even take the second commandment of the New Testament to constitute a complete ethical system. Recognizing the dimensions of altruism undermines that thought. There is no *single* way to be an altruist, and, consequently, the commendation of altruism must be given more specific content. What kind of altruist should we urge someone to be? Moreover, is it right to suppose that the best state of the community (or the entire species) is achieved by having each member (each person) manifest the same altruism profile? You might think the questions have straightforward answers. Along the cognitive dimensions, accuracy is always preferable: ideally people should be aware of the potential impact for others and should understand what others want. For issues of intensity, range, and scope, we ought to aim at golden-rule altruism with respect to all people across all contexts.

The demand for accuracy on the cognitive dimensions is more plausible but still not uncontroversial. Debate about the second part of the proposal arises in obvious ways. It might be valuable for people to develop strong ties with some others—the range of human altruism should have a definite center; from Freud's worries about the "thinning out" of our libido in the development of civilization to familiar philosophical examples about parents who wonder whether they should save the drowning child who is closer, when their own drowning child is farther out and harder to rescue, a spectrum of troublesome cases arouses suspicion about completely impartial altruism.[19] Moreover, in a world with finite resources, the desires of others often conflict. If *A* accurately perceives that both B_1 and B_2 want some indivisible good, it should not be automatic that A's desire should be formed by treating B_1 and B_2 symmetrically. (We may, for example, want *A* to respond to aspects of the history of the situation, including what B_1 and B_2 have previously done.) None of this is to deny that there may be a level at which we want altruism profiles to respond impartially to others, but merely to insist that the impartiality we want cannot be adequately captured as golden-rule

Copyrighted material

altruism toward all people in all contexts.

Further complexities of the notion of psychological altruism will occupy us later. For the present, however, we have enough to begin charting the history of our ethical practices, by understanding how the most basic forms of psychological altruism could have evolved, and how they formed an important part of the social environment in which the ethical project began.

§6. Maternal Concern

Before our human ancestors invented ethics, they had a capacity for psychological altruism. This thesis might be disputed in any of several ways, but the one of immediate concern recapitulates the skepticism about altruism mentioned earlier (§3). Armed with the elements of an account of psychological altruism, the first task is to decide if any such capacity exists, and if it could plausibly be attributed to contemporary human beings, our hominid ancestors, and our evolutionary cousins. Let us begin with the most straightforward case.

Behavior directed toward the survival of young is quite widespread in the animal kingdom, found, for example, among birds as well as mammals. With respect to some types of animals, the hypothesis that this behavior is directed by altruistic desires appears extravagant, for it presupposes the propriety of attributing wants and intentions apparently beyond the cognitive capacities of the pertinent organisms. Nevertheless, we might view the animals as driven by altruistic emotions (or primitive versions of them), generated through the operation of automatic neural mechanisms. Among primates, however, particularly those closest to our own species, our evolutionary cousins the great apes, there is considerable evidence for the ability to have desires and to recognize the desires of others.[20] For the sake of concreteness, we can think of psychologically altruistic dispositions to care for the young as emerging in apelike ancestors of *Homo sapiens,* but it is eminently possible that they evolved much further back in our primate (or even mammalian) past.

Even those who share the orthodox primatological views about the cognitive sophistication of our evolutionary cousins may be skeptical of any hypothesis that parental care is sometimes directed by altruistic desires, in the sense I have explicated in §3. They may wonder, for example, whether any dispositions of this kind could evolve under Darwinian natural selection, or whether the apparently altruistic behavior is really the product of some quite different mechanism. Perhaps the animals are really calculating how to achieve future benefits, violating condition 4 of my account, the anti-Machiavelli condition. Many primatologists take the social organization of primate life to reveal "Machiavellian intelligence," and evolutionary psychologists often propose that increased cognitive powers in hominids reflect the need to manipulate others and

Copyrighted material

psychological states, the glows and pangs, as intertwined with matters of conscience, a point that will be important later). Hence, the skeptic proposes, mothers do the impressive things they do because they want to avoid a future of terrible self-reproach and self-torment.

At least two things cast doubt on the skeptical hypothesis. First, the fact that the mother envisages the future of self-reproach testifies to the motivating power of her recognition of the child's wishes (or, in this instance, more likely the child's *interests*—see §21). It is often preposterous to suppose a mother will reproach herself because she is concerned with attitudes in her society—frequently, those around her would praise her for doing far less than she does, constantly reassure her that she has done more than anyone could possibly expect, and so forth. The drive to pursue every possible avenue comes from within, and it could not be abated by any amount of well-intentioned commendation and comfort. If she fails, the mother will suffer, no matter how much she has done and no matter what others say, and the suffering will stem from her deep desire that the child survive and flourish. So, at least, we might initially believe. On the skeptical hypothesis, however, that desire must be denied. Instead, the mother must be viewed as being able to feel altruistic emotions in response to her child. This ability, and the emotions to which it gives rise, does *not* express itself in a desire for the child's well-being. Instead, the ability leads her to fear a particular type of future state, and the fear replaces the denied desire as the driver of her conduct. We have no grounds for accepting this speculative psychology.

A final—fanciful—way to underscore the point: Our world hardly abounds with clever spirits, willing to offer bargains. Yet the mother might have a particular disposition to react to temptations. Imagine that she were visited by a Mephistophelean figure with a straightforward proposal: "I can give you a pill to ensure you will not feel any guilt should things go badly for your child. The pill will wipe away both the pangs of conscience—you will reflect on your efforts and feel you did your best—and any memory of this conversation and the decision to accept the pill. The downside is truly tiny. The probability of your saving your child if you don't take the pill is p; the probability if you do take the pill is $p - \varepsilon$ (where ε is really infinitesimal). Surely the reasonable thing is to accept?" With respect to many actual mothers, we have no doubt about how they would respond—by telling Mephisto to get lost. They view their future psychological comfort as trivial compared with the value of saving the child—any diminution of the probability of success is a loss for which future amnesia cannot compensate. Their assessment of relative value expresses just the desire for the child's well-being the skeptic attempts to deny.

Psychological altruism is real, it is exemplified in maternal concern, and it originally evolved through the most fundamental type of kin selection. Because it

Copyrighted material

image
not
available

Index

abortion. *See* bioethics
agriculture, 305, 375–376, 389
alienation, 35, 168
altruism: psychological, 5, 17, 19–20, 23–24, 26, 30–31, 35, 39, 42–51, 53, 56–58, 63–67, 71–72, 82–83, 86–88, 100, 105, 111, 131–136, 226–229, 232, 340, 344, 368, 394, 40; failures of, 6–7, 9, 73–74, 86, 91, 93, 98, 100, 103–104, 107, 111–112, 124, 126, 128, 131–132, 135, 137, 143, 150–151, 153, 217, 222–231, 233–240, 242–244, 247–248, 250, 259, 262, 273–274, 276, 279, 294–295, 297–299, 303–305, 307, 309, 311, 327, 339–340, 346, 349–351, 355, 360, 365, 368, 380; biological, 18–19, 37, 49; behavioral, 19, 23, 45–47, 56, 75–76, 79, 82, 86, 94, 131–132, 226, 231–232; vs. egoism, 19–20, 23–24, 31, 34–36, 40, 44–45, 105, 135–136, 230, 274–276; golden-rule, 24, 32, 34, 37, 56, 144, 222–223, 346; physiological basis, 26–27, 39, 47–49, 199, 309; dimensions of, 31–34, 44, 69, 95, 144, 344–345; reciprocal, 50–58, 64–65, 135; quasi, 56; higher-order, 135–137, 217, 232, 291, 327, 339, 345, 371; and asceticism, 227, 289
animals, ethical status of, 180–183, 306–310
anthropocentrism, 66, 310
anthropology, 3, 5, 11, 18, 26, 68, 94, 96, 114, 116, 138, 156, 186
Aquinas, Thomas, 231
archeology, 5, 11, 116
aristocracy, 141, 174, 180, 328
Aristotle, 154, 247, 326
Arrow, Kenneth, 312
atomic physics, 182–184, 272
atomism in ethics, 210–212, 363
autonomy, 207, 255, 325
awe, 81, 84, 92, 113, 131, 201, 226, 231
Axelrod, Robert, 52
Ayer, A.J., 203, 280

Babylon, 117–119, 145, 216
Barkow, James, 36
Batali, John, 56
Beauvoir, Simone de, 147
Bentham, Jeremy, 289, 293, 314

Copyrighted material

Berlin, Isaiah, 249
Bible (Judeo-Christian): Scriptures, 4, 154, 158, 160–161, 165; Deuteronomy, Book of, 119; Exodus, Book of, 119; Genesis, Book of, 130; Jacob, 130; Joseph, 130, 154; Moses, 141, 144, 174; and the practice of slavery, 143, 154–155; Acts, Book of, 144; Luke, Book of, 144; Matthew, Book of, 144; Gospels, 144, 155, 166, 214, 236, 250; Pentateuch, 154; Ham, 154–155, 160; Leviticus, Book of, 167
bioethics, 401–408
biology: ontogeny vs. phylogeny, 330–336; cell biology, 402–403
Boehm, Christoph, 18, 96, 116
Bohr, Niels, 272
boundaries, 120, 190, 214, 303, 306, 309
Boyd, Robert, 53, 87, 109, 112
Brownian motion, 182
brutality, 31, 47, 60, 118, 143, 161, 166, 225, 229, 278, 307, 395

calculus, 234, 293
callousness, 46, 393
Canaan, 154, 167
Carcopino, Jerome, 174
Cardozo, Benjamin, 120
caregiving practices, 38, 152, 385–386, 389–390
causation: between affective and emotional states, 27–30, 77–78, 80–84; of ethical regression, 177–178; in functionalism, 219–220; role in evaluating ethical facts, 291–294; as a way of demarcating ethical communities, 302–305
Cheney, Dorothy, 36, 58
children, and altruistic behavior, 40–42, 132
citizenship, 106, 128, 140, 145–147, 149, 162, 165–166, 174, 177, 212, 375, 378, 382, 390
civilians, 392–393
cloning. *See* bioethics
coalition game, 57–66, 88, 96–97, 228
cognition: role mediating ethical judgments, 25–30, 37–39, 75–80, 198–200; in animals, 39–45, 50, 56, 78–80
Coke, Sir Edward, 146–147
collective: construction of ethics, 2, 191–192, 204–205; use of resources, 123–124, 318–323; acceptance of an ethical code, 234–237, 256
colonial, 109, 154
colonies, 43, 69–70, 146, 153–156, 352
combat, 142, 175, 249, 392–393
commitment, 74–75, 79, 81, 96, 121, 151, 195, 200, 211–212, 233–234, 247, 256, 280, 324, 331, 352, 373, 380, 388, 390–391
communal resources, 125, 129–131, 242, 397
compassion, 140, 143, 160, 174, 214
competition: for resources, 60–62, 248, 255, 350, 375, 390; cultural, 107–115, 128–131, 169–170, 259
compromise, 241, 244, 260, 272, 300, 336
computer simulation, 52, 56, 63
conscience, 40–41, 92–93, 112
consensus building, 175, 334–335, 340–344, 348–360
consequentialism, dynamic, 288–294, 311–313
constructivism about truth, 190–192, 201–203, 245
contractarianism, 273, 279, 344
conversation. *See* discussion; mutual engagement
cooperation: discriminating, 12, 55–57, 88–89, 174; noncooperation, 52–53, 85, 89; nondiscriminating, 55
Crusades (Christian), 144
Cummins, Robert, 219

Darwin, Charles: Darwinian evolution, 12, 36, 90, 107–109, 111–112, 114, 123, 138, 213, 219–220, 237–238;

Copyrighted material

coevolution of genes and culture, 108–109; social Darwinism, 255–256, 258
Dawkins, Richard, 18, 326
deliberation about norms, 111–115, 130–135, 143, 259–261, 286, 308–313, 331–332, 340–345, 372–384
Dennett, Daniel, 114
deontology, 289, 294, 298, 300, 302, 314, 325, 328
Descartes, Rene, 205
de Waal, Frans, 9, 17, 30, 43–44, 50, 57, 59–60, 66, 69–70, 87, 189
Dewey, John, 3, 114, 120, 179, 207, 251, 264, 266, 304–305, 327, 342, 347
Dickens, Charles, 353
discussion: ideal, 96–100, 104–108, 111–112, 339–341; among groups, 270–272, 281–282, 285–288, 296–298, 302–329, 331–338, 342–354, 356–360, 363–369, 394–396; contemporary, 370–373, 377–378, 383–385, 388–391, 394–396, 402–404, 406–410
disobedience, 81–82, 93, 111, 279
dispositions, 36–38, 63–65, 67, 69–74, 78, 87–88, 91, 93, 98–100, 102, 108, 131, 145, 161, 179, 183, 185, 191, 198, 204, 228, 232, 263, 278–279, 303, 335–336, 340
divine commander, 111–115, 119–121, 127–128, 163, 166–170, 207–208, 230–231, 241, 260, 289, 294–298
divinity, 2, 4, 8, 113, 115, 119, 121, 128, 131, 148–150, 163, 165–170, 207–208, 230–231, 260, 285, 289, 298, 311, 313, 328, 372–373, 395
division of labor, 117–118, 122–131, 151n29, 239–241, 243–245, 294–295, 327–329, 350–352
dominance, 59, 69–71, 228
Dummett, Michael, 192

early friendships, 64–66
ecology, 59, 319
economics, 23, 50, 65, 85, 312, 391
education, 146, 149, 193–197, 317–319, 324, 359, 371, 375–377
egalitarian commitments, 96–97, 106, 115–116, 122–123, 127, 217, 248, 297, 321, 324–325, 351–352, 359, 364–366, 373, 377, 389–390, 393–397
egoism, 19–20, 23–24, 31–32, 34, 36, 40, 44–45, 105, 136, 230, 274–276
Egypt, 11, 113, 118–119, 130, 216–217
elitism, 314–315, 328
emotions, 36–37, 49, 76–85, 93–95, 98–102, 198–204; nature of, 25–32, 78–79, 198–201
empathy, 31, 34, 144, 250–251, 345, 371
enforcer (divine), 111, 113–115, 119, 127, 131, 169, 217, 226, 230–231, 241, 260, 274, 294, 297–298
environmentalism, 305, 310, 321
equality, 11, 85, 96, 115, 157, 178, 248, 285, 287, 289, 291, 293–295, 297, 299, 301, 303, 305, 307, 309, 311, 313, 315, 317–319, 321, 323–325, 327, 329, 339, 341, 364, 380, 385, 387, 389, 396
ethical point of view, 80–85, 106, 231–232
eugenics. *See* bioethics
evolution. *See* Darwin, Charles
experiments: psychological, 45–46, 85–86, 101–102; of living, 104–137, 150–153, 213, 217, 226–227, 230, 234–235, 243, 247, 295, 302, 315, 356–360
expertise (ethical), 8, 285–287
explanations, 12–13; how possibly, 60–74, 78–80, 87–94, 99–103, 104–107, 111–115, 122–131, 135–137, 216n8, 223–229, 236–237, 239–242; 242–245, 307–308; how actually, 105–106, 117–119, 173–178, 213–229, 294–296, 374–381; evolutionary, 107–110, 114, 213–214, 237–240

Faustian bargains, 41
Fehr, Ernst, 85, 101
feminism, 146–148, 152, 390
Foot, Philippa, 247, 278
forgiveness, 120, 144–145
Fossil record, 5, 18
Foucault, Michel, 296

Copyrighted material

Freud, Sigmund, 34–35, 92, 95, 163
friendship, early friendships, 64–66
function, etiological approach to, 218, 220
functional conflict, 241–245, 247–249, 252, 260–262, 281–282, 286–288, 296, 333–334, 337–353, 355–356, 360–369. *See also* progress; pluralism (ethical)

Galileo, 148, 181, 189
generosity, 121, 136, 141
genetic determinism, 27, 50, 114, 406
genetic testing, 380, 402–406
Gibbard, Alan, 96, 203–204
Gilgamesh, 118
Goldman, Alvin, 209
Goodall, Jane, 42
Goodman, Nelson, 3, 258
Gould, Stephen, 10
government: autocratic, 148; democratic, 146, 188–189, 265

habits, 97, 331–332, 381–385
Hacking, Ian, 183
Hamilton, William, 18, 49
Hamlet, 3, 92, 94
Hammurabi, code of, 118–119
Harman, Gilbert, 139, 181
health: effect on shaping ethical systems, 91n28, 101, 112, 137, 150–151, 316; as an ethical issue, 296–297, 304–305, 318, 374–381
hedonism, 313, 328, 363
Hitler, 393
Hobbes: Leviathan, 60, 95; state of nature, 154
Hofstadter, Richard, 256
holism, 292–293, 314, 316, 363
Homer: Iliad, 141–142; Odyssey, 141–142, 174; as a model of ancient ethics, 141–142; contrast with Solon, 142–143
hominid, 5, 9, 17–18, 31, 35–36, 47, 59, 61, 64, 66–69, 71, 73–74, 80, 87–88, 90, 92, 96–98, 103, 116–117, 194, 222, 237–238, 251, 275, 279, 297, 304, 337, 367, 393–394
honesty, 246–247, 261–262, 266, 271, 299–302, 358–359, 391–392
Hume, David: Enquiry Concerning the Principles of Morals, 99, 213, 274, 308; problem of induction, 255, 257–258
Hüyük, Çatal, 116–117

institutions (evolution of), 120–122, 128, 137, 146–148, 248–249, 282n32, 295–296, 372–374, 382–385; private property, 120–131, 396–401; marriage, 137, 150–151. *See also* slavery
internalization of ethical norms, 87, 92–96, 106
intuition, 179–181, 183, 205–206, 334–336
irrationality, 272–273, 275

James, William, 3, 92, 176, 185, 384
Jefferson, Thomas, 157
Jex-Blake, Sophie, 147
Jones, Matthew, 213
justice, 73, 105, 119, 139, 141, 201, 251, 258, 274, 312, 334

Kant, Immanuel: Groundwork of the Metaphysics of Morals, 83, 166, 201; Categorical Imperative, 206, 245
Kierkegaard, Søren, 122

Copyrighted material

Kinsey, Alfred, 164
kinship, 45, 71, 120, 223, 259
Kuhn, Thomas, 2, 147, 180, 189, 333, 349
!Kung San, 96, 115–116

Lecky, W.E.H., 143–144, 174, 214
Levi, Isaac, 101, 245, 393
Lewis, David, 168
Lewontin, Richard, 10
lex talionis, 140, 186, 194, 197, 202, 215, 234–235
liberation, 2, 13, 29, 81, 93, 111, 163, 173, 183, 279, 325, 337, 367
liberty, 107, 147, 150, 154, 164, 264, 292–293, 305, 315, 327
Lipit-Ishtar code, 118–119
Locke, John, 397
Lowance, M., 154–155, 157–158
Lumsden, Charles, 108

machiavellianism, 19, 23, 26, 30–31, 36–40, 42, 44, 50, 53, 56, 65, 69–70, 75–77, 79, 82, 86, 136
MacKinnon, Catherine, 146–147
malthusian population control, 320, 395
Marx, Karl, 296
Mathematics: comparison to ethics, 192–193, 203, 205–206; foundationalism, 205
McDowell, John, 198
Mendel, Gregor, 110
Mesopotamian civilization, 11, 113, 118–119, 144, 217, 234
Metaphysics, 83, 166, 201, 263
Mill, John Stuart, 107, 129, 143, 147–152, 164, 263–264, 289, 292–293, 302, 305, 314–316, 325, 327, 351, 399–400
mirroring: primitive, 346–347, 350; extended, 347, 350, 354–357, 364
mutual engagement, 136, 340–349, 356–365, 368, 370–373, 376–381, 384

Nagel, Thomas, 85, 335
naturalism: pragmatic naturalism (general discussion), 3–13, 225, 246, 256–258, 268–270, 281–282, 317–324, 362–369; nonnaturalism, 8, 269–271, 273, 276, 279; counternaturalistism, 281
Nazi, 139, 177, 255
Neurath, Otto, 279–280
neuroscience, mirror neurons, 28–30, 43, 346
Newton, 205, 213
Nietzsche, Fredrich, 92, 95, 143, 211, 276–278, 282, 360–361
nihilism, 269–270
noncognitivism, 30, 201–205
normative: guidance, 67–69, 74–87, 90–103, 111–112, 123–125, 131–137, 221–223, 230–234, 409–410; conclusions, 191–192, 253–259, 338
Nozick, Robert, 397

obedience, 84, 99, 108–109, 111, 160, 167, 231, 264, 289
objectivity in ethics, 139–140, 209–212, 243–245, 267, 312

Parfit, Derek, 322
paternalism, 133–135, 149–153, 227, 301
peacemaking practices, 57, 69–70, 73, 87, 116, 221–222
Peirce, C.S., 176, 210, 246
philosophy (role in ethical project), 8–9, 286–287, 370–373, 410
Plato: Euthyphro, 128, 166; Republic, 128, 274, 367
pluralism (ethical), 131, 210–213, 245–249, 281–282, 306, 318, 337–338, 355, 367–368,

Copyrighted material